

Cross-Linguistic Variation in Nominal Functional Sequences and Pronoun Types*

Paul B. Melchin

University of Ottawa

`pmelc074@uottawa.ca`

1 Introduction

One of the questions that has been debated, explicitly or implicitly, since the early days of generative grammar addresses the extent to which the syntactic structures underlying linguistic expressions are universal. For the past thirty or so years, especially since the advent of works such as Abney's (1987) DP Hypothesis and Grimshaw's (1991) Extended Projection Hypothesis, these underlying structures have taken the form of fixed hierarchies of functional projections; that is, selections of functional heads dominating the lexical item, embedded in a particular order. I will refer to each sequence of functional heads (i.e., each combination of a selection of heads and an order that those heads appear in) as a functional sequence (hereafter, fseq). However, while fseqs have been used as a theoretical tool by numerous syntacticians over the years, there is great disagreement over their structure, universality,

*I would like to thank my supervisors, Rob Truswell and Éric Mathieu for their support and comments throughout this project; the audiences at BWTL 16 and MOLTH, and members of the University of Ottawa Syntax-Semantics Lab for their insightful comments on earlier versions of this paper; my informants for their native-speaker judgements and data, including Basile Roussel (French), Meng Yang (Mandarin), Yukiko Yoshizumi (Japanese), Chung-hye Han and Kyumin Kim (Korean), Saara Huhmarniemi (Finnish), Lev Blumenfeld and Yuliya Manyakina (Russian), and Ana Werkmann Horvat (Croatian); and finally, my parents, Mike and Rosemary Melchin, for their continual support. Of course, any errors herein are entirely my own.

and the very reason for their existence. Some points of disagreement include:

- What kinds of data can be taken as evidence for the structure of an fseq? Proposals include:
 - Purely morphosyntactic data (cross-linguistic ordering facts, complementary distribution, etc.; see Ritter 1991, 1995; Abels and Neeleman 2012, among many others);
 - Semantic composition (see e.g. Zamparelli 2000; Borer 2005);
 - Binding and anaphora (binding properties of a phrase depend on its category; e.g., Déchaine and Wiltschko 2002, 2012; Cowper and Hall 2009);
 - Fixed order of modifiers (each as a specifier to an abstract functional head; e.g., Cinque 1999; Scott 2002).¹
- Are the fseqs of a given category the same, cross-linguistically? There are two sides to this question:
 - Do all languages employ the same inventory of functional heads? That is, is a functional head present even in a language where it is not overtly realized? (This question applies within a language too: do all lexical items of a given category in a given language project the same fseq?)
 - Is the order, or hierarchy, of functional heads the same in all languages, or can the same items appear in different orders in different languages?
- Are fseqs of different categories substantially similar? That is, should nominal and clausal (and perhaps adjectival) fseqs be considered to be parallel, or are they fundamentally different?

¹This could be construed as a form of morphosyntactic data, given that it involves word order within the phrase, or as semantic composition, given that it involves grouping adjectives into semantic categories; however, since it is unclear to which of the two categories it properly belongs, and since it has generated a literature of its own, I list it as a separate category here.

- Many proposed fseqs (e.g., Abney 1987; Grimshaw 1991; van Riemsdijk 1998) are based on an assumption of parallelism between the nominal and clausal domains.
- However, Melchin and Truswell (2013) note differences between the two domains in terms of semantic selection, suggesting that the clausal domain is more layered or articulated than the nominal domain; phase-theoretical considerations (Chomsky 2000, 2005) concur, positing two phase heads in the clausal domain (v^* and C), and only D, or perhaps none, in the nominal domain.

This paper addresses the first two of these questions by examining the syntactic properties of pronouns, compared to noun phrases² in general, cross-linguistically. Since this study focuses on the nominal domain, the question of parallelism between categories will remain unaddressed here. As will be seen, pronouns in different languages differ in various properties, including morphological structure and the potential for adjunction of various nominal modifiers (e.g., adjectives, genitive noun phrases). Further, these differences are not randomly distributed across languages, but correlate with differences in the presence and arrangement of functional material in the noun phrase – that is, with (potentially) different fseqs. In this paper I examine independently-motivated accounts of the fseqs of different languages and show that the pronominal differences between the languages can be derived from these different structures. Specifically, I claim that each language has a set of fseqs for different kinds of noun phrases, and that pronouns corresponding to a given kind of noun phrase will share with it the same functional structure; for example, since personal pronouns are definite noun phrases, singular personal pronouns will pattern with singular definite noun phrases in the language, plural pronouns with plural definite noun phrases, and so on. Thus, by showing that pronominal differences can be explained by these different fseqs, I provide a new source of evidence for certain proposed fseqs as real syntactic entities, and that languages differ in both the functional heads present in the fseq and the order in

²In this paper I will use the expression ‘noun phrase’ as a pre-theoretical short-hand for the nominal fseq or extended projection; the projection of the category N will be referred to as NP.

which they appear, yet these differences are constrained in certain ways.

The structure of this paper is as follows. Section 2 outlines the theoretical assumptions made in this paper. Here I clarify what is meant by fseq, and show that in some cases at least, the internal syntactic structure of pronouns is overtly evident. I also address some influential proposals for the nominal fseq and the structure of pronouns, and show why I reject them. Section 3 presents a case study that addresses a three-way contrast between Japanese, Mandarin and English in terms of modifiability of pronouns, and shows that, while Japanese and Mandarin noun phrases are superficially similar, there are structural differences that explain this pronominal contrast. Section 4 concludes, and provides some speculation on the theoretical consequences of the analyses in this paper.

2 Theoretical Standpoint

2.1 Introduction

In this section I discuss and justify the theoretical views behind this project. In Section 2.2 I discuss the theoretical mechanisms underpinning the analyses, starting with a discussion of the means by which elements enter the derivation and interact with other heads, then moving to a discussion of fseqs and their properties and motivations behind them. I also discuss and reject a previous analysis in which different semantic properties are strictly correlated with different syntactic structures, on the grounds that this makes the wrong cross-linguistic predictions. Section 2.3 provides a brief summary and discussion of the categories of pronouns in the languages to be discussed in detail in this paper. In Section 2.4 I demonstrate some cases where the structure of pronouns can be seen relatively clearly from their overt morphosyntactic properties, demonstrating that pronouns in different languages can have different syntactic structures while still realizing the same maximal projection.

2.2 Heads and Structures

In this subsection I address the theoretical mechanisms used in the analyses here. First I discuss some mechanisms for the spell-out and movement of syntactic heads, showing that there are multiple mechanisms that are compatible with my analyses; then, I discuss the nature of fseqs and extended projections. Finally, I discuss and reject a structure with a similar level of articulation, but in which the heads have different properties.

2.2.1 Movement and Spell-out

When discussing fseqs and syntactic projections of different kinds of words and phrases, it becomes necessary to consider the relationships between the underlying structures and the overt elements that express them in a sentence. Specifically, what is needed is a way for one word to realize multiple terminal nodes. This is necessary in order to constrain the range of possible structures that can be posited for a given expression, and to formulate a proper theory of how overt sentences are built from abstract underlying structures. The most common view is a theory of head movement that is constrained in specific ways; here I will discuss the specific theory of head movement that I adopt here, but I also briefly describe a mechanism of spanning that is also compatible with my analyses.

Head movement is based on the assumption that each morpheme (overt or null) corresponds to a functional head of some category, and that morphemes are concatenated through the successive movement of free morphemes into c-commanding heads. Thus, the order of morphemes is a reflection of the sequence of heads, in that affixes nearer to the base-generated position of the moved head in the fseq will be closer to the root in the resulting word. This movement is crucially subject to the Head Movement Constraint of Travis (1984), which states that movement must be to the *nearest* c-commanding head; in other words, heads cannot be skipped over.

Specifically, I am assuming the particular characterization of head movement assumed in Borer (2005). In this model, every head in the fseq that is present in a given structure

must be filled in some way. Heads can be filled by overt free morphemes or (overt or covert) bound morphemes.³ Bound morphemes require the movement of some free morpheme to their head, in order to provide them with phonological support. Free morphemes in effect block this movement (due to the Head Movement Constraint), but they themselves may move to provide support for higher bound morphemes.

In models of syntax employing head movement, there is a contrast between head movement and phrasal movement: the former is movement from heads to other heads only, while the latter is movement of phrasal constituents from a specifier or complement position to a higher specifier position. However, this contrast runs into problems in light of the theory of Bare Phrase Structure, outlined in Chomsky (1995) and assumed in much subsequent work in the Minimalist Program. In this theory, levels of projection are eradicated, and so there is no difference between a head and a phrasal projection, and thus no structurally-determined head, specifier, adjunct and complement positions in the syntactic structure. In such a framework, positing a contrast between head movement and phrasal movement is problematic (see e.g. Barrie and Mathieu 2012, Haugen and Siddiqi 2013 for discussion of these and other issues with head movement; cf. Matushansky 2006, Roberts 2010 for reworkings of head movement in a contemporary Minimalist framework).⁴ An alternative to head movement is the notion of spanning, introduced by Svenonius (2012), in which the syntax deals only with syntactic features, and phonological material is inserted at PF. In a spanning analysis, a lexeme may realize a span of terminal nodes that are linearly adjacent to one another. Distributed Morphology (Halle and Marantz 1993) also provides mechanisms for lexical items to realize multiple heads, such as fusion; however, see Haugen and Siddiqi (2013) again for some problems with traditional Distributed Morphology approaches.

The analyses in this paper will be presented in the head-movement framework outlined above, but they are also compatible with a spanning analysis, and possibly other analyses.

³Borer (2005) refers to free functional morphemes as f-morphs, and bound morphemes as head features.

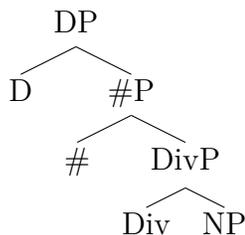
⁴There is more to head movement as well, such as the idea that it achieves the proper order of affixes cross-linguistically, and that the Head Movement Constraint allows certain exceptions; however, these considerations do not bear on the project at hand, so I will not discuss them here.

Note that, due to the Head Movement Constraint, and the fact that a head may not move into a head occupied by a free morpheme, a given word ends up picking up the features from a chain of heads that are in a relationship of immediate c-command; this amounts to saying that it realizes (or, at least is endowed with the features from) a span of heads. In other words, the analyses in this paper can be easily reformulated in a framework utilizing spanning, rather than head movement, and the choice between the two mechanisms has little consequence for this project. Any analysis that allows a single lexical item to realize multiple terminal nodes in an immediate c-command relationship will do. I will stick with a head movement analysis, however, since this is the predominant mechanism used in the literature on this topic.

