

Nominal Projections and Pronoun Types

Paul Melchin

Supervisor: Prof. R. Truswell

Major Research Paper submitted in partial fulfillment of the
requirements for the degree of Master of Arts in Linguistics.

University of Ottawa
August 2013

Contents

1	Introduction	1
1.1	Methodological Preliminaries	2
2	The Models	4
2.1	Déchaine and Wiltschko (2002)	4
2.2	Zamparelli (2000)	6
2.3	Similarities and Differences	7
3	Discussion	8
3.1	Ambiguities and Apposition	9
3.2	The Unified Model	12
3.3	Kratzer (2009) on Binding	16
4	Japanese Third Person Pronouns	19
4.1	Déchaine and Wiltschko (2002) on <i>Kare</i>	19
4.2	Yashima (2013) on <i>Kare</i>	21
4.2.1	Binding <i>kare</i>	21
4.2.2	Epithets	22
4.2.3	<i>Kare</i> as an epithet	23
4.2.4	DP and NP epithets	27
4.3	ϕ P–NP Ambiguity	29
4.4	DP–NP Ambiguity	30
4.5	Pronouns in D-less Languages	31
5	A Feature Geometry for Pronouns	36
5.1	Cardinaletti and Starke (1994)	37
5.1.1	The deficiencies	37
5.1.2	The model	42
5.1.3	Discussion	43
5.2	Previous Geometries	46
5.2.1	Harley and Ritter (2002)	46
5.2.2	Cowper and Hall (2009)	47
5.3	My Proposal	49
6	Conclusion	50

Abstract

The goal of this paper is to formulate a pronoun typology based on the syntactic structure of the noun phrase. I begin by reviewing two previous proposals with the same goal, finding them to be mutually incompatible. In Section 3 I propose a unified model, based mostly on the model in Zamparelli (2000), in which syntactic categories are defined in terms of their semantic type. The binding properties upon which Déchaine and Wiltschko (2002) based their analysis are relegated to an independent mechanism, namely Kratzer's (2009) binding theory. Next I deal with some problematic data from Japanese, in which there appear to be some unwanted, systematic categorical ambiguities; following Chierchia (1998), I show that the pronouns in question are not categorically ambiguous, and in fact such properties are expected given the settings of the Nominal Mapping Parameter of the language. Finally, I introduce the model of Cardinaletti and Starke (1994), showing that the properties they discuss are best incorporated into the unified model in the form of a feature geometry. In this way, I bring together a number of previously incompatible pronominal typologies into a single, unified theory, complete with a proposal for the structure of nominal projections.¹

1 Introduction

Pronouns are one of those rare items in natural language which are relevant to almost every subfield of linguistic research. They provide insights into such various areas of study as binding theory, reference, nominal phrase structure, morphology, and prosody, among others. It is clear that they differ from non-pronominal noun phrases in at least some key ways, and it has long been widely acknowledged that there are different types of pronouns, both within and between languages. Since at least the time of Cardinaletti and Starke (1994) there has been a tradition of describing these different pronoun types in terms of difference in the level of syntactic structure that is projected. A number of authors (e.g. Cardinaletti and Starke 1994; Zamparelli 2000; Déchaine and Wiltschko 2002) have made proposals for how this should be accomplished; the problem is, their proposals are contradictory, positing different structures on the basis of different sets of features. While each analysis contributes worthwhile insights into the study of pronouns and nominal projections in general, they cannot all be reconciled in a single elegant syntactic theory, at least not in the current

¹I could not have finished a project of this size without help from a whole lot of people. In particular I would like to thank my supervisor, Rob Truswell, for his support, invaluable help, and hours of discussion (despite the endless distractions of being a new father); my friends in the MA program for their support and discussion, especially Brendan McDonald for his help with epithets and binding, and Meng Yang for helping me out with the prosodic hierarchy; my informants, Meng Yang, Chung-hye Han, Ayşegül Kutlu, Yukiko Yoshizumi, and Mika Nagai, for their data and judgements; members of the University of Ottawa Syntax-Semantics Research Group for their comments on an earlier version of this project; and finally my parents, Rosemary and Mike, for all the times they convinced me not to drop out of the program. Funding from SSHRC in the form of a research scholarship is gratefully acknowledged: #410-2011-2417, as is funding from an admission scholarship from the University of Ottawa.

form. My goal in this paper is to unify these previous studies and propose a single, coherent theory of pronoun types and nominal projections. Some of the theoretical burden will be carried by a model of the noun phrase² where different types of pronouns realize different amounts of functional structure; however, I will also show that some of the implicational relationships between properties of different types of pronouns are best described in the form of a feature geometry, along the lines of Harley and Ritter (2002) and Cowper and Hall (2009).

The paper will proceed as follows: In Section 2 I will present two existing models of the structure of nominal projections, those of Déchaine and Wiltschko (2002) and Zamparelli (2000), both of which come with pronoun typologies. I will discuss how they propose similar explanations based on very different data, and in their current forms, they are partially incompatible with one another. I will propose changes to Déchaine and Wiltschko's (2002) in Section 3, relegating some of the work of their syntactic structure to the binding theory of Kratzer (2009). The result of this section will be a unified model of the nominal structure. In Section 4 I present data from Japanese, which has pronouns which seem to contradict the claims of the unified model with unwanted systematic ambiguities. I will show that these ambiguities are only apparent, resulting from independent properties of Japanese (and similar languages). In Section 5 I will bring in a third model, that of Cardinaletti and Starke (1994), which is based on an entirely different set of features than those of the unified model. I show that these features can be incorporated into the unified model in the form of a feature geometry. Section 6 concludes. Before beginning, I will go over some of the theoretical assumptions underpinning this research.

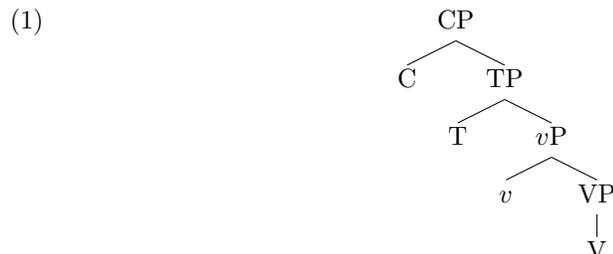
1.1 Methodological Preliminaries

In this study I will be comparing different ways in which pronouns have been classified, and as in any comparison of different theories, this involves choosing between theories based on a certain set of metatheoretical assumptions. Some of these are trivial; clearly, theory A should be preferred over theory B if A is empirically more accurate than B. There is also the metatheoretical tool of Occam's Razor, which states that a more elegant theory should be preferred over a more complex theory (based on some measure of complexity, which is generally stated in terms of the number of entities or the number of types of entities). However, in this study there will be times when I am presented with two theories, both with the roughly same level of complexity and empirical adequacy. In these cases, I am presented with a more difficult question: What makes a good linguistic theory?

In syntactic studies such as this, the question often comes down to what should or should not be explained using the mechanism of (functional) syntactic structure. One implication of a syntactic structure is a rigid c-command

²I will be using the term "noun phrase" to refer to the entire extended projection of the noun, in the sense of Grimshaw (1991); NP is used in the more precise, X-bar theoretic sense as the maximal projection of N.

hierarchy, stated in terms of subcategorization. For example, consider the following (perhaps oversimplified) representation of the clausal spine:



What this tree represents is a set of subcategorization requirements: *C* selects for *TP*, *T* selects for *vP*, and so on. In other words, part of what it means to be *C* is that if it has a complement, that complement is *TP*. Therefore, it would be unexpected for there to be a systematic set of homophonous lexical items of different categories; assuming that subcategorization is represented as a feature, this would require that for every item with feature set *A*, there is a phonologically identical item with a different feature set *B*. This will come into play in Sections 3 and 4, where systematic categorical ambiguity will be avoided unless there is some strong evidence in its favour.

I will be following Grimshaw (1991) and van Riemsdijk (1998) in assuming the existence of extended projections, in the sense that a lexical category (like *N*) shares certain features not only with its own maximal projection (*NP*), but also with a set of functional projections which dominate it (e.g. *DP*). However, contrary to Grimshaw and van Riemsdijk, I will follow Melchin and Truswell (2013) in recognizing that selectional differences between clauses and noun phrases suggest that symmetries between the two are not pervasive. In particular, it appears that the nominal extended projection consists of a single phase, while its clausal counterpart has two (a *C/T* phase and a *v/V* phase). In this paper I will be presenting a number of alternative models of the noun phrase; each is to be understood as a proposal for the structure of the nominal extended projection.

That being said, while syntactic structure is a powerful tool for representing hierarchical relationships between syntactic objects, it is not the only one. Implicational relationships between features can be represented independently of phrase structure in the form of complex feature structures and geometries. In the framework of HPSG (Pollard and Sag 1994), for example, syntactic processes are described almost entirely in terms of these feature complexes, represented as attribute-value matrices; even phrase structural relationships are described as feature specifications of nested lexical entries and inheritance hierarchies. While these theories have a great capacity for descriptive adequacy, they are theoretically too powerful in that it is difficult to constrain the theory so that it does not overgenerate. At the opposite theoretical extreme are frameworks such as Cartographic approach (such as that of Cinque 1999) and Starke's (2009) Nanosyntax, in which every feature is associated with a single syntactic head and its projection, and these heads are in a strict hierarchy of embedding. This

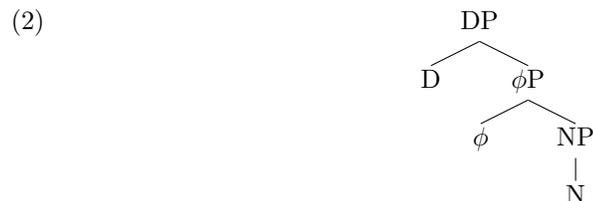
is theoretically very attractive, since it produces a neat, constrained theory, but as I will show throughout this paper, features often do not align themselves so neatly. Between these two extremes are feature geometric approaches such as those described in Section 5.2, where a head in the tree may have a complex feature geometry associated with it, employing the strengths of both syntactic projection and feature structures; this is the kind of model I will be working toward in this paper. In a feature-geometric approach such as this, the challenge is to determine which features are best represented in terms of category and selection (i.e. phrase structure), and which should be treated as part of the geometry, in a hierarchical relationship with other features of the same node. Facing this challenge is one of the goals of this paper.

2 The Models

In this section I will present two different three-layer partitions of the noun phrase: Déchaine and Wiltschko’s (2002) model designed to explain the different properties of different types of pronouns within and between languages, and Zamparelli’s (2000) model primarily intended to capture the semantics of determiners, quantifiers, and predication. The overall syntactic structures of the two models are fairly similar, but the layers in the two models are defined in terms of different grammatical features and they do not line up exactly.

2.1 Déchaine and Wiltschko (2002)

Déchaine and Wiltschko (2002) (henceforth D&W) form a theory of the relationship between pronoun types and nominal phrase structure based on the idea that “pronoun” is not a syntactic primitive. They note that there are different types of pronouns, with different distributions, and their differences derive from the fact that they realize different amounts of syntactic structure; thus, their syntactic categories correspondingly differ. Specifically, D&W propose that the full noun phrase has the following structure (p. 410):



A pronoun³ may realize any of the three maximal projections, DP, φP, or NP; the properties of the pronoun will vary depending on the type of projection it realizes.

³In the interest of clarity, I will refer to all nominal proforms as pronouns; for more specific types of pronoun, I will use pro-DP, pro-NP, etc.

	Pro-DP	Pro- ϕ P	Pro-NP
Internal syntax	D syntax; morphologically complex	neither D syntax nor N syntax	N syntax
Distribution	argument	argument or predicate	predicate
Semantics	definite	–	constant
Binding-theoretic status	R-expression	variable	–

Table 1: Nominal proform typology (Déchaine and Wiltschko 2002, p. 410)

A pro-DP is said to have the external syntax of a DP, and contains ϕ P and NP as subconstituents. Syntactically, they are restricted to argument positions; semantically, they are treated as definite, and therefore, in terms of binding theory, they are R-expressions – thus, they cannot be treated as bound variables. D&W consider pro- ϕ P to be a “cover term” for any intermediate projection between D and N encoding ϕ features (gender, number, sometimes person). As such, they have the syntax of neither determiners nor nouns, and therefore they have a less restricted distribution, as either arguments or adjuncts. D&W claim that ϕ has no inherent semantics (aside from its ϕ features), and therefore, in terms of binding theory ϕ P has the status of a variable. Finally, NPs are treated as lexical nouns, thus restricted to predicate positions. Semantically they are constants,⁴ and their binding properties are supposed to derive from their inherent semantics in a predictable way. The properties of each pro-form are summed up in Table 1.

Extending this analysis to Romance clitic pronouns, D&W claim that in French, the clitic *en* is a pro-NP, and the 3rd person *l*-clitics are pro- ϕ P. The status of *en* is due to the fact that it can stand in for a part of the nominal projection excluding numerals, certain quantifiers and modifying adjectives (Déchaine and Wiltschko 2002, pp. 427-428):

- (3) a. J' ai vu plusieurs livres
I have seen several books
b. J' *en* ai vu plusieurs
I *en* have seen several
- (4) a. J' ai vu un grand livre
I have seen a large book
b. J' *en* ai vu un grand
I *en* have seen a large

In addition, as expected with a pro-NP functioning semantically as a constant, *en* cannot function as a bound variable (5), and it cannot co-refer with an antecedent (6) (Déchaine and Wiltschko 2002, p. 428):

⁴The use of the term “constant” to describe pro-NPs is somewhat misleading; clearly, no pro-form is constant in the sense that its meaning is fixed from sentence to sentence like a lexical noun. D&W probably use the term “constant” as opposed to “variable,” meaning that pro-NP has a fixed meaning in a given context, and may not covary with a quantifier like pro- ϕ P. I will continue to follow D&W’s terminology in this respect.

- (5) * $[Chacun]_i$ pense que Jean $[en]_i$ a vu
 each.one thinks that J. *en* has seen
- (6) * $[Marie]_i$ pense que Jean $[en]_i$ a vu
 M. thinks that J. *en* has seen

Thus, *en*'s status as pro-NP is unproblematic.

Next, D&W show that the French 3rd-person *l*-clitics are pro- ϕ P. First, they may appear both as arguments (7) and as predicates (8) (although as a predicate, only uninflected *le* is possible; see Section 2.3) (Déchaine and Wiltschko 2002, p. 428):

- (7) Jeanne *la* voit
 J. her sees
 "Jeanne sees her."
- (8) Marie est une avocate, et Jeanne $\{*la/le\}$ sera aussi
 M. is a lawyer(FEM), and J. it will.be too
 "Marie is a lawyer, and Jeanne will be one too."

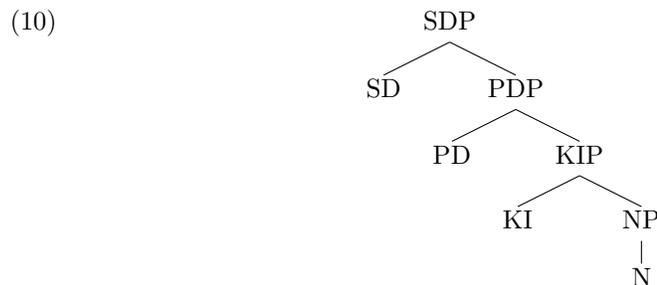
There is a difference in the clitic depending on its function; when it is an argument as in (7), it shows agreement features, while as a predicate it does not (8). D&W claim that this results from the relationship between Case and agreement: if an element has Case (which arguments do, and predicates do not), then it has agreement. Furthermore, these pronouns may be bound (Déchaine and Wiltschko 2002, p. 429):

- (9) $[Chaque homme]_i$ pense que Marie $[l]_i$ a vu
 each man thinks that M. him has seen
 "Each man thinks that Marie has seen him."

Thus, the French *l*-clitics fit D&W's definition of a pro- ϕ P.