2.2.2 The Functional Sequence

In this paper, I assume a fairly flexible model of the nominal fseq, varying within and between languages, within certain constraints. With respect to the composition of nominal denotations, in terms of the organization of aspects of interpretation such as the basic concepts involved, countability, cardinality, and discourse status (definiteness, specificity, and so on), I follow the general approach of Borer (2005), which is based on the structure in (1). However, while I accept Borer’s model in terms of the overall layout of conceptual meaning, division, cardinality, and referentiality in the fseq, I allow more variation in terms which heads are present and their order. To start with, I will introduce Borer’s model:

(1)



In this model, NP consists of an uncategorized root merged with a nominalizer; the result is an undivided, kind- or mass-like denotation. Div divides the NP’s denotation into units

that can be counted, and is absent in mass nouns; this category includes numeral classifiers and plural marking (at least, the canonical plural marking found in European languages). The quantity or cardinality head $\#$ counts the resulting denotation and provides a quantity; this category includes numerals and many quantifiers. Finally, D provides information on specificity and referentiality, and is realized by definite articles, demonstratives and strong quantifiers.

This model was chosen for a few reasons. For one, it is based on principles of co-occurrence restrictions and semantic composition. For example, numeral classifiers and plural markers rarely co-occur in a language or a given noun phrase (and if they do co-occur, the plural tends to have some modificational meaning, such as abundance; see Mathieu 2012b); furthermore, both are necessary in order to count items in many languages. Thus, they are both considered to be of the category Div, which makes nouns countable. Another reason is that the level of articulation in this model allows for precise generalizations regarding correlations between semantic denotation and the amount of structure that is present to a degree that previous accounts are lacking.

However, I differ from Borer in some assumptions about the fseq in (1), and have some theoretical assumptions of my own regarding the relationships between pronouns and non-pronominal noun phrases. Borer assumes that every argumental noun phrase projects a DP; however, I maintain that not all languages have the category D, and in those languages that have D, not all argumental noun phrases project this structure (see discussion of Japanese and Mandarin in Section 3.2). Specifically, I claim that every language has some maximal fseq similar to that in (1), although perhaps lacking D, and perhaps with some other heads included (such as Japanese's CaseP, introduced in Section 3.2.1), as long as the function of such heads does not interfere with the semantic composition outlined above. For example, CaseP may be added to the projection because it contributes information about grammatical function which is orthogonal to the semantics of definiteness and number that are composed by Borer's fseq, but additional nodes of quantification, for instance, would interact seman-

tically with the contributions of Div and # and so are excluded. I must therefore assume, following Ritter (1995), Mathieu (2012a), Wiltschko (2008), and others, that languages may differ in terms of which features are realized on which terminal nodes. Furthermore, I assume that noun phrases of a given semantic type (i.e., a given definiteness, number, and so on) will have the same functional structure in language X, though this may differ from the structure of an equivalent noun phrase in language Y. Specifically, I assume, following Carlson (1977), Ritter (1995), Lyons (1999), and others, that definite and specific indefinite noun phrases have the full nominal projection, but non-specific noun phrases may lack the DP projection (or its equivalent). As for pronouns, I claim that they realize the same structure as non-pronominal noun phrases in the same language. Thus, singular personal pronouns, which receive a definite interpretation, in language X will have the same fseq as singular definite noun phrases in language X. Finally, I claim that there is a fundamental difference between N and D on the one hand, and Div, #, and other intermediate functional heads on the other hand; the exact nature of this difference is hard to pin down, although suggestions are made in Section 4. Crucially, however, it is shown below that pronouns may realize either D(P) or N(P), but not any other categories, so this difference must stem from somewhere.

This builds the view of pronominal structure proposed in Postal (1966), later developed in a DP framework by Abney (1987). Postal proposed that pronouns in English are intransitive definite articles (i.e., definite articles that do not take a nominal complement); in a DP framework, this translates to elements of category D that do not take a complement. This proposal works for English in my analysis, but in languages that do not have D, or for constructions that lack the DP projection, this is clearly not a possibility. Thus, the structure of pronouns in a language must be dependent on the nominal syntax of the non-pronominal noun phrases in the language, as proposed by Postal, but I differ in positing a more complex syntactic structure that varies from language to language.

To demonstrate how this would work, while at the same time illustrating the shortcomings of less-articulated structures, let us consider a model in which the extended projection of N

contains only two projections, NP and DP. Such a model is assumed in such works as Bošković (2008, 2012) and Chierchia (1998), among many others. Here I will discuss Chierchia’s model. Chierchia divides languages into three classes, depending on the denotation of bare NPs in the language. According to Chierchia, in some languages all NPs are arguments, in others all NPs are predicates, and in a third group of languages, NPs may be either argumental or predicative. The generalization is captured using two parameters, $[\pm\text{arg}]$ and $[\pm\text{pred}]$. There are thus three kinds of languages in Chierchia’s system:⁵

- NP[+arg, –pred]: In these languages, all bare nouns may be construed as arguments as a result of the application of the \cap (down) operator in the lexicon, and thus all nouns receive a mass or kind denotation.⁶ These languages allow bare nominal arguments, with no definite or indefinite articles, and have no plural marking, but instead require numeral classifiers when counting items. Examples: Chinese, Japanese.
- NP[–arg, +pred]: In these languages, bare nouns may never be construed as arguments; instead, they are predicates that must be saturated by some D in order to appear as an argument. These languages have both mass and count nouns, and make use of plural marking to count nouns. Examples: Romance languages
- NP[+arg, +pred]: In these languages, the \cap operator is available in the lexicon for mass nouns and plurals, but count nouns enter the syntax as predicates and must receive an article.⁷ Thus, in these languages, plurals, mass, and kind-denoting nouns will be bare, and singular count nouns will require some D. This group includes languages with null definite and indefinite articles, such as Russian (these are differentiated from NP[–arg, +pred] languages by the presence of obligatory plural marking and the lack of numeral classifiers). Examples: Germanic languages (including English), Slavic languages

⁵The fourth logically possible feature setting, NP[–arg, –pred], is impossible since in these languages, NPs would have no denotation.

⁶See Partee 1987, p. 121 for more on this and other type-shifting operations; see also Chierchia 1998, Section 2 for more on Chierchia’s particular semantic ontology.

⁷In Chierchia’s ontology, mass nouns and plurals have the same kind of denotation, except that plural denotations lack the set of atomic entities.

In this classificatory system, and under the assumption that pronouns in a given language match that language's non-pronominal noun phrases in category, one would expect to find at least two different kinds of pronouns, pro-NP and pro-DP. Specifically, the NP[+arg, -pred] languages should have pro-NP, since these languages lack D and so all nouns are maximally NP. NP[-arg, +pred] languages should all have pro-DP, since the DP projection is required for arguments in these languages. In NP[+arg, +pred] languages, the prediction is not clear since both NP and DP may serve as arguments in these languages, but one might expect that pronouns of both categories will be possible, but that they will be pronouns of different semantic types; thus, for example, pronominal forms of mass nouns (and perhaps plural count nouns) may be pro-NP, and those of (singular) count nouns will be pro-DP.

If type-shifting operations can change elements of category NP into definite descriptions, then semantically pro-NP and pro-DP should be equivalent; however, the difference in category leads to the prediction that they will differ in certain syntactic properties. One such difference is that pro-NP, unlike pro-DP, should be amenable to attributive modification by elements such as adjectives and genitive noun phrases, as these tend to adjoin to NP, or at least to elements below DP in the extended projection. Modifiability of pronouns has been proposed as a diagnostic for pronoun category by Fukui (1988), Noguchi (1997) and Bošković (2008), such that pronouns which cannot be modified are analyzed as pro-DP, and those that can be modified are pro-NP, or at least some pro-XP, where XP is a projection dominated by DP.

Using this modifiability diagnostic, a quick glance at the data suggests that this analysis is on the right track. For example, pronouns may not be modified in Romance languages, confirming that these NP[-arg, +pred] languages have pro-DP; This is illustrated with French data below:

- (2) a. *Il d'aujourd'hui est plus intelligent que lui d'hier.
 he of.today is more intelligent than him of.yesterday
 'Today's he is more intelligent than yesterday's him.

- b. *Heureux ils ne vont pas à l'école.
 happy they NEG go NEG to the.school
 'Happy them aren't going to school.'

On the other hand, pronouns may be modified in the NP[+arg, -pred] languages Japanese, Mandarin, and Korean, as seen in (3), suggesting that they are pro-NP languages:

- (3) a. Zuo tian de ta bi jin tian de ta gen cong ming.
 yesterday POSS him COMP.PRT today POSS he more smart
 'Yesterday's he is smarter than today's him.' *Mandarin*
- b. Kyou-no kare-no hou-ga kinou-no kare yori atamagaii desu.
 today-GEN he-GEN way-NOM yesterday-GEN he than smart is
 'Today's he is smarter than yesterday's him.' *Japanese*
- c. Onul-uy ku-ka ecey-uy ku-pota te hyenmyengha-ta.
 today-GEN he-NOM yesterday-GEN he-than more smart-DECL
 'Today's he is smarter than yesterday's him.' *Korean*

However, the variation expected in the NP[+arg, +pred] group of languages does not occur. While one would expect some pronouns to be modifiable and some not, they in fact all disallow modification, as in (4) (Russian data from Richardson 2001, p. 1, Serbo-Croatian data adapted from Runić 2011, p. 39):⁸

- (4) a. *Hän on viisas joka päivä mutta eilisen hän oli viisaampi kuin
 s/he is smart every day but yesterday.GEN s/he was smarter than
 tämänpäivän hän.
 today.GEN s/he
 'S/he is smart every day but yesterday's s/he was smarter than today's
 s/he.' *Finnish*
- b. *Milicia privela p'janogo ego domoj.
 police brought drunk.ACC him.ACC home
 'Police brought drunk him home.' *Russian*
- c. ?*Jesi li ga vidio juče? Jesam, ali je jučerašnji on baš nekako
 are Q him.Cl.Acc seen yesterday Am but is yesterday's he really somehow

⁸The Serbo-Croatian example was judged by my informant as only marginably acceptable, giving an archaic and strange reading only. Similar results were found for Russian, and indeed for the English examples: they are interpretable, but not at all natural.

bio čudan.
 been strange
 ‘Did you see him yesterday? I did, but yesterday’s he was really somehow
 strange.’ *Serbo-Croatian*

As these pronouns all denote count nouns, it is expected under Chierchia’s framework that they would all realize DP; however, even pronouns for count nouns are unmodifiable, as the following English example illustrates:

(5) *I spread sand on my driveway, because gritty it keeps it from getting too slippery.

Thus, while Chierchia’s model explains certain of the data, it does not provide a satisfactory explanation for the NP[+arg, +pred] languages. Furthermore, more recent studies have called Chierchia’s model into question; many languages do not fit into any of the three categories, for example showing co-occurrence of classifiers and plurals, or evidence of D in classifier languages; see Butler (2011), Li (1999), Nemoto (2005), and others. In Section 3.2.2, it will be shown that Mandarin has elements of category D. In order to explain the complexity of the data, a more highly-articulated model is required, and the structure of Borer (2005) has the required abundance of functional nodes to do so.

2.2.3 Which Functional Heads?

Having shown that an fseq articulated with multiple functional heads is necessary to explain the variation between different types of noun phrases, the task remains to justify that the functional heads in (1) are themselves the correct set of heads to be working with. In the system described above, in which each language uses some variation on (1), the structure is used to explain syntactic and morphological facts, but the translation to the semantics is flexible. For example, D is used as an argumentizer in languages that have D, but in languages that do not, other mechanisms are used, perhaps semantic operators as in Chierchia (1998). Déchaine and Wiltschko (2002) propose an influential fseq to account for pronominal and

non-pronominal noun phrases while strictly correlating syntax, morphology, and semantics; however, here I show that this proposal has its shortcomings and should be abandoned.

Déchaine and Wiltschko (2002) (henceforth D&W) form a theory of the relationship between pronoun types and nominal phrase structure based largely on evidence from bound variable anaphora. 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 pronoun may realize any of the three maximal projections, DP, ϕ P, or NP; the properties of the pronoun will vary depending on the type of projection it realizes.

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,⁹ and their binding

⁹The 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.

properties are supposed to derive from their inherent semantics in a predictable way.

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, namely uninflected *le* and inflected *le*, *la*, and *les*, 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):

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

In addition, as expected with a pro-NP functioning semantically as a constant, *en* cannot function as a bound variable (9), and it cannot co-refer with an antecedent (10) (Déchaine and Wiltschko 2002, p. 428):

- (9) * $[Chacun]_i$ pense que Jean $[en]_i$ a vu
each.one thinks that J. *en* has seen
- (10) * $[Marie]_i$ 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 (11) and as predicates (12) (although as a predicate, only uninflected *le*

is possible) (Déchaine and Wiltschko 2002, p. 428):

- (11) Jeanne *la* voit
J. her sees
'Jeanne sees her.'
- (12) 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 (11), it shows agreement features, while as a predicate it does not (12). 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):

- (13) [*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.

As can be seen from the preceding discussion, D&W base their analysis largely on the binding properties of different kind of pronouns, such that a pronoun that can be bound is pro- ϕ P, and one that cannot is either pro-DP or pro-NP (depending on other properties of the pronoun). Thus, they claim that first person pronouns are pro-DP in English, but pro- ϕ P in French, on the basis of the following alleged contrast (pp. 423, 431):

- (14) I_i know that John saw *me*_i, and Mary does too.
= a. I [λx . x knows that John saw me] & Mary [λx . x knows that John saw me]
≠ b. I [λx . x knows that John saw x] & Mary [λx . x knows that John saw x]
- (15) [*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 [λx . x thinks that the police saw me] & Mary [λx . x thinks that the police saw me]
- b. I [λx . x thinks that the police saw x] & Mary [λx . x thinks that the police saw x]

These examples are intended to show that first person pronouns can function as a bound pronoun, and co-vary with different antecedents, in French but not English. In English, the pronoun *me* can (allegedly) refer only to the speaker in (14), and not to Mary in the elided second clause, while in French, both readings are available. However, some English speakers (myself included) do in fact accept the second, bound variable reading in (14), and D&W themselves note that not all French speakers accept the bound variable reading in (15). Thus, it seems that the judgements for bound variable anaphora in first person pronouns are the same in English and French, contrary to D&W's claims, such that in both languages, a bound variable reading is available for some speakers but not others.

A similar situation holds for Japanese pronouns. D&W claim that the third-person pronouns *kare* 'he' and *kanozzyo* 'she' are pro-NP, and therefore cannot act as a bound variable. This is on the basis of examples like those in (16) (p. 417):

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

In section 3, I show that Japanese pronouns are, in fact, pro-NP (or, alternatively, they realize N; the choice between these makes no empirical difference that I can find). However, contrary to D&W's claims, they can in fact act as bound variables, showing that the cor-

relation between category and bound variable status should be dropped. Yashima (2013) demonstrates that, in certain contexts, Japanese third-person pronouns can act as bound variables, as shown in the following examples:

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

Specifically, the contexts in which these pronouns are antilogophoric contexts, or contexts in which the antecedent is not the one from whose perspective the pronoun is evaluated. This is the same set of contexts in which epithets can be evaluated, leading Yashima to claim that these pronouns are in fact epithets; see Dubinsky and Hamilton (1998) for more on epithets and antilogophoricity.

Thus, it seems that positing pronominal categories in an fseq which is itself based largely on anaphoric binding properties of the pronouns is impossible, since both the pro-DP pronouns of English and French, and the pro-NP pronouns of Japanese, can be bound (albeit only under certain conditions for some of the pronouns). This is compatible with a system in which the binding properties of a nominal constituent depend not on its maximal category, but on the featural make-up of the constituent; see Kratzer (2009) for such a theory of binding. Therefore, I reject both the structure in (6) and D&W’s claims that first and second

person pronouns differ from third person pronouns in category in English. See Melchin (2013) for more on D&W, bound variable anaphora, and relationships between Japanese pronouns and epithets. More generally, the above discussion illustrates the infeasibility of a system in which the fseq is strictly correlated with all of syntax, morphology, and semantics; at least one of these correlations must be loosened in order to achieve accurate cross-linguistic predictions, since different languages seem to be able to achieve similar semantic denotations with different syntactic constructions.

2.3 Summary of Pronoun Types

Having established that a Borer-type model is needed to account for the variation between the pronouns and noun phrases of different languages, but before looking closer at particular languages, I will here briefly summarize the pronoun categories of each of the languages (or language families) to be discussed in detail in this paper. This is summarized in (18):

(18)

Language	Pronoun category
English	DP
Hebrew	DP
Algonquian	DP
Mayan	PP
Mandarin	DP (coercible to NP)
Japanese	NP
Korean	NP

Note that the model in (1) predicts a wider distribution of pronoun types than what is observed; that is, we only see pro-NP and pro-DP, but no pro-DivP or pro-#P. The reason for this is not entirely clear, although speculations are given in Section 4.

2.4 Structure of Pronouns

Here I present previous accounts from the literature which show the internal structures of pronouns in two languages: Ritter (1995) for Hebrew, and Déchaine (1999) for Algonquian languages; I also show data from Mayan languages and show that it suggests that pronouns in these languages are pro-PP. In these languages, the syntax and morphology make the internal structures of the pronouns relatively clear, and demonstrate that two languages may have pronouns of the same maximal category (DP), yet the distribution of features across the heads differs between languages, and within Hebrew it depends on the person of the pronoun as well. Note that both of these analyses are based on a somewhat simpler fseq than that in (1), having only two functional projections above N, Num (which can be equated with Borer’s Div head, since it encodes plurality) and D.¹⁰ However, since quantification will not factor into the analyses in this section, the contributions of the # head are minimal and it can be safely excluded from the discussion of these languages.

2.4.1 Hebrew

Ritter (1995) provides a structural analysis of Hebrew personal pronouns. The Hebrew pronominal paradigm is given in (19) (Ritter 1995, p. 419):

(19)

	singular		plural	
	masc.	fem.	masc.	fem.
1st	ani		anaxnu	
2nd	ata	at	atem	aten
3rd	hu	hi	hem	hen

¹⁰In Déchaine (1999), the intermediate projection is labelled PersP; however, since this head contains features for number and person, it is analogous to Ritter’s NumP projection, and for the sake of consistency I will label it as such in this paper.

Ritter determines the maximum level of projection of each pronoun based on co-occurrence restrictions. Specifically, the definite article *ha*, of category D, may not co-occur with first or second person pronouns, but it may co-occur with third person pronouns, with the result being the distal demonstratives, as shown in (20) (Ritter 1995, p. 420):

- (20) a. *ha-ani / *ha-anaxnu
 *the-I / *the-we
- b. *ha-ata / *ha-at / *ha-atem / *ha-aten
 *the-you(m.sg.) / *the-you(f.sg.) / *the-you(m.pl.) / *the-you(f.pl.)
- c. ha-hu / ha-hi / ha-hem / ha-hen
 the-he / the-she / the-they(m.) / the-they(f.)
 ‘that(m.)’/‘that(f.)’/‘those(m.)’/‘those(f.)’

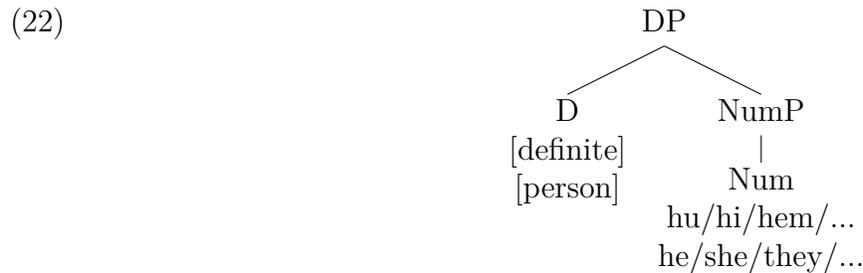
This implies that first and second person pronouns are maximally DP, but the third person pronouns are at most NumP.

In Ritter’s analysis, the nominal features of person, number and gender are not fixed to a specific head in the projection, but are distributed throughout the heads based on availability, while definiteness is always situated in D. Pronouns are analyzed as functional material without an accompanying, lexical N head (following Postal 1966 and others). Thus, she claims that the first and second person pronouns are simply a D head, endowed with the features of definiteness, person, number, and gender, as shown in (21) (Ritter 1995, p. 420):

- (21) DP
 |
 D
 [definite]
 [person]
 [number]
 [gender]

On the other hand, the third person pronouns are assumed to be elements of category Num, since they may co-occur with the definite article to give demonstratives. However, since personal pronouns in Hebrew are definite, they require an element of category D, which

Ritter takes to be null. Furthermore, since demonstratives are not specified for person, the Num head associated with these pronouns is considered to be specified only for gender and number, such that when this head combines with the definite article to result in a demonstrative, there is no person feature; therefore, the person feature associated with the third person pronouns is claimed to be located in the null D as well. Therefore, the third-person pronouns have the structure in (22) (Ritter 1995, p. 421):



Note that, while there is no overt evidence for this null D, it is included for the sake of consistency: Like first- and second-person pronouns, third-person pronouns are definite, and endowed with person features, and these are properties that are associated with full DPs in Hebrew. It would therefore be unlikely that these properties are available for NumPs only in the case of third-person pronouns.