2.2 Zamparelli (2000)

Working along different lines, Zamparelli (2000) proposed a structure for the noun phrase, divided along similar lines to that of D&W. Zamparelli's structure is the following (p. 16):



Here, SD stands for “strong determiner,” PD is “predicate determiner,” and KI denotes a “kind.” SDP is said to be of semantic type e (individual), and it is referential (R-expression). PDP is a property or predicate of the semantic type $\langle e, t \rangle$; as such it appears in such predicate positions as the post-copular position. This layer can be negated or modalized. KIP has the semantic type e , denoting kinds. Zamparelli does little to differentiate KIP and NP.

As evidence for this view, Zamparelli notes that the three-way partition of Italian clitic pronouns results from each type of clitic realizing a different amount of structure. The clitic lo_{+Agr} (with agreement for gender and number) is pro-SDP, lo_{-Agr} (plain lo with no agreement) is pro-PDP, and ne is pro-KIP. This analysis is motivated by their distribution: lo_{+Agr} pronominalizes an element in argument position (11), lo_{-Agr} realizes a predicate (12), and ne stands in for a sub-part of the noun phrase, not including weak determiners (13) (Zamparelli 2000, p. 19):

- (11) Gianni vide [Maria]_{*i*}, ma non {*la*_{*i*} / **lo* / **ne*} riconobbe.
 G. saw M., but (he) not {her_{*FSng*} / one / of.it} recognized
- (12) Quelli sono [due linguisti]_{*i*}, ma non {**li* / *lo*_{*i*} / **ne*} saranno
 Those are two linguists, but not {them_{*MPl*} / such / of.it} will be
 a lungo.
 for long
- (13) Gianni e Maria aiutano molti [linguisti], ma {**li* / **lo* /
 G. and M. help many linguists, but {them_{*MPl*} / one /
*ne*_{*i*}} conoscono [pochi *t*_{*i*}].
 of.them_{*i*}} know [few *t*_{*i*}]

Note that these Italian clitic pronouns behave in the same way as the French pronouns considered by D&W; lo_{+Agr} corresponds to the inflected la in (7), lo_{-Agr} to the uninflected le in (8), and ne to en . If D&W’s model can be mapped directly onto Zamparelli’s (DP=SDP, ϕ P=PDP, NP=KIP), then the status of the latter two pronouns are uncontroversial; however, the authors disagree on the status of the inflected l -pronoun, with D&W lumping it into the intermediate projection with its uninflected counterpart, while Zamparelli places it in the highest projection. The differences in categorization result from differences between the two models, described in the next section.

2.3 Similarities and Differences

The two models have some obvious similarities, both in the overall structural appearance and in the properties associated with each level. However, there are some key differences which show that the models are not interchangeable.

The two models assign similar characteristics to the topmost level, DP or SDP. Both consider this category to be restricted to argument positions, to have definite semantics and referential properties and, thus, to denote entities (type e). However, they differ in their views of the binding properties of this

category: D&W consider DPs to be unable to act as bound variables, due to their referential status. This is why they do not place the inflected *la* in this category, since, as shown in (9), this pronoun can be bound. In Zamparelli’s model, there is no reason why SDP may not be bound; in fact, it makes no sense for PDP to act as a bound variable, as explained below.

The middle layer, ϕ P or PDP, is where most of the differences between the models reside. Both models consider this node to be the locus of (most) ϕ -features. However, in D&W’s model, ϕ P is undefined with respect to many syntactic and semantic properties: it may act as either an argument or a predicate, and it is the only category that may function as a bound variable. On the other hand, Zamparelli treats PDP exclusively as a predicate or property of type $\langle e, t \rangle$; as such, it cannot be a bound variable, since variables must be semantic individuals rather than properties. Thus, the models agree that the predicative, uninflected *l*-clitic belongs in this category. However, Zamparelli may not place the inflected clitic in this category, while D&W are forced to, for the same reason: it may be a bound variable.

The bottom layer, NP or KIP, is broadly similar between the two models. Both consider it to represent a subset of the noun phrase which excludes determiners and quantifiers, and may exclude (certain) adjectives. Thus, the *en/ne* clitic falls into this category in both models, as does the English pronoun *one*. However, there is a key difference in how the layers are defined: D&W’s NP is a predicate, and Zamparelli’s KIP is a type *e* individual. This difference does not seem to lead to differences in their predictions, probably because at that level of embedding, the question of what is and is not a predicate becomes pretty abstract; however, leaving that aside, it is unproblematic for my purposes to collapse KIP and NP in Zamparelli’s system and equate this category with D&W’s NP.

Thus, the comparison of the models can be summed up as follows: NP corresponds to KIP/NP. ϕ P corresponds to both PDP (when it is a predicate) and SDP (when it is an argument, or a bindable variable). DP corresponds to SDP, except that, as mentioned, DP cannot be a variable, while in Zamparelli’s model all variables must be SDP. Thus, there is a many-to-many relationship between the elements of two models.

3 Discussion

The system of Déchaine and Wiltschko (2002) has some built-in systematic ambiguities, particularly between the pronouns analyzed as DP and ϕ P in English and French. D&W support their ambiguity analysis using evidence from “transitive” pronouns with overt NP complements (e.g. *us linguists*), claiming that the complement forces an unambiguous structure. However, analyzing these “complements” as appositive modifiers dispels this claim, and these ambiguities can be eliminated if the binding properties of a pronoun are explained with a mechanism along the lines of Kratzer (2009), allowing syntactic structure to explain mostly semantic category differences. In this way, D&W’s DP and ϕ P

categories can be combined into a single category, paving the way for a unification of the D&W and Zamparelli models.

3.1 Ambiguities and Apposition

There seems to be a systematic ambiguity in the pronouns that D&W treat as pro-DP and pro- ϕ P. For instance, English 1st and 2nd person pronouns are analyzed as typically being pro-DP in their system; part of the motivation for this is the fact that the plural forms of these pronouns may appear with overt complements, as in *us linguists* and *you lovers*, which are assumed to have the structure in (14) (Déchaine and Wiltschko 2012, p. 15):

(14) [DP us [ϕ P -s [NP linguist]]]

However, there are situations where 1st and 2nd pronouns can function as bound variables, such as the following (Rullmann 2004, pp. 160–161):

- (15) a. Only I got a question that I understood
b. You (guys) all think you're smart.

Recall that in their system, only pro- ϕ P can act as a bound variable; this means that these pronouns must be ambiguous between ϕ P and DP. Déchaine and Wiltschko (2012) claim that this ambiguity arises from the fact that a ϕ P may be coerced into being DP (or vice-versa) in certain situations. They support this with the fact that, when the pronouns have a complement as in (14), there is no ambiguity. Particularly, pronouns in this construction may not be bound, as illustrated in (16):

(16) Every guy I've ever dated has wanted us lovers to get married.

An interpretation where *us lovers* both covaries with *every guy* and refers indexically to the speaker is unavailable, unlike the corresponding sentence without the complement *lovers*.

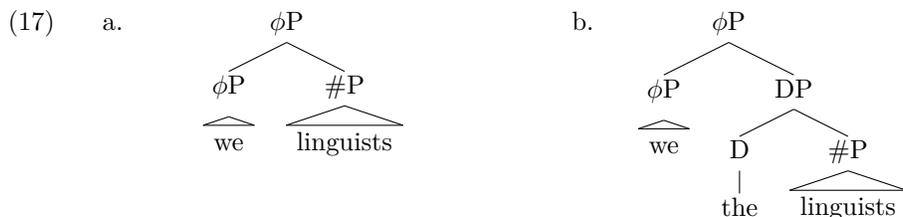
Given these apparently systematic DP- ϕ P ambiguities, one might ask why there are no pronouns which are ambiguous between ϕ P and NP, or indeed between the three.⁵ One way to dispel these ambiguities between DP and ϕ P, while at the same time solidifying the relationships between the layers in D&W and Zamparelli's models, would be to propose an intervening YP between the two (and correspondingly, between SDP and PDP in Zamparelli's model). YP would have the property of denoting an individual (not a predicate) which may be a bound variable. Then, the many-to-many relationship noted above could be reduced to three one-to-one relationships. However, this wreaks havoc on both systems. This node is otherwise unmotivated, and it messes with all of the neat syntactic and semantic solutions that Zamparelli's model provides. In particular, the correspondences between syntactic category and semantic type

⁵In Section 4 I will show that there is in fact a set of systematic ϕ P-NP ambiguities in Japanese, but these are due to other properties of that language, and can be set aside for the moment.

(argument versus predicate status) are lost, or at least complicated, if another type e node is introduced. Also, in such a system there is the possibility for a whole set of ambiguities, yielding a pronominal typology far more complex than what is actually attested – if, as D&W claim, there are systematic categorical ambiguities, then one would expect still more ambiguities once further categories are introduced.

That being said, D&W explain the lack of binding in the presence of an overt complement by claiming that the presence of an overt complement forces the pronoun to surface in D, blocking coercion into ϕ (and therefore blocking the possibility of binding). However, I propose that the lack of binding with an overt complement is due not to the syntactic configuration, but to the semantics: The complement *lovers* is a form of appositive modification, and such modification precludes the possibility of a bound variable interpretation.

In some ways, this suggestion is similar to a recent proposal by Cowper and Hall (2009), although as we shall see, the two proposals differ in some key empirical details. Cowper and Hall develop an alternative model to that of D&W, where ϕP is semantic type e , and $\#P$ is a property, roughly equivalent to Zamparelli’s PDP (their system is explained in more detail in Section 5.2.2). They claim that in an expression like *we linguists* the noun *linguists* is a modifier, rather than a complement as in (14); the structure is thus parallel to that of *we the linguists*, as shown below (Cowper and Hall 2009, p. 115):



According to Cowper and Hall, this can be either restrictive or appositive (non-restrictive) modification. The denotation of *we linguists* will therefore be as follows (p. 115):

$$(18) \quad \llbracket \mathbf{we\ linguists} \rrbracket = \text{the plural individual } y \text{ including the speaker such that} \\ y \sqsubseteq x \& \text{LINGUISTS}(y) \\ = \text{'those of us who are linguists'}$$

In other words, the modifier *linguists* further restricts the set denoted by *we* to include only those who are linguists. Cowper and Hall conclude their analysis by noting that a bound variable has no features of its own, receiving its interpretation from its binder; ϕP s which are modified as in (17) and (18) therefore have too much semantic content to enter into a binding relationship, an idea to which I will return in the discussion of epithets in Section 4.2.2.

However, it seems that the semantics of expressions like *we linguists* may not be one of restrictive modification, but rather must be apposition. If it were restrictive, then one would expect there to be no problem with *we linguists*

referring to a proper subset of *us* in (19); however, this interpretation seems impossible.

- (19) ??The question caused confusion among us, but we linguists all understood perfectly.

The only available interpretation seems to be one where *we linguists* refers to the same set as *us*, suggesting that restrictive modification is impossible in this case. In contrast, if *we linguists* is replaced with *the linguists among us*, which is unambiguously unrestrictive, then the sentence is fine:

- (20) The question caused confusion among us, but the linguists among us all understood perfectly.

The fact that there is a contrast between (19) and (20) indicates that there must be a contrast between *we linguists* and *the linguists among us*. I suggest that this is because the former is appositive modification, while the latter is restrictive. The lack of binding of such an appositive construction is explained in the theory of binding proposed by Kratzer (2009) (which I explain and adopt in Section 3.3). Kratzer notes that descriptive features (such as those contributed by an appositive predicate) cannot semantically combine with the numerical (indexical) features characteristic of bound variables, as demonstrated by (21) (p. 223):⁶

- (21) *No boy was admitted to the Mystery Spot, before he, who had to wait in line for at least two hours, signed a waiver-of-liability form.

Therefore, an analysis of these expressions as appositives explains both their lack of restrictive modification semantics, and the fact that they cannot appear as bound variables; they do not provide proof of the ambiguity and coercion of D&W's analysis. In other words, this is not a genuine syntactic ambiguity, contrary to D&W's predictions, as the failure to bind *us linguists* can be explained without recourse to D&W's proposed syntactic structure.

Of course, the price of this apposition explanation for the lack of binding of pronouns with complements is that we can no longer predict the binding properties of a pro-form using its categorial status alone. D&W's system was based on the assumption that binding theory depends on the syntactic category of an item, and little else, but I have shown that bound-variable status can vary within a single category. If, as I will argue in the next section, the properties of pronouns are best described relative to a model like Zamparelli's, based mostly on semantic type, then it must be shown that there is some category-independent theory of binding with which the model is compatible. In Section 3.3 I will present such a proposal from Kratzer (2009).

⁶At first glance, it seems that this would prevent binding of epithets, since they are more than just an index. However, the additional semantic content of epithets is expressive, rather than descriptive – they do not contribute to the truth-conditional content of a sentence except to introduce a variable. The same restriction does not seem to apply for expressive content as for descriptive content.

3.2 The Unified Model

As discussed above, a large part of D&W’s motivation for their three-way categorization of pronouns comes from their binding properties: DP is an R-expression which cannot be bound; ϕ P can function as a bound variable; and NP may not be bound but it may corefer or not, depending on its semantic content. On the other hand, Zamparelli’s tripartition mostly reflects the semantic type of the element: SDP is a type e entity (or a type $\langle\langle e, t \rangle, t \rangle$ quantifier, which leaves a type e trace after QR); PDP denotes a type $\langle e, t \rangle$ property; and KIP is type e , denoting a kind. Here I will show that, if binding properties are left out of the picture and ϕ P is considered to denote only predicates (so that all individuals are DP), and some non-structural explanation for a pronoun’s binding properties can be found, then the ambiguities mentioned above are resolved and the two systems are equivalent.

Déchainé and Wiltschko (2002) include an analysis of English personal pronouns in which they claim that the 3rd person pronouns are pro- ϕ P, while the 1st and 2nd person pronouns are pro-DP. Their explanation for this is based on their binding status as well as whether or not they can function as predicates. First, they show that 3rd person pronouns may function as bound variables (22), while 1st and 2nd person ones may not (23) (p. 423):

- (22) $[Every\ candidate]_i$ thinks that $[he]_i$ will win.
 $\forall x, \text{candidate}(x), x$ thinks that x will win
- (23) I_i know that John saw me_i , and Mary does too.
 = a. $I [\lambda x . x$ knows that John saw me] & Mary $[\lambda x . x$ knows that John saw me]
 \neq b. $I [\lambda x . x$ knows that John saw x] & Mary $[\lambda x . x$ knows that John saw x]

(22) is uncontroversial; *he* may have a bound variable reading in this kind of context. (23) is meant to show that the English 1st person pronoun may not function as a bound variable; the strict reading (a) is available, where *me* is indexical (refers to the speaker), while the sloppy, bound variable reading (b) is impossible. However, judgements vary on its availability, with some speakers judging (b) to be available, at least in certain contexts. Rullmann (2004) notes that if *and* is replaced with *but* in this kind of sentence, a bound variable reading is easier to obtain (24), and there are other types of sentence where it is easier still (25) (pp. 161-162):

- (24) I think I am John’s favorite ex-wife, but Mary does too.
 (25) We each/all think we’re the smartest person in the world.

In (25) especially the second *we* must be bound, given the logical impossibility of **we’re the smartest person in the world*. So, it seems that, in certain circumstances at least, 1st and 2nd person pronouns in English can be bound.

The other criterion D&W use to categorize 1st and 2nd person pronouns as DP and 3rd person as ϕ P involves their status as arguments or predicates. All

three persons may be used as arguments. However, D&W show that 3rd person pronouns, unlike the others, may function as predicates. They cite the following as evidence (p. 425):

(26) That's [*her*]_{PRED}

However, 1st and 2nd person pronouns can also occur in this position, as D&W point out:

(27) That's [*me*]

(28) That's [*you*]

D&W claim that these are cases of equative structures; however, it seems strange to posit this explanation for these two but not for (26). Presumably, since all three pronouns can occur in this context, this cannot be taken as evidence for a contrast.⁷ Another piece of evidence that they bring forward is that 3rd person pronouns take part in word formation in ways that 1st and 2nd person pronouns cannot (p. 426):

(29) a. [*she*]-male
 [*he*]-goat
 b. *[*me*]-male
 c. *[*you*]-goat

However, Rullmann (2004, fn. 2) notes that there are examples of 1st and 2nd pronouns acting in this way:

(30) the me-generation
 we-society
 you-factor

In each case in (30), the pronoun's contribution to the meaning of the compound is not deictic, relative to the speaker or addressee; rather, it is something more metaphorical. For example, *the me-generation* may refer to a generation of self-centered people. In fact, the same is true in (29a), where the pronouns contribute the gender, not any deictic information. It is difficult to imagine such a meaning for the examples in (29b) and (29c), so it is not surprising that those compounds are ungrammatical. Word-formation does not provide compelling evidence for a categorical distinction between different kinds of personal pronouns in English. Furthermore, it is not clear that an ability to participate in word-formation indicates that a pronoun must be predicative.