Thus, Ritter shows that, while all personal pronouns in Hebrew maximally project DP when they occur as arguments in a sentence, they differ in their internal structure. This is apparent from the fact that third person pronouns, but not those of first or second person, may combine with the definite article to result in demonstratives – clear morphosyntactic evidence, based on complementary distribution, or co-occurrence restrictions, with respect to other functional heads in the noun phrase.

2.4.2 Algonquian

Déchaine (1999) provides an account of the structure of the independent personal pronouns in Algonquian languages.¹¹ Algonquian languages are highly inflected, marking agreement with both subjects and objects. Therefore, independent pronouns are relatively rare in these languages, appearing only in contexts of emphasis or contrast. They are also morphologically complex, as can be seen with the subset of the Plains Cree pronominal paradigm in (23) (Déchaine 1999, p. 28; affixes are underlined):

- (23) a. nîya 'I'
b. nîyannân 'we(excl)'
c. wîya 'he, she, it'
d. wîyawwâw 'they(anim)'

Note that all of these forms consist of the root *îya*, plus agreement. Note further that this agreement is discontinuous: there is a prefix marking only person, and a suffix in the plural forms, marking both person and number.

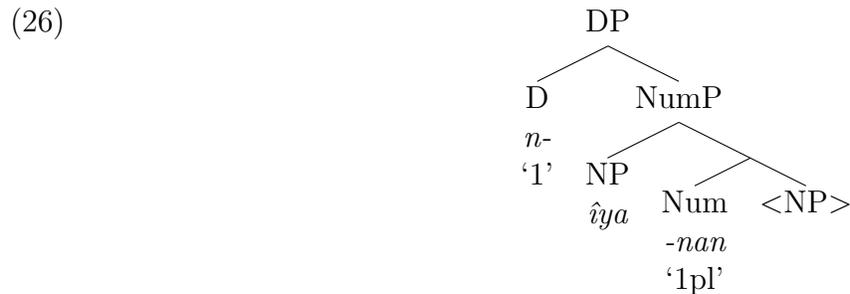
Déchaine claims that pronouns in these languages are DPs. This is due to the fact that they have the same morphological structures as possessed nouns, as shown in (24) and (25):

- (24) a. n-îya 'I/me'
b. n-îya-nân 'we/us(excl)'
- (25) a. n-môhkoman 'my knife'
b. n-môhkoman-nân 'our(excl) knife'

In both of these structures, there is a stem (*môhkoman* 'knife', and the pronominal root *îya*) which takes a prefix, marking person, and (in the case of plural pronouns or possessors) a suffix marking both person and number. The stem is analyzed as NP, and the prefixes and

¹¹Déchaine looks at Potawatomi, Ojibwe, Plains Cree and Blackfoot; with the exception of an additional suffix marking obviation in Blackfoot, these languages all show the same pattern of morphological marking in personal pronouns.

suffixes, as functional material encoding inflectional features, are analyzed as the heads of functional projections dominating NP. Since the suffix includes a number feature, it is analyzed as being of category Num (Déchaine’s Pers; see footnote 10), which in these languages happens to encode person as well; the prefix, encoding only person, is analyzed as D. Thus, after phrasal movement (outlined and justified in Déchaine 1999, Sections 2.2 and 4), the structure is as shown in (26) (Déchaine 1999, p. 29):



The identity of the root NP present in pronouns, *îya* in Plains Cree, is unclear, although it seems to be a marker of emphasis or contrast. This is due to the fact that, although this element does not appear alone, it appears with third person agreement as a marker of contrast. If, as Déchaine (1999, Section 3) argues, “third person” in Algonquian languages is not a person at all, but is instead a kind of default non-person, then the seemingly-third person form *wîya* with the contrastive interpretation can be taken as a non-possessed form of this morpheme. Thus, all personal pronouns in Algonquian can be seen as possessed contrast markers. This fits with the observation, briefly made above, that in Algonquian languages, independent pronouns are only used in contexts of contrast or emphasis anyway; otherwise, verb agreement suffices to identify participants in a sentence.

2.4.3 Mayan

Here I will briefly discuss the pronouns in Yucatec Maya and some other languages in the Mayan language family, which is another family whose pronouns tend to be fairly morphologically transparent. As in the Algonquian language family, Mayan languages have obligatory

agreement markers; thus, overt pronouns are present only in contexts of emphasis or contrast. In Mayan languages there are two sets of agreement markers: “Set A” (or ergative) markers, agreeing with the subject of transitive verbs and intransitive verbs in the imperfective aspect; and “Set B” (or absolutive) markers, agreeing with the object of transitive verbs and subject of intransitive verbs in perfective, subjunctive and extrafocal aspects (Butler 2011). The forms of the pronouns are always related to the Set B agreement markers. However, unlike in Algonquian, the stems to which the agreement markers attach are often prepositional, rather than pronominal, suggesting that pronouns in these languages are in fact pro-PP.

The pronouns in Yucatec Maya are shown below (Butler 2011, p. 20):

(27)

Person	Singular	Plural
First	teen < ti'+en	to'on < ti'+o'on
Second	tech < ti'+ech	te'ex < ti'+e'ex
Third	leti < le ti	leti'o'ob < le ti+o'ob

Each pronoun is a contracted form of a combination of the locative preposition *ti'* ‘to’, and the Set B agreement suffixes (the non-contracted forms are those on the right side of each cell in the table). The third person singular agreement suffix is null, and the plural suffix in the third person is the language’s plural morpheme that occurs in non-pronominal noun phrases; these pronouns must co-occur with the language’s definite article, *le*. Unlike in Algonquian languages, these cannot be analyzed as a possessed stem, as possession in Mayan languages is marked with the Set A affixes, not the Set B affixes seen in the pronouns Butler (2011). Thus, it seems that these are in fact full PPs; this corresponds with the generalization that these emphatic pronouns are adverbial elements in the clause, rather than arguments.¹²

A similar situation can be seen in the Tzeltal pronominal system, shown below, along

¹²While the interpretation of pronouns in Algonquian languages is similarly adverbial, the pro-PP analysis is not posited for these languages because, as illustrated in (24) and (25), the pronouns in these languages do correspond with possessed noun phrases.

with the Set B agreement markers themselves (Shklovsky 2012, p. 131):

(28)

Person	Singular		Plural	
	Agr	Pronoun	Agr	Pronoun
First (exclusive)	-on	jo'on	-on(r)yotik	jo'oryotik
Second	-at	ja'at	-ex	ja'ex
Third	-Ø	ja'	-ik	ja'ik

These pronouns are all formed by the affixation of the Set B agreement marker to the focus particle, *ja'*. In the first person, there is vowel harmony changing the first vowel from /a/ to /o/. The third person agreement suffix is null, so the pronoun is identical in form to the focus particle itself; note that, unlike in Yucatec Maya, Tzeltal does not require the definite article *te* with the third-person pronoun. Thus, Tzeltal is also amenable to analysis as a pro-PP language; again, the Set B agreement markers are used, rather than the Set A markers used on possessed noun phrases, and like the Yucatec Maya locative preposition, the focus particle is an optional adverbial element.

Thus, while this discussion remains as speculation, it seems that the Mayan languages are examples of languages in which the overt, emphatic pronouns are pro-PP; since P has been analyzed as a head in the nominal extended projection (Grimshaw 1991; van Riemsdijk 1998), this is a rare example of a language family in which pronouns realize an element larger than DP. This is likely only possible for pronouns that serve an adverbial (i.e., emphatic, contrastive) purpose, rather than argumental (although note that pronouns serving this purpose will not be pro-PP in all languages, as the Algonquian examples show). And note that unanswered questions remain with respect to these Mayan languages, such as the question of why the Yucatec Maya third person pronouns are preceded by the definite article; perhaps the definite article is an allomorph of the (normally) null third-person agreement marker, which happens to be a proclitic rather than a suffix.

2.4.4 Discussion

Here I have shown analyses of the structure of pronouns in three different language families where the morpho-syntactic properties of the pronouns (i.e., co-occurrence restrictions with respect to nominal functional categories, and overt structure of affixes) give clear evidence as to the structures of the pronouns. This shows that the realization of features by functional heads can vary between languages, and even within a language. Hebrew was an example of the latter, where in the first and second person pronouns, the features [definite], [person], [number] and [gender] are all situated in the D node, while in third person pronouns, [definite] and [person] are in D while [number] and [gender] are in Num. In contrast, in Algonquian, D contains [definite] and [person] features, and Num contains [number] and [person] features.¹³ In addition, unlike Hebrew, Algonquian has an NP root in its pronouns, which seems to be a morpheme of contrast or emphasis. Finally, Mayan languages present a new class of languages with pro-PP, which is expected to be possible only in languages where the overt pronouns are used adverbially, for purposes of contrast or emphasis.

This goes to show that, even in languages where the maximal projection of pronouns is DP, there can be variation in the internal structures of those pronouns. That is, in languages in which noun phrases realize full DPs, the overt material of pronouns may realize either a full DP or something less than that, but never a larger projection than non-pronominal noun phrases; and when the overt material realizes something less, covert material seems necessary to “complete” the DP. Thus, I claim that languages that lack the DP projection in noun phrases will not have DP pronouns. Evidence for this will be seen in the discussion of Japanese in Section 3. Thus, I predict that languages with DPs may show variation in their pronouns, but languages with no DP will be more constrained in that pro-DP will not be an option. On a more general note, the analyses discussed above show that even in languages with the same hierarchy of categories in their nominal fseq, the features can be distributed

¹³Since Algonquian languages have a grammatical gender system based on animacy, rather than a masculine/feminine system, and since independent pronouns all seem to require animate referents, it is difficult to say where the [gender] feature is specified in these languages.

differently among those heads. As we will see, there are languages with different fseqs, and languages where the maximal projection of a pronoun is something less than DP. Thus, while there are constraints on what heads may be absent or available, and in what sequence, and on where features can be realized, there is still a great deal of room for variation in nominal structure, and thus, in pronoun structure and category.

3 Noun Phrases in East Asian

3.1 Introduction

In this section, I compare nominal syntax in Japanese and Chinese (in particular, Mandarin) to illustrate the notion that pronominal category is predictable from the structure of a given language’s nominal extended projection (i.e., differences in the functional sequence or fseq). Noun phrases in Japanese and Mandarin have some superficial similarities, such as the lack of articles and obligatory plural marking, and the presence of numeral classifiers, that have lead some prominent researchers (e.g., Chierchia 1998; Bošković 2008) to group them in the same category. However, on closer inspection, subtle differences in the workings of these elements suggest that the two languages have fundamental differences in their nominal syntax, which are reflected in different fseqs in the two languages. These different fseqs in turn entail different categories for the languages’ pronouns.

Recall that pronouns differ cross-linguistically in various properties, including whether or not they can be modified by adjectives, genitives, and the like. Both Mandarin and Japanese allow much more modification of pronouns than is allowed in English, as shown below:

- (29) a. Zuo tian de ta bi jin tian de ta gen cong ming.
yesterday POSS him COMP.PRT today POSS he more smart
‘Yesterday’s he is smarter than today’s him.’ *Mandarin*
- b. Kyou-no kare-no hou-ga kinou-no kare yori atamagaii desu.
today-GEN he-GEN way-NOM yesterday-GEN he than smart is
‘Today’s he is smarter than yesterday’s him.’ *Japanese*

c. *Today's he is smarter than yesterday's him.

English

This would seem to suggest that in Mandarin and Japanese, pronouns are of category NP, unlike in English, where they are DP; this would be easily explained if the former languages lack the category D. However, recent analyses of Mandarin (e.g., Li 1999; Cheng and Sybesma 1999) involve the presence of the DP projection (at least in definite noun phrases), and there is further evidence that Mandarin pronouns at least sometimes must be of category D. For example, pronouns and proper names, but not common nouns, are permitted in a construction where they are affixed with the plural-like marker *-men* and followed by a numeral and classifier and optional head noun, as shown in (30) (Li 1999, p. 83):

- (30) a. Wo dui ta-men san-ge (ren) tebie hao.
I to them three-Cl (person) especially nice
'(lit.) I am especially nice to them three.'
- b. Wo dui XiaoQiang-men san-ge (ren) tebie hao.
I to XiaoQiang-MEN three-Cl (person) especially nice
'(lit.) I am especially nice to XiaoQiang (them) three persons.'
- c. *Wo dui xuesheng-men san-ge (ren) tebie hao.
I to student-MEN three-Cl (person) especially nice
'(lit.) I am especially nice to three students.'

This suggests that pronouns, like proper names but unlike common nouns, are of category D; this is supported by the fact that pronouns cannot be modified in this environment, as shown in (31):

- (31) *Nan-guo de ta-men san-ge mei you tu can jia huen-li.
sad DE them three-Cl didn't go participate wedding
'(intended) Sad them three didn't participate in the wedding.'

On the other hand, Japanese pronouns can be modified in that language's equivalent of this construction, suggesting that in Japanese, this context does not force the element to be of category D:

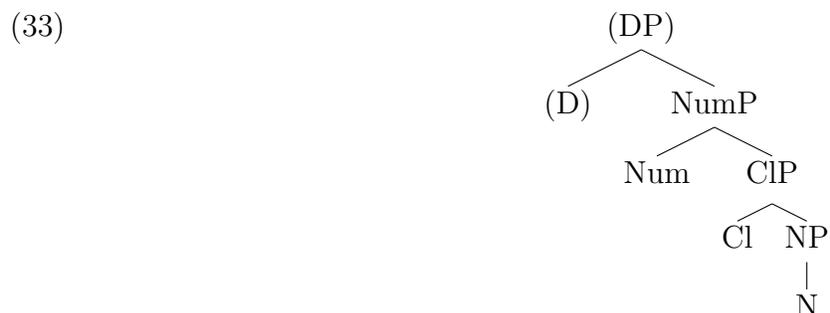
- (32) a. John-wa (ureshii) wareware yo-nin-o mita.
 John-TOP (happy) us four-Cl-ACC saw
 ‘John saw the (happy) four of us.’
- b. John-wa (chiisai) wareware yo-nin-o mita.
 John-TOP (small) us four-Cl-ACC saw
 ‘John saw the (small) four of us.’

Thus, while (29) would suggest that Japanese and Mandarin pronouns are the same category, the contrast in (31) and (32) indicates that there are differences between the two.

In this section I show that the differences between these languages result from two factors: differences in the “default” category of pronouns, and differences in the ease of coercion of pronouns from one syntactic category to another. In Section 3.2 I present and justify the different structures for the different languages, and show how this entails different syntactic categories for pronouns in Japanese and Mandarin. In Section 3.3 I outline a basic theory of coercion and show how it explains the fact that Mandarin pronouns, unlike those in English, can be modified. Finally, in Section 3.2.3 I briefly show that Korean patterns with Japanese, as another example of a language with pro-NP.

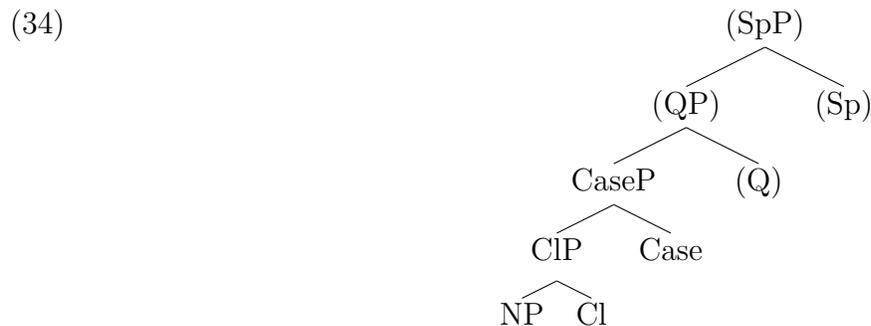
3.2 Structures

In this subsection I introduce the different structures I am assuming for the Mandarin and Japanese nominal fseqs, and provide justification for the differences between them. First I will present the different structures, and point out the differences between them. The structure for the Mandarin nominal fseq is as shown in (33):



In this model, D is the site where pronouns and demonstratives are generated; when the interpretation is definite, N must move to D as well. Numeral classifiers are generated in Cl, while the numeral itself is generated in spec-CIP. The plural-like *-men* is generated in Num, but attaches to nominal elements on their way to D (see Section 3.2.2 for more on *-men*). Lexical nouns are generated in N. The DP projection is absent in indefinite noun phrases. Note that, by the logic outlined in Section 2, this would imply that pronouns realize DP in Mandarin, since they are definite; this is indeed the case (in the absence of coercion), as will be seen below.

For the Japanese nominal fseq I assume the structure in (34).



In this model, once again nouns originate in N. Numeral classifiers are generated in CL, while numerals originate in spec-CIP. Case particles originate in Case, which is also involved in feature checking for specificity. Q is involved in quantification, with non-numeral quantifiers are generated in spec-QP; this projection is present only when necessary. Finally, Sp (short for “specific”) is present only in specific (as opposed to non-specific) noun phrases, and it never attracts elements to its specifier position.

Note the differences between the Mandarin and Japanese fseqs as presented above:

- In Japanese, pronouns are N. In Mandarin, they are D by default, but can be coerced into N under certain conditions.
- Japanese has SpP, indicating specificity, while Mandarin has DP, indicating definiteness. Specificity and definiteness are closely related, as definite expressions are neces-

sarily specific, but the converse is not necessarily true – there can be specific indefinite expressions (Lyons 1999).

- The second-to-top projection in Japanese is QP, while in Mandarin it is NumP. This NumP is the site of the plural marker *-men*, for which Japanese does not have a direct analogy (see Section 3.2.4 for discussion of *-tati*, a similar morpheme in Japanese). The Japanese Q head is generally null, but hosts non-numeral quantifiers in its specifier position.¹⁴
- Japanese has CaseP, and Mandarin does not. This is because Japanese has overt case particles, and Mandarin does not.
- Mandarin is head-initial (left-branching), while Japanese is head-final (right-branching).

For the remainder of this subsection I provide justification of the different structures (33) and (34), drawing on insight from independent analyses of various phenomena in each language, and comparisons between the two languages. I also show evidence that Korean is another example of a language with pro-NP, alongside Japanese.

3.2.1 Japanese Nominal Structure

Here I provide evidence for the structure posited for the Japanese nominal fseq in (34), above. I will draw largely from the work of Watanabe (2006, 2008, 2010), along with some insights from Huang and Ochi (2012). These authors are accounting for the phenomenon of quantifier float in Japanese. In Japanese, a numeral and a classifier (hereafter, Nml+Cl) may appear in a variety of configurations with respect to the head noun, exemplified in (35) (adapted from Watanabe 2006, p. 244):¹⁵

¹⁴Note that in these analyses, numerals are generated in spec-CIP, rather than in #, as in Borer’s model. This may reflect cross-linguistic differences in the nature of numerals, either as heads in the fseq or as phrases in a specifier position; as either seems to fit with the semantic composition of the noun phrase, this does not seem to present a problem. This may suggest that languages differ in terms of how quantification is introduced, as long as there are not multiple contradictory sites of quantification.

¹⁵Watanabe (2006) claims that there is a fourth possible order, where Nml+Cl immediately follows the case particle. However, this order is semantically no different from the floating order, and later works on

- (35) a. John-wa kinoo hon **san-satu**-o katta.
 John-TOP yesterday book three-Cl-ACC bought
 ‘John bought three books.’ *Post-nominal*
- b. John-wa kinoo **san-satu**-no hon-o katta. *Pre-nominal*
- c. John-wa hon-o kinoo **san-satu** katta. *Floating*

Thus, the Nml+Cl *san-satu* may appear following the noun, but preceding the case particle as in (35a); preceding the head noun as in (35b), with the genitive-like particle *no* inserted (this particle is required for any non-clausal modification in the noun phrase; Watanabe 2006); or outside the noun phrase as in (35c), where the clausal adverbial *kinoo* ‘yesterday’ intervenes between Nml+Cl and the rest of the noun phrase.

Watanabe (2006, 2008, 2010) proposes an analysis by which the configurations in (35) can all be derived from the same underlying structure, through massive phrasal movements. That underlying structure is shown in (36) (Watanabe 2006, p. 252):¹⁶



Note that this is identical to the structure in (34) above, except for two things: (i) the top-most projection in Watanabe’s analysis is DP, while in the tree proposed above it is SpP; (ii) in Watanabe’s analysis the DP and QP projections are obligatory, while in (34) they are optional. The reasons for these differences will be discussed below. First, however, I will outline Watanabe’s phrasal-movement analysis of the configurations in (35).