D&W also consider French personal clitic pronouns, using the same bound variable tests as used above for English to show that they are ϕ P in all three persons. That 3rd person clitics can be bound is shown by (9); D&W use (31) to demonstrate that 1st person clitics can be bound as well (p. 431):

⁷It is also not clear that the post-copular phrase in these examples is best analyzed as a predicate, but the proper analysis of these sentences is outside the scope of this paper.

- (31) [Je]_{*i*} pense que la police [m]_{*i*} a vu, et Marie le pense aussi.
 I think that the police me have seen, and M. it thinks also
- a. I [$\lambda x . x$ thinks that the police saw me] & Mary [$\lambda x . x$ thinks that the police saw me]
- b. I [$\lambda x . x$ thinks that the police saw x] & Mary [$\lambda x . x$ thinks that the police saw x]

Here, D&W claim that both bound and unbound readings are available, unlike the parallel English example (23). However, they note (fn. 24) that not all speakers accept the bound variable interpretation (b). Therefore, since judgments seem to be mixed in both English and French, there seems to be no reason to posit different categorical status of 1st and 2nd person pronouns in the two languages on the basis of this test. D&W also point out that French 1st and 2nd person pronouns can take part in word-formation. Since this is also the case with their English equivalents, this further suggests that English and French pronouns should be considered the same category.

To sum up so far, I have shown that there is little motivation for any claim based on binding properties or argument-predicate status that 1st and 2nd person pronouns contrast with 3rd person pronouns in English. Therefore, by D&W's model, they should be considered the same syntactic category – either DP or ϕ P. Now the task remains to determine in which of the two categories they belong.

D&W consider all DPs to be R-expressions, so they cannot be bound, while ϕ Ps are all variables. However, epithets are an example of nominal phrases which are transparently DP, which can be bound by a quantifier.⁸ Consider (32):

- (32) [*Every boy's*]_{*i*} mother thinks [*the little bastard*]_{*i*} should be sent to bed early.

For every boy x , x 's mother thinks x should be sent to bed early

Here, the presence of the determiner *the* (D&W's D, Zamparelli's SD) shows that the epithet *the little bastard* should be considered a full DP. However, it can also be bound. Therefore, the conclusion is unavoidable: in at least certain circumstances, DPs can act as bound variables, contra D&W's claims.

Now that DPs can be bound, the only remaining difference between D&W and Zamparelli's models is the fact that ϕ P may be either an argument or a predicate, while PDP must be a predicate. However, the only reason that D&W had placed any type e pronouns in ϕ P in the first place was to satisfy their binding requirements without positing separate, homonymous DP and ϕ P forms of the same pronoun. If we consider all type e pronouns to be DP, and all predicative pronouns (like the uninflected *le* in (8)) to be ϕ P, then in fact the two models converge.

⁸Binding of epithets is subject to the antilogophoricity constraint; see Section 4.2.2 for more on the binding and coreference properties of epithets.

The data D&W cite from the Salish language Halkomelem actually supports this hypothesis. Halkomelem is a language in which the D and ϕ elements of the 3rd person pronoun $tú-tl'ò$ are both overt morphemes ($tú = D$, $tl'ò = \phi$). D&W show that, with both morphemes present, the pronoun may function as an argument, but not a predicate (33), while when only the ϕ P is present, it functions only as a predicate (34) (p. 413):

- (33) a. [Lám]_{PRED} [tú-tl'ò]_{ARG}
 go D-PHI
 ‘He goes’
 b. *[Tú-tl'ò]_{PRED-cha} te Bill kw'e may-th-óme
 D-PHI-FUT D B. C help-TRANS-2.OBJ
- (34) a. *[Lám]_{PRED} [tl'ò]_{ARG}
 go PHI
 b. [Tl'ò]_{PRED-cha} te Bill kw'e may-th-óme
 PHI-FUT D B. C help-TRANS-2.OBJ
 ‘It will be Bill that helps you’

D&W note that the pro-DP $tú-tl'ò$ cannot function as a bound variable. They attribute this to the fact that it is pro-DP and therefore an R-expression; however, there may be an independent reason. D&W mention that overt pronouns in Halkomelem are emphatic pronouns. It may be the case that emphatic pronouns in general may not be used as bound variables.⁹ Consider (35), with emphasis on *he*:

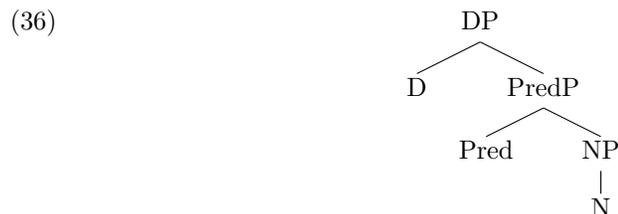
- (35) [Every man]_i thinks HE_{?i} is the smartest in the world.

In this example, there is a strong preference for an indexical reading of *he*, referring to an individual in the physical or discourse context, rather than a bound variable reading. The pro- ϕ P $tl'ò$ cannot be a bound variable either, for the simple reason that it is a predicate, with neither the semantic type e nor the argumental distribution to allow it to function that way. The crucial thing that the Halkomelem data shows is that, in a case where there is transparent morphology indicating the D and ϕ heads, the DP is unambiguously an argument, and ϕ P a predicate, in accordance with Zamparelli’s model.

In this section I have demonstrated that if categorical differences are used to account for such properties as semantic type (individual versus predicate), and if binding properties are left to some other mechanism (such as that of Kratzer 2009, explained below), then the models of Déchaine and Wiltschko (2002) and Zamparelli (2000) can be combined. The result uses the syntactic structure of Zamparelli’s model, and deals with the binding facts from D&W in non-structural ways. From here on, for convenience I will refer to this merged

⁹In terms of Cardinaletti and Starke’s (1994) pronominal typology, it appears that these Halkomelem pronouns, like the emphatic *HE* in (35), are strong pronouns, which may not be bound due to the fact that they already have a semantic range-restriction; see Section 5.1 for more on Cardinaletti and Starke’s typology and how it relates to the model being developed here.

model as the “unified” model. The unified model has the three projections that are common to the two models, which I will refer to as follows: the top layer (DP/SDP) is DP, as it is “traditionally” known; the middle layer (ϕ P/PDP) is PredP, since it has no clear association with ϕ -features (or at least, no more so than the other layers), it is unambiguously predicative, and PredP is a more transparent label than PDP; and the bottom layer (KIP/NP) is NP, again in keeping with familiar notation. The structure of the noun phrase will thus be as in (36):



3.3 Kratzer (2009) on Binding

The unified model proposed above is only viable if it can be shown that binding can be explained independently from the categorical structure of a pronoun, contrary to the claims of D&W. This is the view in the traditional binding theory (from Government-Binding theory, Chomsky 1981); Kratzer (2009) presents a newer version which is compatible with the unified model. In Kratzer’s analysis, bound variable pronouns (including “fake indexicals,”¹⁰ as in (24) and (25)) are created in one of two ways: they may be born with a defective feature set and have ϕ -features transmitted to them from a λ operator; alternatively, they may be born fully specified, and receive their bound interpretation via a context-shifter, in the form of a person feature on the binder. The former mechanism is available only for locally-bound anaphora, while the latter applies freely. Through these two mechanisms, binding can be explained in a way that fits with the unified model of DP structure.

In Kratzer’s model, the actual binders are λ -operators on v and other functional heads. The evidence for this comes from German sentences like those in (37), where a bound variable reading of the pronoun is available only in (37c) (Kratzer 2009, pp. 191–192):

- (37) a. Ich bin der einzige, der t *meinen*
 1SG be.1SG the.MASC.SG only.one who.MASC.SG 1SG.POSS.ACC
 Sohn versorg-t.
 son take.care.of-3SG
 “I am the only one who is taking care of my son.”
 b. Du bist der einzige, der t *deinen*
 2SG be.2SG the.MASC.SG only.one who.MASC.SG 2SG.POSS.ACC

¹⁰Fake indexicals are those 1st and 2nd person pronouns which are ambiguous between a referential reading and a bound variable reading; when bound, they look like indexicals but are not, hence the name.

- Sohn *versorg-t*.
 son take.care.of-3SG
 “You are the only one who is taking care of your son.”
- c. Wir sind die einzigen, die *t unseren* Sohn
 1PL be.1/3PL the.PL only.ones who.PL 1PL.POSS.ACC son
versorg-en.
 take.care.of-1/3PL
 “We are the only ones who are taking care of our son.”
- d. Ihr seid die einzigen, die *t euren* Sohn
 2PL be.2PL the.PL only.ones who.PL 2PL.POSS.ACC son
versorg-en.
 take.care.of-1/3PL
 “I am the only one who is taking care of my son.”

The thing that sets (37c) apart is the fact that, due to the syncretism in the German verb agreement paradigm, the person and number features on the possessive pronoun and the embedded verb match. In the other examples, on the other hand, the features clash. Thus, it seems that the availability of a bound variable interpretation depends on ϕ -feature compatibility with the nearest verb (exactly why this should hold is not important to our current concerns). Support for this notion comes from dialects where these sentences do not involve a ϕ -feature clash; (38) is from a dialect of German with a different agreement paradigm which results in matching ϕ -features, while (39) is from Dutch, which has no person distinctions in the plural. In these examples, a bound variable interpretation is readily available (p. 192):

- (38) Ihr seid die einzigen, die *t euren* Sohn
 2PL be.2PL the.PL only.ones who.PL 2PL.POSS.ACC son
versorg-t.
 take.care.of-2PL
 “I am the only one who is taking care of my son.”
- (39) a. Wij zijn de enigen die *ons* best gedaan
 1PL be.PL the only.ones who 1PL.POSS best done
hebben.
 have.1/2/3.PL
 “We are the only ones who have done our best.”
- b. Jullie zijn de enigen die *je* best gedaan
 2PL be.PL the only.ones who 2PL.POSS best done
hebben.
 have.1/2/3.PL
 “You are the only ones who have done your best.”

Finally, to show that it is indeed feature clash with the verb, rather than a nominal antecedent, which leads to the unavailability of a bound variable interpretation in (37) (with the exception of (37c)), Kratzer shows (40), where the possessed DP is now the subject of the embedded clause, leading to singular

agreement on the embedded verb and thus a ϕ -feature clash (p. 192):

- (40) Wir sind die einzigen, die *unser* Sohn t
 1PL be.1/3PL the.PL only.ones who.PL 1PL.POSS.NOM son
 versorg-*t*.
 take.care.of-3SG
 “We are the only ones who our son is taking care of.”

In this example, unlike (37c) a bound variable reading is unavailable, suggesting that it is indeed ϕ -feature compatibility with the nearest verb which leads to the availability of a bound variable interpretation.

Long-distance binding operates under different mechanisms from local anaphora, a fact demonstrated by the contrast between the German sentences (41) and (42) (Kratzer 2009, pp. 212).

- (41) Du bist der einzige, der *deinen*
 2SG be.2SG the.MASC.SG only.one who.MASC.SG 2SG.POSS.ACC
 Aufsatz versteht.
 paper understands
 “You are the only one who understands your paper.”
- (42) a. Du bist der einzige, der glaubt, dass
 2SG be.2SG the.MASC.SG only.one who.MASC.SG believes that
 jemand *deinen* Aufsatz versteht.
 somebody 2SG.POSS.ACC paper understands
 “You are the only one who thinks that somebody understands your
 paper.”
- b. Du bist der einzige, der jemand
 2SG be.2SG the.MASC.SG only.one who.MASC.SG somebody
 kennt, der *deinen* Aufsatz versteht.
 knows who.MASC.SG 2SG.POSS.ACC paper understands
 “You are the only one who knows somebody who understands your
 paper.”

In (41), binding is impossible due to a ϕ -feature mismatch between the pronoun and the verbal inflection. However, in the long-distance contexts of (42), the same mismatch does not block binding, and these pronouns have an available bound variable reading. Thus, there must be some different mechanism for non-local binding.

Kratzer claims that the mechanism driving long-distance binding of fake indexicals is a version context-shifting. Unlike the minimal pronouns involved in local binding as described above, pronouns which are non-locally bound are born with an indexical (1st or 2nd person) feature (hence, they are fully-specified). Their interpretation comes from a context-shifting rule, which Kratzer takes to be a kind of λ -abstraction. Unlike regular λ -abstraction, which changes variable assignments, this indexical λ -abstraction shifts the context parameter to change the reference of 1st and 2nd person pronouns (p. 213):

- (43) *Indexical λ -abstraction*
 $\llbracket \lambda[\mathbf{1st}] \alpha \rrbracket^{g,c} = \lambda x \llbracket \alpha \rrbracket^{g,c'}$, where c' is like c , except possibly that $\text{speaker}(c') = x$.
 $\llbracket \lambda[\mathbf{2nd}] \alpha \rrbracket^{g,c} = \lambda x \llbracket \alpha \rrbracket^{g,c'}$, where c' is like c , except possibly that $\text{addressee}(c') = x$.

Here, the features [1st] and [2nd] denote the speaker and addressee in c . These context shifters are considered to be 1st and 2nd person features on the binder.

Thus, Kratzer has developed a system of binding which does not specifically require a particular syntactic structure of the pronoun, making it compatible with the unified model proposed above. As long as the feature unification or context shifting does not result in an incompatible set of ϕ -features, binding is possible. The different mechanisms for local and long-distance binding also explain why the two types of binding have different characteristics, such as special reflexive forms for local anaphora and often laxer restrictions on long-distance binding. In this way, binding theory is explained with reference to features, but not categorical features in particular, so it can be seen as independent of the syntactic structure of the items involved.

4 Japanese Third Person Pronouns

In this section I will discuss the Japanese 3rd person pronouns, *kare* ‘he’ and *kanozyo* ‘she.’ These pronouns pattern somewhat differently from those in English and French. In particular, they may only function as bound variables in a limited set of environments, namely antilogophoric contexts. Yashima (2013) shows that these pronouns have properties of epithets, which are also bindable only in antilogophoric environments. Furthermore, Yashima claims that these pronouns are ambiguous between DP and NP epithets, which have somewhat different properties. Recast in the model of Déchaine and Wiltschko (2002), this translates to an ambiguity between NP and ϕ P, which is a previously-unseen ambiguity; I will propose that this is due to the lack of a D head in Japanese. I will then show that, in terms of the unified model developed in Section 3.2, this translates to an ambiguity between DP and NP. Finally, I will present type-shifting operations from Chierchia (1998) with which it can be shown that this ambiguity is not between two categories, but in fact results from the fact that certain languages lacking D have these type-shifting operations available as a last resort.

4.1 Déchaine and Wiltschko (2002) on *Kare*

Déchaine and Wiltschko (2002) cite *kare*¹¹ as the canonical pro-NP. They show that it has the distribution of a noun, and that it cannot function as a bound variable, although it can corefer.

¹¹In the interest of brevity, I will talk only about *kare* ‘he’ in this section; however, *kanozyo* ‘she’ has equivalent properties and should receive the same analysis.

First, D&W demonstrate that *kare* has the distribution of a bare noun. In particular, it can appear with an attribute adjective, a possessor,¹² or a demonstrative (p. 417):

- (44) a. *tiisai kare*
 small he
 “he who is small”
 b. *watasi-no kare*
 I-GEN he
 “my boyfriend”
 c. *kono kare*
 this he
 “this guy here”

Since these modifiers can attach to NP, but not PredP or DP, the fact that they can occur with *kare* leads D&W to the conclusion that it is a pro-NP.