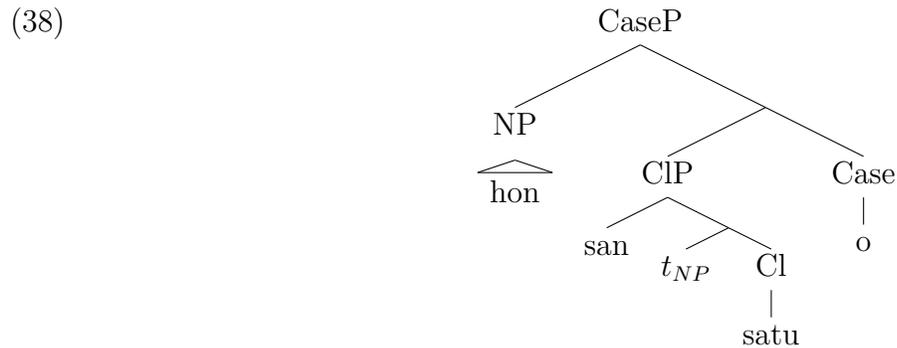
the subject (e.g., Huang and Ochi 2012, as well as Watanabe 2008, 2010) do not treat it as a separate phenomenon, so I will follow the later versions of the analysis here in sticking to the three orders in (35).

¹⁶In Watanabe’s work, the CIP projection is actually labeled #P; I am calling it CIP here for the sake of consistency with the rest of the structures in this paper. I retain Watanabe’s DP label, rather than my SpP, in the discussion of Watanabe’s version of the theory, since the difference between them is theoretical, rather than simply nominal.

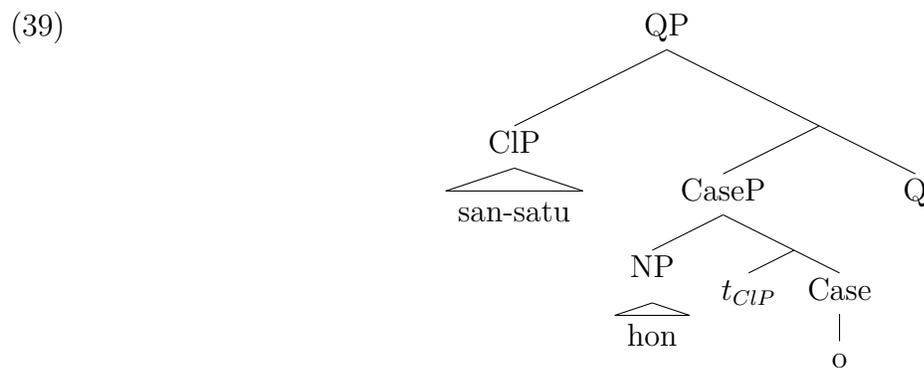
According to Watanabe (2006, pp. 253–256), the orders in (35) are derived as follows. First the following structure is generated, with the classifier occupying the head Cl and the numeral in spec-CIP:



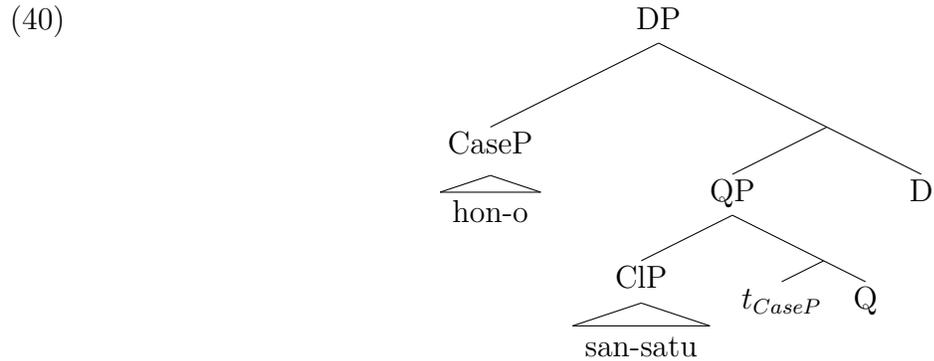
Next, CaseP merges and NP moves to its specifier position, triggered by an EPP feature on the Case head:



If no further movement occurs after the Q and D heads are merged, the postnominal order in (35a) is obtained, with the Nml+Cl between the head noun and the case particle. However, there is optional raising of CIP to spec-QP, which Watanabe speculates may be related to mass-count encoding, as shown:



This gives the prenominal order in (35b), if the merger of D does not trigger movement (the morpheme *-no* is considered to be inserted by a post-syntactic morphological rule, discussed in detail in Watanabe 2010). If D enters a feature-checking relationship with CaseP, arguably for reasons of specificity, then CaseP moves to spec-DP:



This results in an order where the Nml+Cl follows the case particle. Finally, CIP may scramble outside the DP entirely, giving the quantifier float order in (35c).

Watanabe (2006) provides extensive justification for the agreement relations triggering each of the phrasal movements described above. Here I will only go into the specificity-related features triggering the movement to spec-DP in (40). According to Watanabe, the floating order in (35c) differs from the other two orders in allowing only a non-specific interpretation. This can be illustrated with examples like (41) (Watanabe 2006, p. 298):

- (41)
- a. Gakusei **san-nin**-ga eigo-ga umai.
 student three-Cl-NOM English-NOM good
 ‘Three students are good at English.’
 - b. **San-nin**-no gakusei-ga eigo-ga umai.
 - c. *?Gakusei-ga **san-nin** eigo-ga umai.

Watanabe follows Diesing (1992) in claiming that an individual-level predicate like ‘be good at English’ is compatible only with a specific noun phrase; thus, the impossibility of (41c)

suggests that the post-case-particle order can give only a specific reading.¹⁷ These specificity differences are the basis of arguments used by Huang and Ochi (2012) in proposing their own analysis of Japanese noun phrases, which influenced my own proposal in (34) above. Here I will briefly outline Huang and Ochi’s proposal.

Huang and Ochi (2012) present a kind of hybrid analysis, using a version of Watanabe’s noun phrase structure of (36) (with important differences that are discussed below) for some orders and the more standard adjunction analysis (proposed by Saito et al. 2008, among others, and shown below) for others.



Specifically, Huang and Ochi posit that for the pre-nominal order of (35b), the adjunction structure (42) is involved, and the post-nominal and floating orders (35a) and (35c) have a variation on Watanabe’s structure, where DP, QP and CaseP are replaced by the same XP that is optional in (42). This XP is agnostically left undefined, but Huang and Ochi claim that it is present in specific noun phrases, and absent when the noun phrase is non-specific.

I will discuss only briefly Huang and Ochi’s arguments for claiming that the adjunction structure obtains in the pre-nominal order. There are two main arguments. One involves scope relations between Nml+Cl and universal quantifiers like *subete* ‘all’, which may co-occur in a noun phrase only in certain configurations. This argument loses force in the face of Watanabe (2010), in which such co-occurrences are discussed in detail. The other argument involves ‘N’-ellipsis’ structures, which are also discussed in detail in Watanabe (2010) and shown to be unproblematic. Again, I will not discuss these debates further as

¹⁷This cannot be strictly true; the sentence *No-one is good at English* would seem to be a counterexample (Robert Truswell, p.c.). Perhaps a more accurate generalization would be to claim that non-specific existential sentences are disallowed with individual-level predicates, but non-existential sentences do not have a specificity restriction.

they do not directly bear on the matter at hand. However, Huang and Ochi present new data on specificity that I describe below.

Recall from the discussion of (41) that Watanabe notes that the floating orders forces a non-specific reading. According to Watanabe, the other two orders are ambiguous in specificity. However, Huang and Ochi (2012) argue that the post-nominal order of (35a) forces a specific reading (or at least it strongly prefers a specific reading). Evidence is provided in (43) (pp. 7–8):

- (43) a. **San-nin**-no itoko-ga iru hito-wa te-o agete kudasai.
 three-Cl-GEN cousin-NOM have person-TOP hand-ACC raise please
 ‘Those of you who have three cousins, please raise your hand.’
- b. Itoko-ga **san-nin** iru hito-wa te-o agete kudasai.
 cousin-NOM three-Cl have person-TOP hand-ACC raise please
 ‘Those of you who have three cousins, please raise your hand.’
- c. ??Itoko **san-nin**-ga iru hito-wa te-o agete kudasai.
 cousin three-Cl-NOM have person-TOP hand-ACC raise please
 ‘Those of you who have three cousins, please raise your hand.’

This context forces a non-specific reading of the noun phrase, and (43c), with the post-nominal order, is marginal in this context, suggesting that this order is most compatible with a specific reading. Huang and Ochi follow other authors in making the generalization in (44) (p. 9):¹⁸

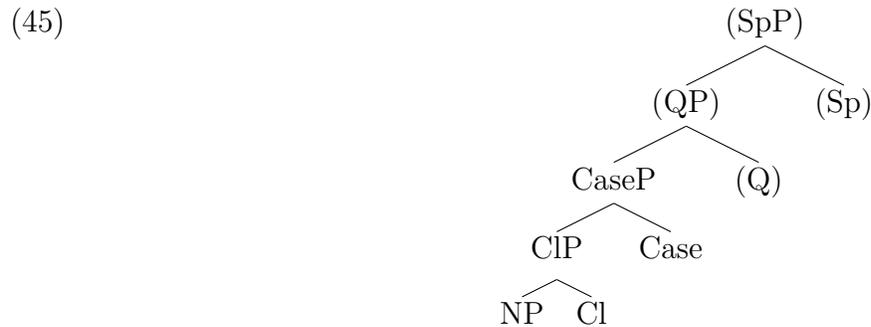
- (44) A specific indefinite nominal has a larger structure than a non-specific indefinite nominal.

This “larger structure” is indicated in Huang and Ochi’s trees as XP. They claim that the pre-nominal order is ambiguous in specificity because the XP is optional here. Recall that for the post-nominal and floating orders, they adopt a version of Watanabe’s analysis that differs from (36) in that CaseP, QP and DP are replaced with the single projection XP. On

¹⁸Such claims can be found in Carlson (1977), Ritter (1995), and many others.

Huang and Ochi's analysis, when the structure in (37) is generated, the NP must move so that the numeral and classifier can be adjacent; if there is an XP present, NP moves to spec-XP, giving the post-nominal order (35a), and if XP is not present, NP moves outside the nominal projection, resulting in quantifier float.

Thus, in light of these arguments, I propose the structure in (34), repeated as (45), for the Japanese nominal fseq:



My reasoning is as follows. From Watanabe (2006, 2008, 2010) I retain the unified analysis of the three configurations in (35), unlike Huang and Ochi (2012), in which there are two different underlying structures. However, from Huang and Ochi I draw the insight that, in light of the generalization in (44), it makes no sense to posit a DP that is present even in non-specific noun phrases, and especially not if it only undergoes agreement, triggering movement to spec-DP, in non-specific sentences, as Watanabe (2006) proposes. Thus, I follow Huang and Ochi in claiming that there is a functional category that is present only in specific noun phrases, which I call SpP. Similarly, I claim that QP is present only when it is required, for example when it undergoes agreement with CIP to trigger movement to its spec position, or to host non-numeral quantifiers.

On this analysis, the post-nominal and pre-nominal orders (35a) and (35b) receive the same phrasal-movement analysis as given by Watanabe (i.e., (38) and (39), respectively), modulo the lack of DP. The floating order (35c) occurs when either CaseP or CIP moves out of the noun phrase entirely from the configuration in (39), since there is no spec-DP to move into (The present data do not decide between CaseP and CIP moving out of the QP; I will

leave the question open). The floating order is impossible when SpP is present, which follows if SpP is a phase head, as has been suggested for DP (e.g., Chomsky 2005), which simply does not enter into any agreement relations with any of its subconstituents. If, as Watanabe argues, phrasal movement is triggered by agreement, then this means that no elements could raise to spec-SpP, and thus, by the Phase Impenetrability Condition (Chomsky 2000, p. 108), elements could not be extracted from SpP and the floating order is impossible.¹⁹

I have now shown why I have been assuming the structure in (34)/(45) for the Japanese nominal fseq. Note that on this analysis, pronouns are forced to realize N, like nouns – thus, the category of pronouns can be inferred from their distribution, which directly explains the modification facts noted above. This is because they occur in the same syntactic configurations as nouns. First, like nouns, Japanese pronouns must occur with a case particle when they appear in a sentence, as shown in (46):

- (46) Wareware-*(ga) piza-*(o) tabeta.
 we-NOM pizza-ACC ate
 ‘We ate pizza.’

Thus, the empirical data force the conclusion that pronouns must realize some constituent lower than Case in the fseq (see Neeleman and Szendrői 2007 for a similar analysis). In addition, also like nouns, pronouns may occur with a following Nml+Cl in the noun phrase, even if they are modified, as shown above in (32) and repeated in (47):

- (47) a. John-wa (ureshii) wareware yo-nin-o mita.
 John-TOP (happy) us four-Cl-ACC saw
 ‘John saw the (happy) four of us.’
 b. John-wa (chiisai) wareware yo-nin-o mita.
 John-TOP (small) us four-Cl-ACC saw
 ‘John saw the (small) four of us.’

¹⁹This analysis does not explain why the post-nominal order is restricted to specific readings as Huang and Ochi (2012) note; however, the fact that Watanabe did not report this restriction, combined with the ‘??’ on (43c) (rather than complete ungrammaticality), suggests that this is more a preference than a restriction, which is compatible with this analysis.

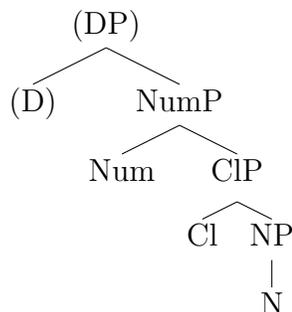
This means that pronouns realize something lower than #P, which leaves only N. This line of reasoning does not exclude the possibility that Japanese pronouns realize the entire NP, by some mechanism of phrasal spell-out. However, since the evidence presented here demonstrates that Japanese pronouns show the same syntactic behaviour as lexical nouns, I will assume that, like nouns, they are of category N. It will be seen below that the construction in (47) is disallowed in Mandarin. Thus, pronouns are by default N heads in Japanese, which explains why they accept modification as freely as lexical nouns in this language. This analysis not only fits the evidence from modification shown in Section 3.1, but it is also forced by the structure in (34)/(45), which was developed for the independent purpose of explaining quantifier float.

3.2.2 Mandarin Nominal Structure

Here, I provide justification for the structure posited in (33) for the structure of the Mandarin nominal fseq. I will draw largely from the work of Li (1999), who develops a proposal to account for the distribution of the plural-like marker, *-men*, in Mandarin.²⁰ Previous accounts (e.g., Iljic 1994) analyzed *-men* as an “associative plural” marker rather than a true plural; that is, *N-men* would mean ‘N and N’s associates’, rather than meaning strictly ‘plural N’. However, Li shows that, while the morpheme can have this meaning with proper names and pronouns, when *-men* is affixed to a common noun it signals as a combination of plurality and definiteness; Li provides a unified account of these two denotations, and other properties of *-men*, using the structure in (33), repeated here as (48):

²⁰See Section 3.2.4 below for discussion of a similar morpheme in Japanese, and explanation of why this morpheme does not warrant the same analysis as *-men*.

(48)



While *-men* may seem at first glance like an ordinary plural marker, Li (1999) notes four peculiarities. First, it can only attach to animate nouns (nouns denoting humans, or pronouns). Second, it cannot co-occur with Nml+Cl, as shown in (49) (examples from Iljic 1994, p. 93):

- (49) a. san-ge pengyou
three-Cl friend
'three friends'
- b. *san-ge pengyou-men
three-Cl friend-MEN
'three friends'

Third, a noun phrase with *-men* is always definite (Li 1999, p. 78):²¹

- (50) a. Wu qu zhao haizi-men.
I go find child-MEN
'I will go find the children.'
- b. Wu qu zhao haizi.
I go find child
'I will go find the/some child/children.'

To further illustrate this point, Li cites Iljic (1994), who notes that *-men* is disallowed in existential sentences, whether positive (51) or negative (52) (p. 94):

- (51) a. *You ren-men.
have man-MEN

²¹Note that in Mandarin, as in Japanese, bare noun phrases are unspecified for number.

‘*There are the men.’²²

- b. You ren.
have man
‘There is someone / there are some people.’

(52) a. *Mei you ren-men.
NEG have man-MEN
‘*There are not the men.’

- b. Mei you ren.
NEG have man
‘There is nobody.’

The fourth peculiarity is the fact that, when a proper name is suffixed with *-men*, it can have an associative or collective reading, denoting a group consisting of the person denoted by the proper name and others associated with that person (53) (Li 1999, p. 78):

(53) XiaoQiang-men shenme shihou lai?
XiaoQiang-MEN what time come
‘When are XiaoQiang and the others coming?’

However, Li notes (p. 80) that the pure plural reading can also obtain when *-men* is suffixed to a proper noun. Thus, in addition to the reading in (53), *XiaoQiang-men* can mean ‘multiple people with the name or attributes of XiaoQiang’, which happens to be the only reading available when a plural marker occurs on a proper name in languages like English, with only the pure plural reading.

Li claims that these properties²³ can be accounted for by proposing that *-men* is a plural marker generated in Num that can only attach to a noun in D. The second part of this explanation, the claim that *-men* can only attach to an item in D, while accounting for the

²²This example, and (52a), are grammatical on certain readings; however, with the desired existential reading, they are disallowed in English as in Mandarin.

²³The first peculiarity, the fact that *-men* can only attach to animate nouns, is not accounted for by this structure *per se*. However, Corbett (2000) notes that in many (or perhaps all) languages, the expression of number is constrained by the animacy hierarchy, such that in a given language, more animate nouns can receive comparatively more number specification than less animate nouns. Thus, the fact that *-men* can only attach to animate nouns is typologically quite normal; the Korean plural morpheme *-tul* has the same animacy restriction.

facts listed above, is largely stipulative, although see below for some possible explanations for it. The analysis of *-men* as a plural marker in Num, on the other hand, follows fairly standard analyses of plurality cross-linguistically (see, e.g., Ritter 1991, 1995, and references cited therein). Thus, *-men* may only attach either to a noun that moves through Num on the way to D, or to an item generated in D, thus deriving the fact that noun phrases containing *-men* are always definite. This could be accomplished by stipulating that *-men* is licensed only in the presence of DP, as proposed by Ghomeshi (2003) for the Persian plural marker. Furthermore, this explains the prohibition in (49b), against *N-men* following *Nml+Cl*, since, if *-men* always ends up in D, it must always precede items in *ClP*.

The derivation of the fourth peculiarity, that proper names may receive a collective plural reading with *-men*, involves differentiating between items generated in D, and items that end up there through head movement. Recall from above that *-men* can give two readings, a collective plural and a pure plural, when suffixed to a proper name. Common nouns, on the other hand, may receive only the pure plural reading. The reading obtained when *-men* is attached to a pronoun is difficult to determine; it is the ordinary reading seen with plural pronouns, but as Iljic (1994) points out, there is a sense in which all plural pronouns are collective plurals. For instance, the English pronoun *we* does not denote multiple speakers, as one might expect from the plural form of the first person pronoun; instead, it denotes the speaker and some associated group, which is identical to the collective reading obtained with proper names. Thus, for common nouns *-men* is a pure plural, for proper names it is either pure or collective plural, and for pronouns it is only a collective plural. Li (1999) claims that the pure plural reading obtains when *-men* attaches to an item that has moved from N to D, and the collective reading obtains with items generated in D. Following Longobardi (1994), Li assumes that proper names are normally generated in D. However, Li claims that they are ambiguous, and may also be generated in N, with different semantics; in her words, “a proper name can be base-generated in D to refer to a definite individual by name. In addition, it can function like a common noun, base-generated in N, and denote person(s)

with the same name (*I met two Bills at the party. I like the Bill you like*) or denote person(s) with the same characteristics” (Li 1999, p. 84).²⁴ Thus, the two readings of *-men*, and the restrictions on their distribution, are explained.