In D&W’s model, NPs have the semantics of nouns, which are constants; therefore, they cannot be bound. This appears to be the case for *kare*, demonstrated by the fact that in (45), the pronouns cannot have a bound variable reading (p. 417):

- (45) a. **Daremo_i-ga kare_i-no hahaoya-o aisite-iru*
 Everyone-NOM he-GEN mother-ACC love-PRES
 “Everyone loves his mother.”
 $\neq \forall x, x$ loves x ’s mother
 b. **Dono zyosei_i-mo [kanozyo_i-ga tensai-da to] omotte-iru*
 every woman-PART she-NOM genius-COP COMP think-PRES
 “Every woman thinks she is a genius.”
 $\neq \forall x, \text{woman}(x), x$ thinks x is a genius

Therefore, D&W conclude that *kare* cannot be bound (although I present evidence to the contrary below).

As for coreference properties, in D&W’s model whether a pro-NP can corefer depends on the inherent semantics of the pronoun. They claim that *kare* has the features [MALE] and [MARRIAGEABLE AGE], so therefore it has referential properties and it should be expected to corefer. This prediction is borne out (p. 418):

- (46) a. *John_i-ga kare_i-no hahaoya-o aisite-iru*
 J.-NOM he-GEN mother-ACC love-PRES
 “John loves his mother.”
 b. *Mary_i-ga [kanozyo_i-ga tensai-da to] omotte-iru*
 M.-NOM she-NOM genius-COP COMP think-PRES
 “Mary thinks she is a genius.”

Therefore, D&W conclude that *kare* must be analyzed as pro-NP.

¹²The presence of a possessor forces it to have the lexical meaning of ‘boyfriend’ (for *kanozyo*, ‘girlfriend’), which will be discussed in Section 4.2.4.

4.2 Yashima (2013) on *Kare*

D&W's analysis of *kare* as pro-NP largely hinges on the observation that it cannot be bound. However, Yashima (2013) demonstrates that in certain environments, binding is in fact possible. The environments in question are antilogophoric, which happen to be the same environments in which binding of epithets is possible (Dubinsky and Hamilton 1998). Yashima shows that *kare*'s binding properties (and certain other puzzling facts about it) suggest that *kare* is an epithet. However, the lack of antilogophoric restrictions on the pronoun in coreference situations leads Yashima to conclude that there are two types of epithets, DP and NP, which have different properties, and *kare* is ambiguous between the two.

4.2.1 Binding *kare*

Citing examples like (45), D&W claim that *kare* cannot function as a bound variable. However, Yashima (2013) gives some examples of situations in which it can be bound. Yashima notes that in the following sentences, a bound variable reading is possible (p. 13):

- (47) a. [Dono nooberusyoo zyusyoo sakka]_i-ga kare_i-no kuruma-de
which Nobel.prize winning author-NOM he-GEN car-in
kita-no?
came-Q
“Which Nobel Prize winning author came in his car?”
- b. Dono gakusei_i-mo sensyuu kare_i-o suisensita sensei-ni
every student-PART last.week he-ACC recommended teacher-DAT
orei-o okutta.
present-ACC sent
“Every student sent a present to the teacher who recommended him last week.”
- c. [Sono ondai-ni haitta] zyosi gakusei-no
that music.college-DAT entered female student-GEN
daremo_i-ga [kanozyo_i-no sainoo-o mottomo yoku hikidasite
everyone-NOM she-GEN talent-ACC most fully bring.out
kureru] sensei-ni dea-e-ta.
do.the.favor teacher-DAT meet-can-PAST
“Every female student who entered that music college was able to meet a teacher who could bring out her talent to the full extent.”

These examples show that *kare* can be bound, but only in a subset of the environments in which English personal pronouns can be bound. However, the set of environments in which *kare* can be bound does correspond with that of epithets.

4.2.2 Epithets

Anaphoric epithets are definite DPs which can be bound and coreferential in a limited set of circumstances. Dubinsky and Hamilton (1998) cite the following sentences as examples of an epithet coreferring with an antecedent in the same sentence (p. 685):

- (48) I wanted Charlie to help me, but the bastard wouldn't do it.
- (49) Irving was besieged by a horde of bills and the poor guy couldn't pay them.
- (50) Although the bum tried to hit me, I couldn't really get too mad at Harry.

In these sentences, the epithets *the bastard*, *the poor guy* and *the bum* are acting like pronouns, coreferring with the R-expressions *Charlie*, *Irving*, and *Harry*, respectively. However, each of the epithets above is free (non-c-commanded) in its domain, so this is not a case of binding. The following examples show that in many cases, epithets may not be bound by a local DP (Yashima 2013, p. 19):

- (51) *John_i realizes that the sissy_i is going to lose.
- (52) *John_i thinks that I admire the idiot_i.

These data led many researchers to suspect that epithets are a specific type of pronoun which may not be locally bound. However Dubinsky and Hamilton (1998) show that this is not the case.

Dubinsky and Hamilton provide examples like the following, which show that epithets, unlike ordinary referential noun phrases, are not subject to Condition C (p. 687):

- (53) John_i ran over a man (who was) trying to give the idiot_i directions.
- (54) Through an accumulation of slipups, John_i (inadvertently) led his students to conclude that the idiot_i couldn't teach.

That variable binding, not only coreference, is available is illustrated by sentences like (55), where the referential DP is replaced with a quantified DP (quantificational phrases have no referent) (Yashima 2013, p. 20):

- (55) Every stupid driver_i ran over a man (who was) trying to give the idiot_i directions.

Structurally, these sentences are no different from (51) and (52) in that the epithet is c-commanded (bound) by its antecedent. Dubinsky and Hamilton claim that the key difference is that (51) and (52) represent logophoric contexts, where the antecedent is the perspective-bearer in the clause containing the epithet, while in (53)–(55), this is not the case – in other words, the environments where epithets may be bound are antilogophoric. The following sentences illustrate this contrast (Dubinsky and Hamilton 1998, p. 688):

- (56) a. *It was said by John_i that the idiot_i lost a thousand dollars on the slots.
 b. It was said of John_i that the idiot_i lost a thousand dollars on the slots.
- (57) a. *According to John_i, the idiot_i is married to a genius.
 b. Speaking of John_i, the idiot_i is married to a genius.
- (58) a. *John_i told us of a man (who was) trying to give the idiot_i directions.
 b. John_i ran over a man (who was) trying to give the idiot_i directions.
- (59) a. *Despite an accumulation of slipups, John_i asked his students to conclude that the idiot_i could teach.
 b. Through an accumulation of slipups, John_i (inadvertently) led his students to conclude that the idiot_i couldn't teach.

Note that in each example, the two sentences are structurally equivalent (particularly in terms of c-command and binding). The crucial difference is (anti)logophoricity: in each of the (a) examples, the antecedent is the bearer of perspective in the clause containing the epithet, while in the (b) examples, this is not the case.

In this respect, epithets contrast with R-expressions, which may not be bound even in antilogophoric environments (Dubinsky and Hamilton 1998, p. 688):

- (60) *John_i ran over a man (who was) trying to give the president_i directions.
 (61) *Through an accumulation of slipups, the teacher_i (inadvertently) led his students to conclude that John_i couldn't teach.

Furthermore, like standard Condition B pronouns, epithets may not be locally bound (p. 688):

- (62) *John_i shaved the idiot_i.
 (63) *John_i embarrassed the idiot_i.

These observations lead Dubinsky and Hamilton to conclude that epithets are Condition B pronouns, with the added restriction that they must obey the constraint, which they define as follows (p. 689):

- (64) *Antilogophoricity constraint for epithets*
 An epithet must not be anteceded by an individual from whose perspective the attributive content of the epithet is evaluated.

4.2.3 *Kare* as an epithet

Yashima (2013) provides evidence that *kare* is an epithet. In addition to the distributional evidence, there are semantic facts about *kare* which suggest that it is not simply a pronoun. First, note that all of the epithets discussed above (*the idiot*, *the bastard*, *the poor guy*, etc.) are expressives with some kind of attributive content – specifically, they all carry negative connotations. Yashima

notes (p. 22) that one fact about *kare*, which would be puzzling for a pronoun, is the fact that it is not used for social superiors, especially those who the speaker has some sort of personal relationship with. This is not surprising if *kare* is considered to be an epithet, with its associated negative attributive content. Furthermore, I mentioned above that *kare* and *kanozyo* have the implications ‘boyfriend’ and ‘girlfriend’ (respectively) in certain contexts. This, too, makes more sense if they are epithets rather than pronouns, since that kind of semantic drift is much more common with lexical than functional items. Thus, the semantics of *kare* suggest that it is more than simply a spell-out of ϕ -features. However, as with English epithets, this extra semantic content is expressive rather than descriptive (truth-conditional), and so it does not block blocking; see footnote 6 above.

Now let us turn to *kare*’s binding properties. In Section 4.1, I showed (45) and (46), repeated below, which demonstrate that *kare* is unable to function as a bound variable in unexpected situations, although it can corefer in the same contexts (Déchaine and Wiltschko 2002, pp. 417–418):

- (65) a. **Daremo_i-ga kare-no hahaoya-o aisite-iru*
 Everyone-NOM he-GEN mother-ACC love-PRES
 “Everyone loves his mother.”
 $\neq \forall x, x$ loves x ’s mother
- b. **Dono zyosei_i-mo [kanozyo_i-ga tensai-da to] omotte-iru*
 every woman-PART she-NOM genius-COP COMP think-PRES
 “Every woman thinks she is a genius.”
 $\neq \forall x, \text{woman}(x), x$ thinks x is a genius
- (66) a. *John_i-ga kare_i-no hahaoya-o aisite-iru*
 J.-NOM he-GEN mother-ACC love-PRES
 “John loves his mother.”
- b. *Mary_i-ga [kanozyo_i-ga tensai-da to] omotte-iru*
 M.-NOM she-NOM genius-COP COMP think-PRES
 “Mary thinks she is a genius.”

This evidence led D&W to conclude that *kare* must be pro-DP: it has the inherent semantics necessary to corefer, but as a semantic “constant” it cannot be a bound variable. However, in Section 4.2.1 I presented (47), repeated below, which demonstrated that in structurally similar situation, *kare* can in fact be bound (Yashima 2013, p. 13):

- (67) a. [*Dono nooberusyoo zyusyoo sakka*]_{*i*}-ga *kare_i-no kuruma-de*
 which Nobel.prize winning author-NOM he-GEN car-in
 kita-no?
 came-Q
 “Which Nobel Prize winning author came in his car?”
- b. *Dono gakusei_i-mo sensyuu kare_i-o suisensita sensei-ni*
 every student-PART last.week he-ACC recommended teacher-DAT

- orei-o okutta.
 present-ACC sent
 “Every student sent a present to the teacher who recommended him last week.”
- c. [Sono ondai-ni haitta] zyosi gakusei-no
 that music.college-DAT entered female student-GEN
 daremo_i-ga [kanozyo_i-no sainoo-o mottomo yoku hikidasite
 everyone-NOM she-GEN talent-ACC most fully bring.out
 kureru] sensei-ni dea-e-ta.
 do.the.favor teacher-DAT meet-can-PAST
 “Every female student who entered that music college was able to meet a teacher who could bring out her talent to the full extent.”

In (67), the pronouns are bound just as in (65). The difference is that, in (65), they are in logophoric contexts: *kare* is in a clause where its attributive content is evaluated from the antecedent’s point of view. On the other hand, the sentences in (67) have non-logophoric contexts, where that is not the case.

As additional evidence to show that *kare* can be bound in non-logophoric contexts, Yashima shows that when the antecedent is an R-expression, both coreference and binding interpretations are possible. This is evident with sentences like (68) (p. 25):

- (68) (Subete dansi gakusei no nakade) Taroo_i-dake-ga kare_i-no sainoo-o
 all male student of among T.-only-NOM he-GEN talent-ACC
 tatau hyookasuru sensei-ni deatta.
 highly admire teacher-DAT met
 “(Of all the male students) only Taro met a teacher who thought highly of his talent.”

Yashima notes that (68) has both strict and sloppy readings; in other words, both sentences in (69) are felicitous denials of (68) (p. 25):

- (69) a. Iya, boku-mo Taroo-no sainoo-o tatau hyookasuru
 no I-also T.-GEN talent-ACC highly admire
 sensei-ni deatta.
 teacher-DAT met
 “No, I also met a teacher who thought highly of Taro’s talent.”
- b. Iya, boku-mo boku-no sainoo-o tatau hyookasuru
 no I-also I-GEN talent-ACC highly admire
 sensei-ni deatta.
 teacher-DAT met
 “No, I also met a teacher who thought highly of my talent.”

The strict, coreferential reading, where *kare* refers to *Taro*, is denied by (69a). (69b), on the other hand, denies the reading where *kare* is a variable bound by the quantified DP *Taroo-dake*, which means “Taro is the only *x* such that *x* met a teacher who thought highly of *x*’s talent.” Therefore, it is clear that in

non-logophoric contexts like (68), *kare* may be either coreferential or bound. In this respect, *kare* patterns with the epithets we saw in Section 4.2.2.

However, we have already seen sentences like (66), where *kare* is able to corefer even though it is situated in a logophoric context. To further illustrate this point, Yashima notes sentences like the following (p. 10):

- (70) (Subeteno dansi gakusei no nakade) Taroo_i-dake-ga Hanako-ga
 all male student of among T.-only-NOM H.-NOM
 kare_i-o uttaeru to omotteiru.
 he-ACC sue COMP think
 “(Of all the male students) only Taro thinks Hanako will sue him.”

In this sentence, *kare* is in a logophoric context, and only a coreferential (strict) reading is possible; this is illustrated by the fact that only the first sentence in (71) is a felicitous denial of (70) (p. 11):

- (71) a. Iya, boku-mo Hanako-ga Taroo/kare-o uttaeru to omotteiru.
 no I-also H.-NOM T./he-ACC sue COMP think
 “No, I also think that Hanako will sue Taro/him.”
 b. Iya, boku-mo Hanako-ga boku-o uttaeru to omotteiru.
 no I-also H.-NOM I-ACC sue COMP think
 “No, I also think that Hanako will sue me.”

In (71a), the strict reading is denied, where *kare* corefers with *Taroo*; in contrast, (71b), which denies the sloppy bound variable reading, is infelicitous because there is no bound variable reading available here.

Yashima summarizes these observations as follows (p. 28):

- (72) Semantic binding of *kare/kanozyo* is subject to the antilogophoricity constraint, whereas coreferential interpretations of *kare/kanozyo* are not.

This contrasts with both ordinary Condition B pronouns, where both binding and coreference are possible regardless of logophoricity, and the epithets we saw in Section 4.2.2, with which binding and coreference are both constrained by antilogophoricity.

It is worth noting that even the antilogophoric environments in which *kare* can be bound are less restrictive than those for English epithets. Consider the English equivalents of the first two sentences of (67), with epithets replacing *kare*:

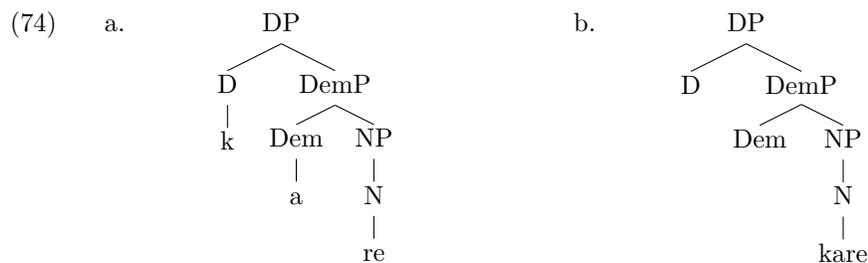
- (73) a. ??Which Nobel Prize winning author_i came in the bastard_i's car?
 b. ??Every student_i sent a present to the teacher who recommended the idiot_i.

Here, a covarying interpretation is marginal at best. This could be due to the evaluative content of the epithet clashing with the non-specificity of the quantified DP. For example, (73a) is interpretable only if the speaker considers

every Nobel Prize winning author to be a bastard; similarly, (73b) only works if every student in question is considered to be an idiot. Note that in (55), where the epithet could be bound with no problem, the presence of *stupid* in the antecedent ensures that the expressive connotations of *the idiot* are appropriate for every possible referent. Now the question arises as to why *kare* does not present the same difficulties. It seems that *kare* is associated with a different kind of evaluative content than English epithets. The only evidence for expressive content that Yashima notes is the fact that *kare* is typically not used for social superiors; thus, if *kare* has evaluative content, it relates only to the referent's position on some social scale, which can be easily presumed to be uniform for a given group of people. English epithets, on the other hand, evaluate the referent's intelligence, tolerability, and other such properties which tend to vary widely on an individual basis. Therefore, it is not surprising that English epithets are more resistant to binding than *kare* is.

4.2.4 DP and NP epithets

Yashima (2013) explains (72) with the claim that *kare* is categorically ambiguous: it may be either a DP epithet or an NP epithet. The syntactic structures of the two are as follows (p. 34):



DP epithets may be bound and they may corefer, but they must also obey the antilogophoricity constraint. NP epithets may also corefer, but they cannot be bound, and they are not restricted by the antilogophoricity constraint. The properties of the two types of epithets are summed up as follows (p. 35):

(75)

	antilogophoric	can be bound	can corefer
DP epithets	Yes	Yes	Yes
NP epithets	No	No	Yes

Yashima claims further that the reason DP epithets may be bound is that semantic binding of an epithet is possible only in the presence of a functional head; a functional head also brings about the requirement of antilogophoricity. Thus, a bare NP epithet may not be bound, and it is not subject to the antilogophoricity constraint.

Yashima's categorical ambiguity analysis is supported by the behaviour of *kare* when modified by a demonstrative, a functional head. The presence of a

demonstrative forces an NP interpretation of the epithet, because it must reside in Dem, which is filled in a DP epithet but empty in an NP. A Japanese DP containing the demonstrative *sono* ‘that (mesial),’ like its English counterpart, may be bound. When *kare* or *kanozō* is modified by *sono*, it acts like an antilogophoric pronoun, as shown by the contrast between the following sentences (pp. 35–36):

- (76) *?Daitooryoo-to kekkonsita zyosei_i-wa moto-supai-no
 president-with got-married woman-TOP former-spy-GEN
 sono-kanozō-ga hihansareru to omotteiru.
 that-she-NOM be.criticized COMP think
 “The woman who got married to the president thinks that she, who is
 a former spy, will be criticized.”
- (77) Daitooryoo-to kekkonsita zyosei_i-wa moto-supai-no
 president-with got.married woman-TOP former-spy-GEN
 sono-kanozō-ga hihansita sinbun-kizi-ni koogisita.
 that-she-GEN criticized newspaper-article-DAT protested
 “The woman who got married to the president protested a newspaper
 article that criticized her, who is a former spy.”

In the logophoric context of (76), both binding and coreference are impossible due to the presence of the demonstrative. If *sono* were absent, *kanozō* would be interpreted as an NP epithet, and the possibility of coreference together with a lack of the antilogophoricity constraint would make this example grammatical; however, the overt functional head forces the epithet to be antilogophoric. On the other hand, (77) is in a non-logophoric context, so coreference is possible. That a Dem+NP epithet combination may also be bound in a non-logophoric context is illustrated by (78) (p. 36), where the first *kare* is the lexical word meaning ‘boyfriend’ and the second is contained in a DP with the functional head *sono*, and is therefore bindable:

- (78) Dono kare_i-mo sono kare_i-o suisensita zyosei-to tukiatteiru.
 every he-PART that he-ACC recommended woman-with be.dating
 “Every boyfriend is dating a woman who recommended him.”

Furthermore, modification by an adjective, which is a lexical (rather than functional) head, does not impose an antilogophoricity requirement on an NP epithet, as shown by the following example (p. 37):

- (79) Taroo_i-wa wakai kare_i-ga riidaa-ni erabareru-beki da to
 T.-TOP young he-NOM leader-as be.selected-should COP COMP
 omotteiru.
 think
 (Lit.) “Taro thinks that young he should be selected as leader.”

In (79), *wakai kare* is in a logophoric environment, yet it may still corefer, as expected for an NP epithet.

To sum up: Yashima proposed that *kare* is ambiguous between an NP epithet and a DP epithet (in addition to the ‘boyfriend’ reading, which is clearly NP and not an epithet). Both kinds may corefer; the presence of a functional head (which is part of the DP epithet) both allows it to be bound, and introduces the antilogophoricity constraint. When *kare* is modified by a demonstrative or adjective, it must be NP, while a bare *kare* which is bound (and therefore, antilogophoric) must be DP – a bare, unbound *kare* is the only ambiguous case. The examples in this section fit this analysis.

4.3 ϕ P–NP Ambiguity

In Section 4.1 I showed arguments from Déchaine and Wiltschko (2002) which claim that *kare* is pro-NP. However, this was based on the belief that *kare* cannot function as a bound variable, which Yashima (2013) showed to be false. As presented above, Yashima claims that *kare* is an epithet which is ambiguous between two structures, each with its own properties: the DP epithet may corefer, may be bound, and is subject to the antilogophoricity constraint; the NP epithet may corefer, but it may not function as a bound variable, and it has no antilogophoricity constraint. Here I will show that what D&W had called pro-NP was actually the NP epithet incarnation of *kare*, and that in terms of their model, the DP epithet *kare* can be analyzed as a special epithet pro- ϕ P; this leaves us with a ϕ P-NP ambiguity, which was previously unseen. However, recall that in Section 3.2 I showed that D&W’s argumental ϕ Ps are in fact best analyzed as DP; under this view, we are confronted with another ambiguity, one between DP and NP, the first of its kind in the unified model. However, recall that I am trying to eliminate all systematic categorial ambiguity from the model; to this end, I will adopt the type-shifting rules from Chierchia (1998). In this way, this ambiguity, too, can be eliminated in favour of an analysis where both forms of *kare* are pro-NP, due to the fact that Japanese lacks the category D and its NPs are argumental, rather than predicative. In this way, all systematic ambiguities can be removed from the unified model. I will begin working toward this conclusion by translating Yashima’s work into the framework of D&W.

As explained in Section 4.1, D&W analyzed *kare* as pro-NP on the basis of its noun-like distribution, its lack of bindability and the fact that it can corefer. Pro-NPs have the distribution of a bare noun; NP epithet *kare* satisfies this requirement in that it can be modified by an attributive adjective, a possessor or a demonstrative. Semantically, pro-NPs are constants, meaning they cannot function as bound variables, as is the case with NP epithets. The coreference properties of pro-NPs depend on the inherent semantics of the pronoun; since *kare* has the features [MALE] and [MARRIAGEABLE AGE], it has coreference properties and as expected, it can corefer. Therefore, it seems unproblematic to claim that the NP epithet version of *kare* is a kind of pro-NP. However, it is still an epithet, as is apparent from the facts that it undergoes semantic drift (as mentioned above, *kare* means ‘boyfriend’ in certain contexts, which is easier to explain if it has lexical content, rather than the pure functional content of a pronoun), and that it may not be used for people of higher social status,

suggesting it has the expressive content expected of an epithet. In other words, NP epithets like *kare* are a subtype of pro-NP.

Now we turn to the DP epithet version of *kare*. In terms of D&W’s model, this version should be classified as pro- ϕ P, rather than pro-DP. Recall from Section 2.1 that pro-DPs have the distribution of an argument, definite semantics, and the binding-theoretic status of an R-expression, meaning they cannot be bound. Pro- ϕ P, on the other hand, may be an argument or predicate, and they function as bound variables. We have seen no evidence that *kare* may function as a predicate; it has been an argument in every context we have observed so far – however, this does not decide between the two categories, since pro- ϕ P may function as either argument or predicate. The crucial fact here is that it may function as a bound variable, which only pro- ϕ P may do. Thus, it seems clear that Yashima’s “DP epithet” reading of *kare* is pro- ϕ P in D&W’s model. Again, the fact that it is an epithet imposes some further restrictions on it, including the social restrictions noted above, and in this case, the added restriction of the antilogophoricity constraint.

In Section 3.1 I discussed the fact that in D&W’s system, there were systematic ambiguities – many pronouns, such as English and French 1st and 2nd person pronouns, seemed to be ambiguous between the categories DP and ϕ P. I noted that, in these languages, there were no pronouns straddling the line between ϕ P and NP. However, this seems to be the case for *kare*. This raises the question of what is different about *kare*, or perhaps the Japanese language in general, which causes there to be this ambiguity where we had not seen it before; before getting to this, however, I will rework the preceding analysis in the framework of the unified model.

4.4 DP–NP Ambiguity

Recall that in Section 3.2 I showed that if binding theory is left out of the picture, and if all type e individuals are analyzed as DP and ϕ P is unambiguously predicative, then the models of Déchaine and Wiltschko (2002) and Zamparelli (2000) converge. Thus, in the unified model, the properties of the different categories are as follows: Pro-DP is type e and therefore has the distribution of an argument; it may corefer, and in many cases (but not all) it may function as a bound variable; and it has the syntax of a full DP, meaning that it cannot be modified by articles, adjectives, etc. Pro-PredP is used for type $\langle e, t \rangle$ predicates; it may corefer or perhaps be bound, but it may only enter this sort of relationship with other predicates, not individuals. Pro-NP is used for semantic individuals denoting types, and they may corefer with other types, but it may not be a bound variable, and it typically needs to be embedded in some DP. In this model, Yashima’s NP epithets are still analyzed as pro-NP, but his DP epithets are better analyzed as pro-DP than pro-PredP (unlike in D&W’s model, where they are best analyzed as pro- ϕ P).

In the unified model, it makes no sense to analyze the DP epithet reading of *kare* as pro-PredP. This is because it is always an argument, acting as a type e individual rather than a predicate. It may be bound, as is expected for

pro-DP; the fact that this binding comes with an antilogophoricity constraint is unsurprising, given its epithet status and the claims of Dubinsky and Hamilton (1998). On the other hand, the NP epithet reading of *kare* is still best analyzed as pro-NP, on the basis of its semantic type, its distribution, and its binding properties. The NP epithet *kare* still corefers with individuals, meaning it is semantic type *e* (and therefore, not pro-PredP). It may co-occur with demonstratives, possessors and attributive adjectives, which is possible for pro-NP but not pro-DP. Finally, it may not be bound (unless it is embedded in a DP which can be bound for other reasons, such as the presence of a demonstrative like *sono* ‘that’), as expected for pro-NP.

Now we are presented with another kind of ambiguity which we have not seen before: an ambiguity between DP and NP. Note that the DP- ϕ P ambiguities discussed above are no longer ambiguities in the unified model: the pronouns in question are unambiguously pro-DP, and the apparent categorical ambiguities actually depend on the binding properties of the pronoun, which have some non-structural explanation. Now, however, we see a case of a DP-NP ambiguity in the unified model. That we would see ambiguities between these types, skipping over PredP, is unsurprising given the fact that DP and NP are both semantic type *e*, unlike PredP. That being the case, it may come as a surprise that we have not seen any DP-NP ambiguities before now. The next step is to explore the cross-linguistic distribution of this sort of ambiguity and, if they prove systematic in some languages, to determine whether they are caused by some properties of those languages.

4.5 Pronouns in D-less Languages

Up to this point I have been exploring the nature of the ambiguity of *kare*, as described by Yashima (2013). First I recast his analysis in terms of D&W’s model, and then in the unified model, where it is best analyzed as a DP-NP ambiguity. I have been trying to eliminate all such systematic ambiguities from the system, and here I will show that, using Chierchia’s (1998) Nominal Mapping Parameter, this can in fact be accomplished. What appears to be a categorical ambiguity between DP and NP is in fact unambiguously NP, which may undergo type-shifting processes to obtain properties of DP. This kind of type-shifting is available only in languages lacking the DP projection entirely, including Mandarin and Korean (but excluding languages like Turkish, which lack articles but still have DPs, as we will see). Thus, it can be predicted which languages will have the same apparent ambiguity as Japanese, where pronouns have properties of both NP and DP.

Chierchia (1998) proposes a theory of the semantics of NPs which describes how different languages map NP to different types of interpretations. Languages may map NP onto arguments, predicates, or both; semantic type-shifting operations may shift the interpretation, but these operations are only permitted when there is no overt morpheme to perform the same function. This explains why Japanese NPs (including *kare*) may take on DP-like qualities in certain situations; it also predicts that other languages with no determiners or overt

number marking will also have pronouns which have both NP and DP properties. This prediction is borne out, as Mandarin and Korean pronouns pattern with Japanese. In this section I will summarize Chierchia's theory, and show data from Mandarin, Korean, and Turkish which support it.

Chierchia claims that the interpretation of N (and therefore NP) in a language is determined by two features: $[\pm\text{arg}]$ and $[\pm\text{pred}]$. These features should be understood as parameter settings for a language, not as features of a particular syntactic head. If a language is $[\text{+arg}]$, NPs may be mapped into arguments (specifically, for common nouns, they are mapped onto kinds), while if a language is $[\text{+pred}]$, its NPs may be mapped onto properties of type $\langle e, t \rangle$. Chierchia discusses a number of operators for shifting semantic types; there are two type-shifting operators in particular, " \cup " and " \cap ", which are available in the lexicon to shift back and forth between these two types (kinds to predicates and predicates to kinds, respectively). Every combination of values for the two features is predicted to be possible except NP $[-\text{arg}, -\text{pred}]$ (which would mean that NPs are uninterpretable), and all combinations are attested in well-known languages. Chierchia refers to crosslinguistic variation in the settings of this pair of features as the Nominal Mapping Parameter.

A language of type NP $[\text{+arg}, -\text{pred}]$ would have all its NPs referring to kinds. It would allow all its bare NPs to be arguments. In addition, Chierchia shows that kinds have a mass (as opposed to count) denotation, which is in complementary distribution to plurality; therefore, in NP $[\text{+arg}, -\text{pred}]$ languages, all nouns will be mass nouns, and numerals with nouns will be impossible without a classifier to make the noun countable. Mandarin and Japanese fit the description of this kind of language, since they allow bare NPs to appear in argument positions and require classifiers with their numerals.

In NP $[-\text{arg}, \text{+pred}]$ languages, all NPs enter the derivation as predicates and so they may not appear as arguments without some element of category D. In this type of language, one would expect some nouns to have count extensions and others to be mass, so these languages are predicted to have plural marking for the count nouns. It is also expected that the determiners that license NPs in argument positions may be null in some cases; however, null morphemes need some kind of licensing, so there would be restrictions on the distribution of "bare" NPs. Chierchia claims that French and Italian are examples of NP $[-\text{arg}, \text{+pred}]$ languages, where French has no null D (so every argument has some overt D), while Italian has a null D, so "bare" NP arguments are permitted only in a restricted set of environments, specifically those in which the null D is licensed by a lexical head, following the Empty Category Principle (ECP; see Chomsky 1981).

The final type, NP $[\text{+arg}, \text{+pred}]$, allows NPs to be either arguments or predicates. If an NP is realized by an argument, it may be predicativized by the operator " \cup ", which yields a mass interpretation. This predicts that all bare nominal arguments will be mass nouns. On the other hand, if a noun is a predicate, it will have an atomic extension, and will thus be a count noun (since mass nouns are derived more economically by being realized directly as arguments). So, count nouns will be able to take plural marking, and will need

some determiner to be realized as arguments. In addition, the operator “ \cap ” may apply to plural count nouns (Chierchia shows elsewhere that this operation applied to a singular nouns does not yield kinds) to give a kind interpretation; therefore, bare plurals may also appear as arguments. English is a language of this type.

Before turning to how this theory relates to pronoun types, I will go over some of the other type-shifting operations Chierchia discusses. In particular, I will look at ι and ID. The operator ι “selects the greatest element from the extension of a predicate and constitutes typically the meaning of the definite article” (Chierchia 1998, p. 359); ID is the inverse of this operation, turning a specific entity back into a property. These operators are like “ \cap ” and “ \cup ”, respectively, in that they shift back and forth between elements of type e and $\langle e, t \rangle$; however, unlike “ \cap ” which yields kinds, the ι -operator creates a particular unique and salient entity. Chierchia claims that these (and other operators) are available only when there is no overt morpheme which performs the same function. In other words, ι may only apply to NP in languages which have no definite article. Thus, one expects a bare NP to pick out a specific individual only when there is no definite article to do so.

The connection to pronoun types is as follows: In an NP[+arg, –pred] language, one would expect that pronouns (like other arguments) realize NP. Given the unavailability of any definite article, the equivalent operator ι is available to give these NPs the semantics of DP.¹³ Thus, one would expect that in these languages, pronouns have properties of both NPs and DPs. So, unlike in English or Romance, these languages would have the same pronominal forms able to be modified by attributive adjectives (like NPs) and able to be bound variables (like DPs). As shown above, this is the case for Japanese; this pattern is also attested in both Mandarin and Korean.¹⁴ These are both NP[+arg, –pred] languages which allow bare NP arguments and have numerical classifiers instead of plural marking. As shown in (80), Mandarin pronouns take attributive adjectives just like common nouns, and in (81) we see that personal pronouns may be bound.¹⁵

- (80) a. Nian qing de nü hai'r bu hui kai che.
 young PRT girl not know.how drive
 “Young girls don’t know how to drive.”
 b. Nian qing de wo bu hui kai che.
 young PRT 1SG not know.how drive
 “(lit.) Young me didn’t know how to drive.”

- (81) Mei ge ren_i dou ren wei ta_i zui cong ming.
 every CL person PART believes 3SG most smart
 “Everyone_i thinks he_i is the smartest.”

¹³This operator must apply before “ \cap ” in the lexicon; like “ \cap ,” ι applies to predicates, not entities.

¹⁴Thanks to Meng Yang and Chung-hye Han for their generosity in providing data on Mandarin and Korean, respectively.

¹⁵Tones are omitted from the Mandarin data.

Likewise, Korean allows modification of its pronouns:

- (82) a. cak-un ku
 small-ADNOMINAL he
 “small he”
 b. mantwu-lul manhi mek-un ku
 dumpling-ACC much eat-ADNOMINAL he
 “he who ate lots of dumplings”

Binding of pronouns in Korean is less straightforward, but some speakers allow a bound variable interpretation in sentences like (83), and binding of fake indexicals is possible in certain configurations as well (Chung-hye Han, p.c.):

- (83) a. %Motwu-ka_i nongkwucang-eyse ku-uy_i umloyswu-lul
 Everyone-NOM basketball.court-at he-GEN beverage-ACC
 masi-ess-ta.
 drink-PAST-DECL
 “Everyone_i drank his_i water at the basketball court.”
 b. %Motwu-ka_i ku-ka_i nolay-lul cal ha-n-tako
 Everyone-NOM he-NOM singing-ACC well do-PRES-COMP
 sayngkakha-n-ta.
 think-PRES-DECL
 “Everyone_i thinks that he_i sings well.”
- (84) Na-man_i nay-ka_i ihayha-n cilmwn-ul pat-ass-ta.
 I-only I-NOM understand-ADNOM question-ACC got-PAST-DECL
 ‘Only I_i got a question I_i understood.’

Note that the pronouns in Mandarin and Korean are unlike those in Japanese in that, for speakers who allow binding, there are no antilogophoricity constraints. This can be explained if Mandarin and Korean pronouns are straightforwardly pronouns, rather than epithets; for some reason, unlike these two languages, the only pronouns in Japanese also happen to be epithets. Thus, the predictions of Chierchia’s system are borne out.¹⁶

To show that this behavior is not simply due to a lack of D, but to the interpretation of NPs in the language, I will consider Turkish. Chierchia discusses a class of languages which are NP[+arg, +pred], but unlike English, they lack overt articles. In these languages, like English, bare NPs may be either predicates or arguments, and the former will have a count interpretation and the latter will be mass nouns. These languages will also have plural marking. Unlike English, however, they will be able to make use of the *ι*-operator, so that bare NPs will be freely available with a definite meaning. Chierchia notes

¹⁶It could be argued that this clustering of properties in Japanese, Mandarin and Korean is due to an areal relationship between the three, rather than their parameter settings with respect to the interpretation of NPs. Unfortunately, there are few well-described languages outside East Asia lacking determiners or plural marking to test this. Some candidate languages in which this could be tested include Bini (Nigeria), Canela-Krahô and Pirahã (Brazil) (Dryer 2011a,b).

that Russian and other Slavic languages¹⁷ fall into this category; Turkish does as well. It turns out that pronouns in these languages pattern like those in English, and other [+pred] languages, in that their pronouns may not act as NPs; as shown in (85), Turkish pronouns may not take attributive modifiers:¹⁸

- (85) *Küçük o pizza ye-di.
 small 3SG pizza eat-PAST
 “(intended lit.) Young he ate pizza.”

Modification of the pronoun is impossible, suggesting that pronouns in Turkish, like English, are pro-DP, or at least some projection larger than NP.

The morphological shapes of pronouns in [+pred] and [-pred] languages show a contrast which supports this analysis. Neeleman and Szendrői (2007) notes that the way that such features as number and case are realized on a pronoun can indicate the amount of material that the pronoun realizes. If a feature is realized as an agglutinative affix on a pronoun, then the associated projection is not a part of the pronoun itself; features realized suppletively are a part of the pronoun’s own projection. Consider the Mandarin and Turkish pronoun inventories:¹⁹

- | | | | | |
|------|----|-----------------|----|----------------|
| (86) | a. | <i>Mandarin</i> | b. | <i>Turkish</i> |
| | | wo – 1SG | | ben – 1SG |
| | | women – 1PL | | biz – 1PL |
| | | ni – 2SG | | sen – 2SG |
| | | nimen – 2PL | | siz – 2PL |
| | | ta – 3SG | | o – 3SG |
| | | tamen – 3PL | | onlar – 3PL |

Note that Mandarin plural pronouns are formed by adding the plural suffix *-men* to their singular counterparts, while Turkish plural pronouns are formed suppletively. This suggests that Mandarin pronouns realize NP, as predicted from the language’s NP[+arg, –pred] parameter setting. In contrast, Turkish pronouns must realize some larger projection, at least NumP (how NumP relates to the other projections in the unified model is unclear at present, although both Zamparelli and D&W associate plurality with the middle projection, PredP).

Thus, it seems that the only languages in which pronouns may have the properties of NP are the [-pred] languages, such as Mandarin and Japanese. In all [+pred] languages, regardless of the setting of [\pm arg], and regardless of whether or not overt articles appear in the language, pronouns must realize

¹⁷Serbo-Croatian, another Slavic language, patterns with Japanese and Mandarin (and not Turkish and other Slavic languages) in that its pronouns may have NP-like properties (Željko Bošković, p.c.). This is problematic for the present analysis if it turns out that Serbo-Croatian behaves otherwise as a [+pred] language; however, further pursuit of this problem is outside the scope of this paper.

¹⁸Thanks to Ayşegül Kutlu for her generosity in providing Turkish data.

¹⁹Mandarin data from Neeleman and Szendrői (2007), Turkish data from Ayşegül Kutlu (p.c.).

some larger projection and have the properties of DP. To put it differently, since there is a correlation between [+pred] and the possibility of count nouns, the possibility for pro-NP to take on DP-like qualities exists only in languages where all nouns are mass – there is an inverse correlation between the presence of number marking and the possibility of this type shift. Exactly why this should be the case is unclear. It may be the case that the category D (and the projection DP) are available in all and only the [+pred] languages, regardless of whether or not there are overt articles of this category. In this analysis, pronouns must be DP if the language has DP, and they are NP otherwise. However, further exploration of this issue requires empirical research which is beyond the scope of this paper.²⁰

In addition to confirming Chierchia’s predictions about the interpretations of NPs in different languages, the discussion above also explains the apparent DP-NP ambiguity noted in Section 4.4. There, it was noted that such an ambiguity is strange because it is not attested in any of the languages D&W examine; furthermore, it was the only remaining systematic ambiguity, since the DP- ϕ P ambiguities were eliminated in the unified model. In fact, this ambiguity, too, is only apparent: It is not the case that these pronouns are ambiguous between NP and DP. Instead, categorically they are unambiguously NP, but the absence of determiners in these languages allow type-shifting, so that NPs can take on the properties of DPs. In this way, all systematic categorical ambiguities can be eliminated from the unified model, and the behavior of NPs and DPs can be determined by other properties of the language in question.

5 A Feature Geometry for Pronouns

In this section I will start by introducing one final model of the noun phrase with its own pronoun typology, that of Cardinaletti and Starke (1994). This typology is based on a set of features orthogonal to those discussed above, and the model of the the noun phrase is incompatible with the unified model I developed in Section 3; however, the implications and properties that they discuss cannot be ignored. I am forced to conclude that the properties discussed in Cardinaletti and Starke (1994) are largely prosodic, and thus orthogonal to the syntactic conclusions discussed up to this point; instead, they must have some non-structural explanation. I conclude that there is a sort of feature geometry, in the sense that some (syntactic, semantic, prosodic, etc.) features are dependent on others in an implicational hierarchy. To this end, I will start by describing the system of Cardinaletti and Starke (1994), including the features under discussion

²⁰An alternative to this Nominal Mapping Parameter-based approach would be to follow Bošković (2012) in his claim that certain languages lack the D projection entirely, and combine it with Chierchia’s (1998) notion of blocking, such that operators like ι are available only in languages lacking the corresponding determiner. Like the present analysis, this would predict that Japanese, Mandarin, and Korean pronouns should have properties of both NP and DP; however, Bošković also analyzes Turkish and Slavic languages as D-less, while as I have shown, Turkish (and, following Chierchia, at least some Slavic languages) patterns with English and other DP languages in the behaviour of its pronouns.

and the model they propose. Then I will discuss some pronominal feature geometries in the literature (Harley and Ritter 2002; Cowper and Hall 2009) and show that, while they may be correct, they are lacking the sorts of prosodic features involved in Cardinaletti and Starke (1994). Finally, I will present my own proposal, formulating a feature geometry to overlay on the unified model.

5.1 Cardinaletti and Starke (1994)

Here I will describe the system of Cardinaletti and Starke (1994) (henceforth C&S). I will start by describing the deficiencies that the model is meant to account for. Then I will present the structure they propose for the noun phrase. Finally, I will show that what C&S describe are best explained in terms of prosodic features, and I will discuss the relationship between these features and the unified model.

5.1.1 The deficiencies

C&S discuss a number of properties of various types which divide pronouns into three types: strong, weak, and clitic pronouns. Clitic pronouns are deficient with respect to weak pronouns, which are in turn deficient compared to strong pronouns. The set of properties in question includes distributional, interpretational, prosodic, and morphological features. Most of these properties distinguish weak (and therefore clitic; weak and clitic pronouns are collectively referred to as deficient pronouns) from strong pronouns; only a small set of prosodic and distributional features set clitic pronouns apart from their weak counterparts.

The first set of properties setting strong pronouns apart from deficient pronouns involves syntactic distribution. First, they note that deficient pronouns may not occur in the base-generated, θ -position; the following examples show base and derived positions for the subject, indirect object and direct object, respectively, in Italian (Cardinaletti and Starke 1994, p. 46):²¹

- (87) a. {essa_D; lei_S; Maria} forse l'ha fatto {*essa_D; lei_S; Maria}
 {it_D; she_S; M.} maybe it.has done {it_D; she_S; M.}
 da sola
 DA alone
- b. Non dirò mai {loro_D; *a loro_S; *a Gianni} tutto
 not I.will.say never {them_D; to them_S; to G.} everything
 {*loro_D; a loro_S; a Gianni}
 {them_D; to them_S; to G.}
- c. Gianni {li_D; *loro_S; *questi studenti} stima {*li_D;
 G. {them_D; them_S; these students} estimates {them_D;
 loro_S; questi studenti}
 them_S; these students}

²¹In these examples, a subscripted D indicates a deficient pronoun, S indicates a strong pronoun. Non-pronominal noun phrases are included for comparison.

As (87) shows, deficient pronouns are disallowed from base-generated positions, while strong pronouns pattern with non-pronominal nominals. The same pattern holds for peripheral positions; (88) shows that deficient pronouns are disallowed in the context of clefts, left dislocation, right dislocation, and isolation, respectively (p. 47):

- (88) a. È { *ess_D; le_S; Maria } che è bella
 it.is { 3.SG.FM_D; 3.SG.FM_S; M. } that is pretty
 b. { *ess_D; le_S; Maria }, lei è bella
 { 3.SG.FM_D; 3.SG.FM_S; M. }, she/it is pretty
 c. *pro* arriverà presto, { *ess_D; le_S; Maria }
 she/it will.arive soon, { 3.SG.FM_D; 3.SG.FM_S; M. }
 d. Chi è bella? { *ess_D; le_S; Maria }
 who is pretty? { 3.SG.FM_D; 3.SG.FM_S; M. }

Another contrast is that strong pronouns, but not weak pronouns, allow c-modification (modification of the projection as a whole, as opposed to NP-internal modification like attributive adjectives and relative clauses), as the following French example shows (*il* is analyzed as deficient, *lui* as strong) (p. 47):

- (89) a. { vraiment; seulement; ... } lui/*il
 b. lui/*il { seul; aussi; ... }

The final distributional contrast is that strong, but not deficient, pronouns may be coordinated (again from French; *elles* is strong, *les* is deficient) (p. 46):²²

- (90) a. Jean les trouve belles
 J. them.FEM finds pretty
 b. *Jean les et celles d'à côté trouve belles
 J. them.FEM and those besides finds pretty
 c. Jean trouve elles belles
 J. finds them.FEM pretty
 d. Jean trouve elles et celles d'à côté belles
 J. finds themFEM and those besides pretty

C&S sum up these distributional asymmetries by proposing that deficient, but not strong, pronouns must occur in a special, derived position (as opposed to a base-generated or peripheral position), and that they are incompatible with c-modification and coordination.

In terms of semantic interpretation, C&S note a number of semantic properties which distinguish strong from deficient pronouns. They note that while strong pronouns may freely occur with contrastive stress or ostension (non-linguistic pointing), deficient pronouns may do so only if the referent is an entity which is already prominent in the discourse. Thus, (91a), which has a deficient pronoun and no discourse antecedent, contrasts with (91b), which has

²²C&S note that (90c) is grammatical only if *elles* is construed as contrastive.

no antecedent but a strong pronoun, and the rest of (91), where a deficient pronoun is grammatical with a discourse antecedent (p. 49):²³

- (91) a. *J' ai vu Marie puis je ⇒ I' ai vu
 I have seen M. then I her have seen
 b. J' ai vu Marie puis j' ai vu ⇒ elle
 I have seen M. then I have seen her
 c. Mets-toi ici et regardes cette maison. Tu ⇒ la vois bien
 Come here and look.at this house. You it see well
 maintenant?
 now?
 d. Mais, tu ne vois donc pas ce livre? Bien sûr que je ⇒ le
 But, you don't see this book? Of course that I it
 vois.
 see

They also note that strong pronouns are disallowed in expletive, or non-referential, positions (even if it doubles a deficient pronoun) (p. 50):

- (92) a. Il est arrivé un grand malheur.
 he is arrived a big disaster
 b. *Lui (il) est arrivé un grand malheur.
 he (he) is arrived a big disaster
 c. Il pleut.
 he rains
 d. *Lui (il) pleut.
 he (he) rains

Similarly, deficient, but not strong, pronouns may have an impersonal reading. In (93), deficient *on* and *ils* may receive either an impersonal or referential interpretation, while strong *lui* may only be interpreted referentially (p. 51):

- (93) a. **On** t' a vendu un livre pas cher.
 theyNONREF/wEREF you have sold a book not expensive
 b. **Ils** m' ont vendu un livre pas cher.
 They me have sold a book not expensive
 c. **Eux** (ils) m' ont vendu un livre pas cher.
 They (they) me have sold a book not expensive

They also note that in non-referential dative constructions as in (94) (examples from French and Slovak), where the pronoun acts as a discourse marker and has no referent, deficient pronouns (the pre-verbal clitics) are allowed, but strong pronouns (the phrase-final pronouns in (94)) are disallowed (p. 51):²⁴

²³An arrow indicates ostension.

²⁴C&S translate both sentences in (94) as roughly meaning “By Joves, I’ll give him a blow he’ll remember!”

- (94) a. Je vais (**te**) lui foutre une de ces claques (***à toi**)!
 I will (you) him give one of these smacks (to you)
- b. Ja (**ti**) mu dám takú facku (***tebe**)!
 I.will (you) him give such a.smack (to you)

The final semantic distinction C&S note is the fact that strong pronouns may refer only to human referents, while deficient pronouns may refer to either human or non-human entities.

These semantic properties are not unrelated. C&S note that expletive, impersonal, and non-referential dative pronouns are similar in that they have no referent; the fact that strong pronouns may not occur in these contexts suggests that they must be referential in some way that deficient pronouns do not require. In a similar vein, the fact that strong pronouns may refer without a discourse antecedent is another sign of their “stronger” referentiality with respect to deficient pronouns. C&S conclude that strong pronouns always bear their own range-restriction (a restriction on the range of possible referents); deficient pronouns may not bear their own range restriction, and are either rangeless (non-referential) or receive their restriction from an element prominent in the discourse (p. 54).²⁵ They speculate that the final distinction, the requirement that strong pronouns have animate referents, may indicate that [+human] is the default range (cross-linguistically) if nothing else restricts the range.

The next set of properties contrasting strong from deficient pronouns involve prosodic restructuring. Specifically, deficient, but not strong, pronouns are subject to processes such as sandhi rules. For example, French liaison appears possible only with deficient pronouns; in (95a), *elles* may be deficient, but in complex inversion contexts such as (95b), its (homophonous) strong counterpart is required and liaison is impossible (underlining indicates the consonant is pronounced, meaning liaison has occurred) (p. 55):

- (95) a. Elles ont dit la vérité.
 they have spoken the truth
- b. *Quand elles ont-elles dit la vérité?
 when they have-they spoken the truth?

Furthermore, deficient pronouns may form a single prosodic unit with an adjacent word, while strong pronouns (like non-pronominal nominals) may not, as shown in (96), where underlining indicates prosodic domains (p. 55):

- (96) a. A mange beaucoup.
 A. eats a lot
- b. ⇒Lui mange beaucoup.
 he eats a lot
- c. Il mange beaucoup.
 he eats a lot

²⁵In terms of Kratzer’s (2009) binding theory, it appears that deficient pronouns are minimal pronouns (born without ϕ -features, only receiving these features from other elements in the discourse).

Finally, deficient pronouns, unlike strong pronouns, may undergo phonological reduction, as the following English example shows (where *you* is ambiguous between a strong and weak pronoun, but coordination forces a strong pronoun reading) (p. 55):

- (97) a. I saw **'ya** in the garden.
 b. *I saw **'ya** and John in the garden.
 c. I saw **you** and John in the garden.

Each of the processes here are varieties of prosodic restructuring; the generalization is that deficient pronouns may undergo such restructuring processes, and strong pronouns may not.

The final distinction that C&S note between strong and deficient pronouns is morphological. The two forms are often homophonous (like English *you*, discussed above), and if there are distinct forms, they are most often suppletive (like French deficient *il* versus strong *lui*). However, there are cases where seemingly transparent morphology distinguishes the two, as in the Italian contrasting deficient/strong pair *loro/a loro*, and Slovak *ho/jeho*. From these examples, C&S conclude that the morphology of deficient pronouns is less than or equal to that of strong pronouns.

The pronominal typology of C&S is tripartite; there are a few properties which further divide the deficient class of pronouns into weak and clitic pronouns. The first of these has to do with distribution: Weak pronouns appear in positions where full projections (XP) may appear, while clitic pronouns appear in head positions. For examples, in Olang Tirolese (a dialect of German), the weak pronoun *es* may appear in V2-initial position, while the clitic *s* may not (p. 60):

- (98) a. **Es** isch toire.
 it is expensive
 b. ***S** isch toire.
 it is expensive

In addition, the Italian clitic pronoun *gli*, unlike its weak counterpart *loro*, may (and, in fact, must) be merged with, and picked up by, the verb (p. 60):

- (99) a. Non **loro* dirò mai **loro** tutto.
 No to.them I.will.say never to.them everything
 b. Non **gli** dirò mai **gli* tutto.
 No to.him I.will.saw never to.him everything

Finally, the sentence-initial position in Slovak is generally reserved for topicalized and subject XPs, and so the weak *ono* but not clitic *mu* may appear in this position (p. 60):

- (100) a. **Ono** *mu* *to* bude pomáhat'.
 it to.him it will help

	morphology	distribution		interpretation	prosody		X-bar
	+reduced	in FP at SS	*coord, ...	no range	reduction rules	no word-stress	X ^o
clitic	1	+	+	+	+	+	+
weak	2	+	+	+	+	-	-
strong	3	-	-	-	-	-	-

Table 2: Properties of the three types of pronouns (Cardinaletti and Starke 1994, p. 69)

- b. ***Mu** to bude pomáhat'.
to.him it will help

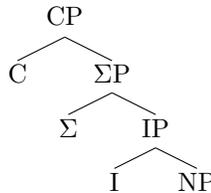
C&S conclude that the distributional differences between clitic and weak pronouns may be captured if the former are considered to be syntactic heads, while the latter are full projections. They also note a prosodic difference between the two: Weak, but not clitic pronouns, may receive lexical word-stress. Finally, clitic pronouns are morphologically reduced with respect to weak pronouns, if there is a morphological difference between the two.

Thus, the generalizations that C&S make about the distinctions between the pronoun types are as in Table 2. Notice that the only factors distinguishing clitic from weak pronouns relate to word stress and X-bar status; in every other way they pattern with the weak pronouns. Next, I will discuss how C&S explain this tripartition structurally.

5.1.2 The model

To explain these deficiency relationships between the different types of pronouns, C&S claim that weak pronouns are lacking some syntactic head which houses all the properties which weak pronouns are lacking with respect to strong pronouns; an additional head is missing in clitic pronouns. The reasoning for this kind of explanation comes from the morphological relationships between the three kinds, mentioned above. The result is a three-layered model of the nominal projection, similar to those of Déchaine and Wiltschko (2002) and Zamparelli (2000), discussed in Section 2. C&S's proposed model is as in (101) (p. 86):

(101)



Strong pronouns (and non-pronominal nominals) realize the entire structure of (101). Weak pronouns are missing the CP layer, and clitic pronouns are missing both CP and ΣP, having a maximal projection of IP. C&S do not discuss the properties of I; they include it mainly to reinforce the symmetry between nominal and clausal projections.

In considering the properties of C, the element present only in strong pronouns, C&S base their analysis largely on the assumption that in the Italian strong pronoun *a loro*, the morpheme *a*, a “dummy marker,” is a transparent realization of this head. Noting a resemblance between this morpheme and the dummy marker in complex complementizers (e.g. the *of* in *instead of*), they label this head C(omplementizer). In terms of syntactic properties of C, the fact that weak pronouns (lacking C) must surface in some derived position suggests that C houses the projection’s Case feature, meaning that elements lacking C must move to a position where they can inherit a functional Case feature, which generally means the specifier of some AgrP. Thus, C is similar to the K (case) head proposed by some authors. Semantically, C&S claim that C holds the referential index, resulting in the range restrictions on strong pronouns (and hence their ability to refer on their own, and inability to have non-referential functions like impersonal pronouns). In the absence of other range features, C receives the default range [+human] to avoid having a featureless head. They also suggest that index is the interpretation of Case, so that a single Case feature explains all the syntactic and semantic facts about strong pronouns. Finally, they suggest that the presence of CP makes a pronoun a sufficiently large syntactic constituent to resist prosodic restructuring processes.

To define the properties of Σ , present in weak but not clitic pronouns, C&S consider the Slovak strong-clitic pairs *je-ho*, *ho* and *je-mu*, *mu*. Since these are some of the only examples of a transparent morphological relationship between clitic and non-clitic pronouns, and there is no reason to treat the *je-* prefix as the dummy marker C, they consider this prefix to be an overt realization of Σ . They note that the same morpheme appears on the copula in emphatic, affirmative contexts. Given this parallel, they consider the two to be the same morpheme, which appears to have polarity (affirmative versus negative) and focus features, as well as prosodic features. Since clitic pronouns are lacking this Σ , they must associate with some prosodic features, and the only way to do this is head-movement resulting in incorporation; hence, clitic pronouns must surface in some X° position. And since Σ holds the projection’s prosodic features, clitics cannot receive lexical stress.

5.1.3 Discussion

It is clear that C&S’s model (101) cannot be reconciled with the unified model. The most obvious evidence for this is the fact that the inflected clitic pronouns in French and Italian are analyzed as DP in the unified model, but IP (the smallest projection) in C&S’s model. The problem is that the models are defined in terms of entirely different sets of features. The unified model is based on the semantic type of the projection: Type *e* entities denoting kinds are NP, type $\langle e, t \rangle$ properties are PredP, and type *e* (or generalized quantifier) argumental entities denoting individuals are DP.²⁶ (101), on the other hand, is based on a whole slew of other semantic, syntactic, morphological and prosodic properties

²⁶The model of D&W, which is based on binding-theoretic status rather than semantic type, is also incompatible with C&S’s model.

pertaining to arguments. However, the implicational relationships between the sets of features associated with each type of pronoun in C&S’s analysis cannot be ignored, so there must be some way to combine the two models to form a more complete theory of pronoun types.

On top of this incompatibility with the unified model, the universality of C&S’s model is brought into question by data from Japanese. Japanese pronouns appear to have properties of both strong and weak pronouns. Like strong pronouns, Japanese pronouns may be coordinated:²⁷

- (102) Watasi-to tomodachi-ga piza-o tabeta.
 I-and friend-NOM pizza-ACC ate
 “Me and a friend ate pizza.”

However, if a pronoun appears in an argument position, it (like non-pronominal nouns) must be accompanied by a case particle:²⁸

- (103) a. Watasi-ga piza-o tabeta.
 I-NOM pizza-ACC ate
 “I ate pizza.”
 b. *Watasi piza-o tabeta.
 I pizza-ACC ate
 c. *Watasi-ga piza tabeta.
 I-NOM pizza ate

This suggests that the pronoun realizes some element smaller than the maximal nominal projection. This is supported by Neeleman and Szendrői (2007), who analyze the pronoun itself as realizing NP and the case particle as a K head, so that *watasi-ga* has the structure $[_{KP} [_{NP} \textit{watasi}] \textit{ga}]$. If we equate Neeleman and Szendrői’s K with C&S’s C, this means that the Japanese pronoun is at the largest a weak pronoun. However, this means that it should not be able to be coordinated; (102) should be impossible. A possible argument could be made that the coordinator *to* is itself some kind of case particle. But Vermeulen (2005) provides evidence against this hypothesis; in certain situations, *to* may be suffixed to every element being coordinated, and in these contexts, the last *to* must be followed by an appropriate case particle (Vermeulen 2005, p. 2):

- (104) John-ga [Mary-to Bill-to]-o mita.
 J.-NOM M.-and B.-and-ACC saw
 “John saw Mary and Bill.”

Assuming each nominal projection has at most one K (or C) node, (104) shows that *to* must be some other category. Thus, Japanese pronouns show qualities of

²⁷Thanks to Yukiko Yoshizumi and Mika Nagai for generously providing Japanese data and judgements for this section.

²⁸Arguments may in fact occur without a case particle in casual speech; in this register, all sentences in (103) are grammatical (although it would be unusual to omit case on one but not both arguments). However, no such register restrictions exist for coordination sentences like (102), so the point remains.

both strong and weak pronouns, which should be impossible in C&S's system. This apparent categorical ambiguity is reminiscent of that already discussed for Japanese pronouns in Section 4, which was resolved with last-resort type-shifting rules, following Chierchia (1998). However, those rules only worked for shifting semantic types, which do not seem to be involved in this case.

The Japanese data are only a problem given C&S's analysis of the distributional freedom of strong pronouns as being due to their having an inherent Case feature. However, it is conceivable that the distributional properties have a prosodic explanation. This sort of analysis would be based on the following assumptions: The reason that only strong pronouns resist restructuring is that only they have the status of a certain size of prosodic constituent, most likely that of phonological phrase; only constituents of this kind can receive stress; an item can only be stressed if there is some reason for it to do so; and there are certain syntactic positions which can be occupied only by items which can receive stress. So, not just any pronoun may qualify as a phonological phrase – it needs some reason to justify receiving stress, such as contrast or ostention. Pronouns with no range-restriction in the C&S sense (Kratzer's (2009) minimal pronouns) are too unspecified to receive stress, and so they may not qualify as a phonological phrase, and must be deficient pronouns.

If this prosodic analysis is on the right track, it must be shown that the positions available only to strong pronouns must be occupied by a prosodic phrase-sized constituent (in the sense described above: a constituent which can, or must, bear stress). For the peripheral positions, such as clefting and dislocation, this is an intuitive claim, since items in these positions are typically emphasized or focused anyway, which tends to involve stress. It is also reasonable to expect that conjoined items receive stress. Consider coordination at the sub-word level: in a word like *underinTERpreted*, the primary stress is on the fourth (capitalized) syllable. However, in a contrastive conjoined expression such as *both UNder- and OVer-interpreted*, the primary stress is on the first syllable. The presence of conjunction has forced stress to move to the prefix, suggesting that it forces phonological phrase-level status on the conjuncts. Finally, there are the base-generated θ -positions to consider, available only for strong pronouns. It may be that base-generated positions in general must be occupied by phonological phrase-sized constituents, and that all smaller constituents must move to a position where they can become part of another phonological phrase. This is similar to C&S's analysis of the movement of clitic pronouns, only at a higher prosodic level: Just as a clitic must move to a head position to receive word-level stress, a weak pronoun must undergo XP-movement to a specifier position to receive phrase-level stress. If this analysis is on the right track, then all the properties dealt with by C&S are in fact prosodic. Given the correlation between prosodic weight, morphological size and syntactic structure, it is not surprising that the wide range of properties they discuss should be correlated in such a way.

In its current form as a sequence of functional projections, the hierarchy proposed by C&S is too inflexible to be united with the unified model. One way to represent the same set of relations in a more flexible way is to treat

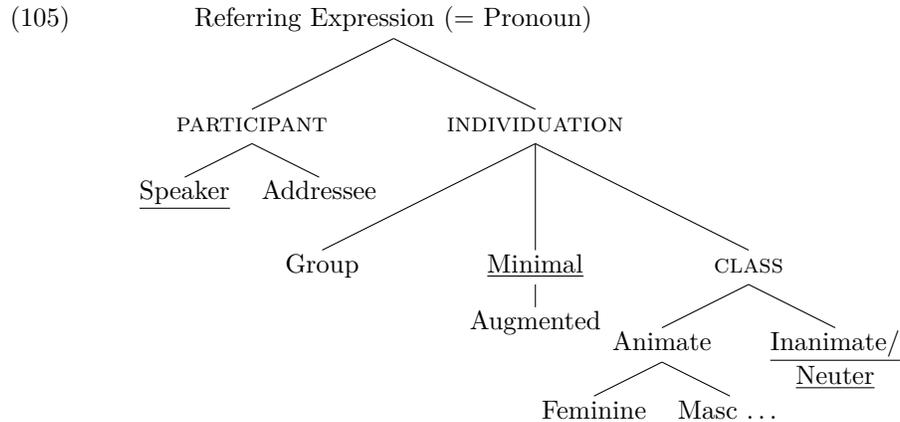
it as an implicational feature hierarchy, rather than a syntactic structure – in other words, as a feature geometry. In this way, the properties associated with strong pronouns could be construed as dependent on those associated with weak pronouns, and the same for clitic pronouns, without positing additional syntactic structure to house these properties. If all the properties in question are in fact prosodic, as suggested above, an explanation outside of syntactic structure is desirable anyway. Since all the pronouns C&S consider are analyzed as pro-DP in the unified model, the entire feature geometry could be located in the DP layer (or, for D-less languages like those discussed in Section 4.5, located in NP). Before fleshing out this proposal, I will review some pronominal feature geometries already in the literature to show that there is a precedent for such an idea.

5.2 Previous Geometries

In the syntax, a collection of features on a single terminal node is generally construed as a largely unstructured bundle. However, many studies assume (often implicitly) a minimal amount of organization, separating semantic, syntactic, morphological, and phonological features. In addition, for some time there has been a tradition of organizing segmental phonological features into a feature geometry. This raises the question of how organized the rest of the feature bundle should be. Geometries for syntactic pronominal features have been proposed by such authors as Harley and Ritter (2002) and Cowper and Hall (2009); I will discuss these proposals here, showing that while some version of a syntactic feature geometry is compatible with the unified model, the prosodic features underpinning C&S’s deficiencies have no place in this kind of geometry.

5.2.1 Harley and Ritter (2002)

The feature geometry proposed by Harley and Ritter (2002) is a response to the standard view that the set of features associated with a syntactic entity is unstructured. Such a view generates combinations of features which are unattested in natural language. For example, if the set of pronominal features contains the six features [1], [2], [sg], [pl], [masc], and [fem], one would expect 2^6 different possible feature combinations, when in fact the number is considerably smaller. Of course, few would deny that some of these features are mutually exclusive (such as [sg] and [pl], or [masc] and [fem]), but even taking that into account, the problem persists. Instead, Harley and Ritter propose that the implicational relationships between different types of features are better captured with the feature geometry in (105) (Harley and Ritter 2002, p. 486):



The features in this geometry are divided into three groups, dominated by the nodes in small caps. Underlining signals that the node is a default, unspecified value for its mother node. All of these features are privative, or monovalent.

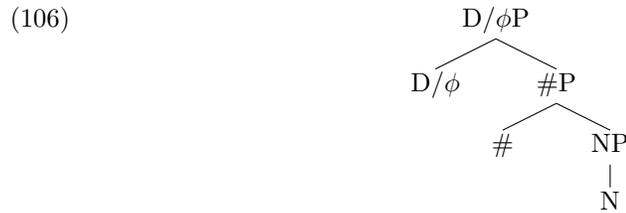
Note that all the features included in this geometry are what are commonly known as ϕ -features, relating to person, number, and gender. Person is given by the PARTICIPANT node: Speaker is 1st person, Addressee is 2nd person. “3rd person” is construed as a lack of person PARTICIPANT features. 1st person inclusive pronouns have both the Speaker and Addressee nodes. Number is determined by the dependents of the INDIVIDUATION node, with singular, plural, dual, etc. denoted by various combinations of Group, Minimal, and Augmented. Animacy and gender are in the dependents of CLASS. Different pronouns contain different subtrees of this geometry, in a way that generates the attested combinations of ϕ -features; markedness (defined in terms of the number of nodes in a tree, such that the more nodes present, the more marked a feature combination is) determines the relative frequencies of different pronominal paradigms.

This geometry deals with different kinds of features than what I will be including in my geometry (I am mainly concerned with the features discussed by C&S, which turn out to be more prosodic than syntactic). However, it is an illustration of the strength of this form of representation as opposed to syntactic structure, in that it allows one to talk about implicational relationships between different kinds of features independently of one another; for example, Augmented is dependent on Minimal, and gender features are dependent on animacy, yet at this level the two sets are independent of one another. Since the prosodic features required to account for C&S’s deficiencies do not seem to interact directly with these ϕ -features, there is no reason why a prosodic feature hierarchy could exist side-by-side with this one in the unified model.

5.2.2 Cowper and Hall (2009)

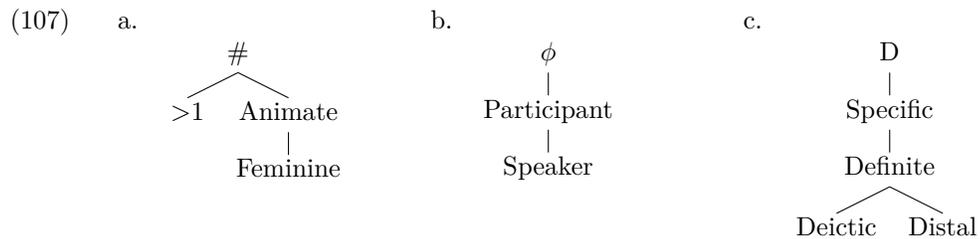
Cowper and Hall (2009) is an undertaking somewhat similar to the present project, in which they adapt the theory of Déchaine and Wiltschko (2002) to better account for the data. Before proposing their feature geometry, they make

some changes to the D&W model to account for systematic ambiguities which are reminiscent to those in Section 3, but different in some key ways. NP is unchanged from the D&W model (unlike the unified model, where I follow Zamparelli (2000) in treating NPs as type *e* kinds rather than predicates). As in the unified model, Cowper and Hall analyze the predicative and argumental instances of ϕ P as separate projections. The predicative variety is relabeled #P, which is more or less equivalent to my PredP, while ϕ P is a label for type *e* argumental pronouns which may be bound. DP is still used to label unbindable, referential argumental pronouns. Thus, they posit four heads instead of three: D, ϕ , #, and N. However, only a maximum of three layers are present in any single pronoun, since DP and ϕ P disjunctively dominate #P. Thus, the model is as in (106), where the top layer may be headed by D or ϕ , but not both:



Like D&W, Cowper and Hall analyze English 3rd person pronouns as categorically distinct from 1st and 2nd person pronouns; they label 3rd person pronouns as ϕ P and 1st and 2nd person pronouns as DP.

Cowper and Hall's feature geometry is different from that in Harley and Ritter (2002) in that it is distributed among different syntactic heads, rather than all stemming from the same root. It also differs in that only marked (non-default) values are present; thus, a tree with the Animate node and lacking Feminine will be considered to have the default value, Masculine. Thus, the feature geometries are as in (107) (Cowper and Hall 2009, p. 98):



As in Harley and Ritter's (2002) model, plurality (here indicated by the feature >1) and gender/animacy are sister nodes, while person (Participant) features are separate. However, unlike Harley and Ritter, Cowper and Hall associate these subtrees with separate syntactic heads. They also include the referential and demonstrative features of D, to achieve the proper featural specification for various articles, demonstratives, and indexical pronouns.

While this analysis is similar to the unified model in a few ways, it runs into some of the same problems as D&W’s original analysis. First, as mentioned above, Cowper and Hall follow D&W in maintaining a categorical distinction between English 1st and 2nd person pronouns on the one hand and 3rd person on the other, while I have shown in Section 3 that such a distinction is unnecessary; their different properties come from other, non-categorical features of the pronouns. In addition, their analysis involves a treatment of English pronouns with overt complements, such as *us linguists*, which runs into empirical issues, as I discussed in Section 3.1. For these reasons, I will not adopt Cowper and Hall’s analysis; still, it is a promising example of how a feature geometry might be integrated into a syntactic structure of the sort which I have been developing.

5.3 My Proposal

Recall that I am attempting to explain those pronominal features described in C&S without resorting to syntactic categorial differences. The reason I am avoiding a categorial analysis is because the analysis of this type proposed by C&S is incompatible with the unified model; all three of C&S’s pronoun types are best described as pro-DP.²⁹ Instead I am resorting to a pronominal feature geometry reminiscent of those described in Section 5.2. C&S’s system differentiates three types of pronouns, but I will show that the system is simplified if clitic pronouns are treated as weak pronouns with the head extracted via X° -movement, motivated by the lack of a feature defining prosodic word status. This leaves me with the strong-weak deficiency to account for, which I will also explain in terms of prosodic features of D.

First I will look at the relationship between weak and clitic pronouns. C&S describe no semantic interpretational differences between these two types; as shown in Table 2, the only differences are that clitics, unlike weak pronouns, receive no word-level stress and have the distribution of syntactic heads, rather than full phrasal projections. C&S claim that the difference stems from the fact that Σ hosts prosody features which are absent in IP, and so the head must incorporate into a head which does have the desired features. In my analysis, I claim that there is some prosodic feature, call it ω , such that when this feature is present, the projection receives word-level stress. When this feature is absent, the head must undergo X° -movement to a location where this feature is present, cliticizing onto some host, thus formalizing the relationship between lack of lexical stress and X° status.

The next task is to formalize the relationship between strong and weak pronouns. Looking again at Table 2, it is clear that there is a bigger set of differences between these types, including syntactic, semantic, and prosodic differences. C&S reduced the entire set of syntactic and semantic differences to a single functional case feature in C, which is interpreted semantically as a referential index. The reason for this analysis was that there are certain unlicensed

²⁹In fact, D&W also discuss how the C&S model might be integrated into their own; they describe C&S’s three pronoun types as instances of pro- ϕ P.

positions where only strong pronouns are permitted. The other difference is that weak pronouns may undergo prosodic restructuring and phonological reduction rules. They are vague on the prosodic details, but it appears the difference is at the level of phonological phrase; strong pronouns must be their own phonological phrase, while weak pronouns may merge with a neighbouring phonological phrase (but they do not have to), as seen in (95). It may be that the distributional and prosodic differences have the same source: only items of sufficient prosodic weight can occur in the unlicensed positions characteristic of strong pronouns. That being the case, the only feature distinguishing strong from weak pronouns is a prosodic feature which specifies the constituent as a phonological phrase. I will label this feature PHP, which is obligatorily present in strong pronouns and absent in deficient pronouns.

Thus, I can explain the deficiencies discussed in C&S in terms of two prosodic features, ω and PHP, which have a hierarchical relationship. ω is absent in clitic pronouns, but present in weak (and, by implication, strong) pronouns. PHP is absent in clitic and weak pronouns and obligatorily present in strong pronouns. The hierarchy is stated in (108):

$$(108) \quad \text{PHP} \succ \omega$$

The presence of PHP implies the presence of ω . This relationship can be translated into a feature hierarchy as in (109):

$$(109) \quad \begin{array}{c} \text{D} \\ | \\ \omega \\ | \\ \text{PHP} \end{array}$$

I am locating these features in D because they seem to be features of DP pronouns, but not PredP or NP pronouns (in [+pred] languages at least; [-pred] languages like Japanese seem to have these features in N, as discussed in Section 4.5). A pronoun is strong iff both features are present, weak pronouns lack PHP, and a pronoun lacking both features is a clitic. Thus, the hierarchical relationship between the three types of pronouns can be captured while maintaining the syntactic structure of the unified model. And since these features are prosodic, rather than syntactic or semantic, they are orthogonal to those described in Section 5.2, and there is no problem with having this geometry overlaid on, or adjacent to (i.e. separate branches stemming from the same terminal node) whichever geometry is preferred (the task of determining the exact relationship between these features and those discussed in Section 5.2 is beyond the scope of this paper).

6 Conclusion

In this paper, I discussed three separate proposals to create a pronoun typology in terms of a specific syntactic model of the noun phrase. Each one is based on

a real set of distinctions between different sets of pronouns, but the three are mutually incompatible, in that the syntactic structures they proposed could not be unified into a single model in their current forms. The solution I chose was to keep the syntactic model from one of the proposals, and explain the other two in terms of some mechanism other than the syntax. Of the three models, Zamparelli (2000) based his largely on the assumption that there should be a correlation between syntactic category and semantic type (individual versus property, argument versus predicate, kind versus specific entity); this is the syntactic structure I adopted for the unified model. Déchaine and Wiltschko's (2002) model is largely based on binding properties of the pronouns; I showed that such binding contrasts are better described in terms of Kratzer's (2009) binding theory than syntactic structure. Finally, the properties motivating the model of Cardinaletti and Starke (1994) turned out to involve prosodic weight more so than syntactic features, which I analyzed as a small prosodic feature geometry. In other words, the unified model uses syntactic structure to explain only semantic type differences, while binding and C&S's distributional differences are best described in terms of the presence and absence of ϕ -features and prosodic features, respectively.

One problem that I faced over the course of this analysis was that Japanese pronouns (and, it turned out, Mandarin and Korean pronouns) appeared to be systematically ambiguous between NP and DP, even in the unified model. Such systematic ambiguities are problematic, in that it is unexpected that there would be a set of homophonous items which can have two distinct sets of categorial and selectional features. However, I showed that these pronouns are in fact unambiguously NP, and that NPs are expected to take on DP-like properties in languages with no articles or count nouns, following Chierchia (1998). The fact that all systematic categorial ambiguities have thus been removed from the model suggests that my analysis is on the right track. Further research is needed to determine the cross-linguistic validity of my claims. In particular, it remains to be seen whether the similarities between Japanese, Mandarin, and Korean are in fact due to their parameter settings, rather than an areal relationship; this requires research into the properties of pronouns in some obscure languages, since there are few languages of the appropriate type outside of Eastern Asia. In addition, more research should be done to ensure that article-less, [+pred] languages do in fact pattern with Turkish and not with [-pred] languages like Japanese with respect to the categorial status of their pronouns. Finally, it remains to be seen just how other features, such as Case and ϕ -features, should be integrated into the geometry, and precisely how this interacts with Kratzer's (2009) binding theory. This may involve reformulating Kratzer's binding theory in terms of a ϕ -feature geometry like that of Harley and Ritter (2002), rather than the unstructured feature bundles Kratzer deals with. All in all, however, the success so far in unifying such different pronominal typologies into a unified model with no systematic ambiguities suggests that this is a fruitful line of research.

References

- Bošković, Ž. (2012). On NPs and clauses. Ms., University of Connecticut, Storrs.
- Cardinaletti, A. and M. Starke (1994). The typology of structural deficiency: A case study of the three classes of pronouns. *University of Venice Working Papers in Linguistics* 4(2), 41–109.
- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics* 6, 339–405.
- Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Berlin, Germany: Mouton de Gruyter.
- Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Cowper, E. and D. C. Hall (2009). Argumenthood, pronouns, and nominal feature geometry. In J. Ghomeshi, I. Paul, and M. Wiltschko (Eds.), *Determiners: Universals and Variation*, pp. 97–120. Amsterdam: John Benjamins.
- Déchaine, R.-M. and M. Wiltschko (2002). Decomposing pronouns. *Linguistic Inquiry* 33(3), 409–442.
- Déchaine, R.-M. and M. Wiltschko (2012). When and why can 1st and 2nd pronouns be bound variables? Ms., UBC.
- Dryer, M. (2011a). Coding of nominal plurality. In M. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*, Chapter 33. Munich: Max Planck Digital Library. Available online at <http://wals.info/chapter/33>. Accessed on 2013-08-08.
- Dryer, M. (2011b). Definite articles. In M. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*, Chapter 37. Munich: Max Planck Digital Library. Available online at <http://wals.info/chapter/33>. Accessed on 2013-08-08.
- Dubinsky, S. W. and R. Hamilton (1998). Epithets as antilogophoric pronouns. *Linguistic Inquiry* 29(4), 685–693.
- Grimshaw, J. B. (1991). Extended projection. ms., Brandeis University. Published in Grimshaw (2005).
- Grimshaw, J. B. (2005). *Words and Structure*. Stanford, CA: CSLI Publications.
- Harley, H. and E. Ritter (2002). Person and number in pronouns: A feature-geometric analysis. *Language* 78(3), 482–526.
- Kratzer, A. (2009). Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2), 187–237.

- Melchin, P. and R. Truswell (2013). Noun phrases and non-projecting heads. Paper presented at the Canadian Linguistics Association, 2 June 2013.
- Neeleman, A. and K. Szendrői (2007). Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38(4), 671–714.
- Pollard, C. and I. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Rullmann, H. (2004). First and second person pronouns as bound variables. *Linguistic Inquiry* 35(1), 159–168.
- Starke, M. (2009). Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36.
- van Riemsdijk, H. (1998). Categorial feature magnetism: The endocentricity and distribution of projections. *The Journal of Comparative Germanic Linguistics* 2(1), 1–48.
- Vermeulen, R. (2005). Coordination in Japanese: A case of syntax-phonology mismatch. Ms., University College London.
- Yashima, J. (2013). Third-person pronouns and variable binding in Japanese. Ms., UCLA.
- Zamparelli, R. (2000). *Layers in the Determiner Phrase*. New York: Garland Publishing.