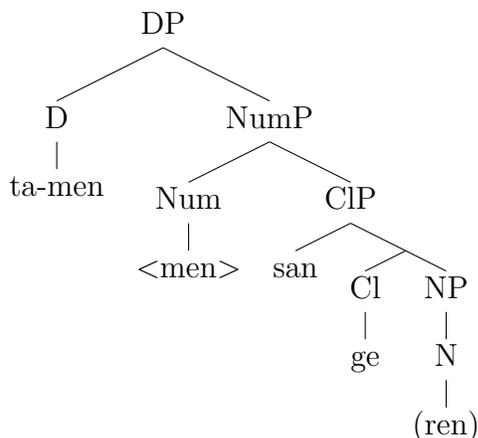
These claims are corroborated by the fact that there is a structure involving *-men* in which only items generated in D are possible, and items moved from N are disallowed, and the only possible reading in this structure is that of collective plural. This structure is one in which a pronoun or proper name suffixed with *-men* may precede a Nml+Cl sequence, with an optional following head noun. This structure is exemplified in (54) (Li 1999, p. 83):

- (54) a. Wo dui ta-men san-ge (ren) tebie hao.
 I to them three-Cl (person) especially nice
 ‘(lit.) I am especially nice to them three.’
- b. Wo dui XiaoQiang-men san-ge (ren) tebie hao.
 I to XiaoQiang-MEN three-Cl (person) especially nice
 ‘(lit.) I am especially nice to XiaoQiang (them) three persons.’
- c. *Wo dui xuesheng-men san-ge (ren) tebie hao.
 I to student-MEN three-Cl (person) especially nice
 ‘(lit.) I am especially nice to three students.’

Note that (54b) is available only on the collective reading (‘XiaoQiang and two friends’), not the pure plural reading (‘three XiaoQiangs’). Along with the ungrammaticality of the structure with a common noun in (54c), this demonstrates that only the collective reading is available in this structure. This follows if the presence of Nml+Cl blocks movement from N to D, since the Cl head is filled; thus, the structure of (54a) must be as in (55):

²⁴This does not exactly reflect the analysis proposed by Longobardi (1994). Longobardi’s analysis involves different readings of proper names depending on their position in N versus D at S-Structure, but he claims that in all cases they are base-generated in N. I will stick with the Li-style coercion analysis instead, given that the current theoretical model does not distinguish different levels of representation, unlike the GB model assumed in Longobardi’s analysis.

(55)



Note that, in this construction, the pronoun realizes the D head, and cannot have been generated in N. Further evidence for the status of pronouns as D in this construction comes from the fact that pronouns may not be modified in this context, as shown in (56), repeated from (31) above:

- (56) *Nan-guo de ta-men san-ge mei you tu can jia huen-li.
sad DE them three-Cl didn't go participate wedding
'(intended) Sad them three didn't participate in the wedding.'

Thus, Li's analysis explains the relevant facts about the distribution of *-men* and Nml+Cl with respect to common nouns, proper names and pronouns. The model that I adopt, (33)/(48), is the same as Li's, except for some slight modifications involving the location of numerals. Li places numerals in spec-NumP, following Ritter (1991, 1995). However, by analogy with the structure proposed above for Japanese (34)/(45), I propose that the numeral is in fact in spec-ClP. Crucially, this analysis also explains why, unlike Japanese pronouns, Mandarin pronouns may not be modified by adjectives as easily as nouns can: pronouns are generated in D, rather than N. In Section 3.3, I will show why Mandarin pronouns are more readily modified than those in English. First, however, I will demonstrate that Korean patterns with Japanese in terms of pronoun modifiability, and thus pronominal category.

3.2.3 Korean Pronouns are N

Here I briefly discuss data from Korean that suggest that pronouns in this language are of category N(P), as in Japanese. Recall from the discussions of Japanese and Mandarin that there is a structure that is only possible in languages where pronouns are maximally NP; this is the structure in which a pronoun is modified by an adjective and counted by a Nml+Cl. In Japanese, this structure is possible, as shown in (47), while in Mandarin it is disallowed due to the fact that the presence of Nml+Cl forces the pronoun to be generated as D, as shown in (65c). This structure is permissible in Korean, as the following data show:²⁵

- (57) a. Ku(tul) sey myeng-i kyelhonsik-ey ka-cian-ass-ta
they three Cl-NOM wedding-to go-NEG-PST-DECL
'Them three didn't go to the wedding.'
- b. Sulphun kutul-un kyelhonsik-ey ka-cian-ass-ta
sad they-TOP wedding-to go-NEG-PST-DECL
'Sad them didn't go to the wedding.'
- c. Sulphun ku(tul) sey myeng-i kyelhonsik-ey ka-cian-ass-ta
sad they three Cl-NOM wedding-to go-NEG-PST-DECL
'Sad them three didn't go to the wedding.'

Thus, Korean pronouns must also be of category N(P).

3.2.4 Residual issues

In the Japanese and Mandarin fseqs proposed and defended above ((34) and (33), respectively), Mandarin is claimed to have a NumP projection, and Japanese does not. This Num in Mandarin is the site where the plural marker *-men* is generated. The lack of a NumP projection in Japanese is justifiable only if there is no equivalent plural morpheme in Japanese. Japanese does in fact have the plural-like morpheme *-tati*, which at first glance looks like an equivalent of *-men*. However, some researchers (Nakanishi and Tomioka 2004, Ueda and

²⁵The plural marker *-tul* may be present in all cases, although it is only required in (57b), and the other examples sound better without it; this is due to the fact that it is not necessary when there is another overt marker of plurality, such as Nml+Cl.

Haraguchi 2008) have shown that there are some crucial differences between *-tati* and *-men*. Here I will present those differences and claim that this demands an alternative analysis for *-tati*. However, although I will suggest some possibilities, providing a proper analysis for *-tati* is outside the scope of the present paper.

The first difference is that *-tati*, unlike *-men*, may co-occur with Nml+Cl (under certain conditions). Recall from (49b) that *-men* is ungrammatical in the presence of Nml+Cl; the Japanese equivalent is degraded as well, as shown in (58) (Nakanishi and Tomioka 2004, p. 113):

- (58) ??san-nin-no gakusei-tati
 three-Cl-MOD student-TATI

However, this degraded status obtains only when the number is small and precise. When a larger, vague number is used, the result is grammatical, while the Mandarin equivalent with *-men* is still ungrammatical, as shown in (59) (Nakanishi and Tomioka 2004, pp. 119–120):

- (59) a. *chau-guo 200-ge haizi-men
 more-than 200-Cl child-MEN
 ‘more than 200 children’ *Mandarin*
- b. 200-nin-izyoo gakusei-tati
 200-Cl-or.more student-TATI
 ‘200 or more students’ *Japanese*

Thus, unlike in Mandarin, there must be no structural element blocking the derivation of a noun phrase with both Nml+Cl and a common noun suffixed with *-tati*.

Another difference is that *-tati*, unlike *-men*, does not force a definite interpretation. The denotation of a noun phrase with *-tati* may seem at first glance to be definite; however, syntactic tests show that it is not necessarily definite. Specifically, it can occur in existential sentences (60a), it can appear with *donna* ‘what kind of’ (60b), and it can antecede a sluiced *wh*-phrase (60c), all of which should be impossible with definites (Nakanishi and Tomioka 2004, pp. 120–123)

- (60) a. Kooen-ni gakusei-tati-ga ita.
park-LOC student-TATI-NOM existed
'There were students in the park.'
- b. Donna gakusei-tati-ga kita-no?
what.kind.of student-TATI-NOM came-Q
'What kind of student came?'
- c. Inoue-sensei-no ie-ni kodomo-tati-ga atumatta-to-kiita-kedo,
Inoue-prof.-GEN house-LOC child-TATI-NOM gathered-COMP-heard-while
watasi-wa dono kodomo-tati-ga sira-nai
I-TOP which child-TATI-NOM know-NEG
'(I) have heard that children gathered at Prof. Inoue's house, but I don't know
which children.'

Thus, *-tati* should not be analyzed as forcing a definite interpretation.

Nakanishi and Tomioka note further that *-tati* can receive the “associative” or collective reading with both proper names and common nouns (pp. 124–126):

- (61) a. Taroo-tati-wa moo kaetta.
Taroo-TATI-TOP already went.home
'The group of people represented by Taroo went home already.'
- b. Kyoō kooen-de gakusei-tati-no demo-ga atta.
today park-LOC student-TATI-GEN demonstration-NOM existed
'Today, there was a demonstration by (the) students (and possibly non-students)
at the park.'

Mandarin *-men*, on the other hand, receives the collective reading only with pronouns and proper names, as discussed in the previous subsection. Furthermore, Ueda and Haraguchi (2008) observe that multiple instances of *-tati* can co-occur on a single noun (62), while this is impossible for *-men* (63) (p. 237):

- (62) a. gakusei-tati-tati
student-TATI-TATI
'the students and their associates'

- b. Taroo-tati-tati
Taroo-TATI-TATI
'Taroo and his associates and their associates'
- (63)
- a. *xuesheng-men-men
student-MEN-MEN
 - b. *XiaoQiang-men-men
XiaoQiang-MEN-MEN

Thus, the differences between *-tati* and *-men* suggest that they should not receive the same analysis. Specifically, the justifications for calling *-men* a realization of Num, which attaches to an element in D, do not apply in the Japanese case. Again, I will not attempt to provide a full analysis of the status of *-tati* here. However, such an analysis may well follow Wiltschko's (2008) claim that plurality may be realized in different places in the nominal fseq in different languages. For example, Butler (2011) showed evidence that the plural marker in Yucatec Maya is in fact adjoined to DP, rather than a head in the fseq. Perhaps *-tati*, too, is best characterized as an adjunct, which would explain the possibility of multiple *-tatis* in (62), although *-tati* will adjoin lower in the fseq in order to obtain the proper word order. Whatever the eventual analysis, it will suffice for now to say that *-tati*, unlike *-men*, is not an instance of Num, and so the Japanese nominal fseq lacks this projection.

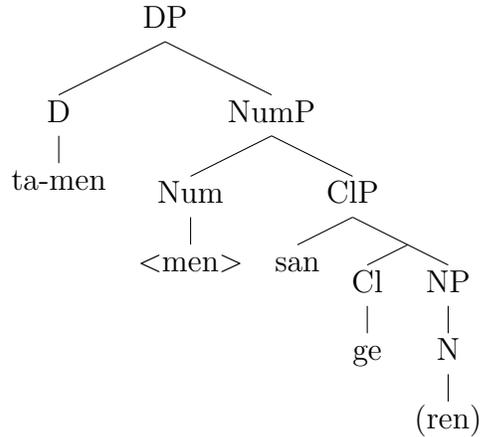
3.3 Coercion

In Section 3.2, I presented structures for the Japanese and Mandarin fseqs, motivated by independent concerns regarding the interactions of numerals, classifiers and plurality, which entail that pronouns in Japanese are N, and those in Mandarin are D. However, this does not explain why Mandarin pronouns are like those in Japanese in that they are modifiable, while English pronouns, which are also of category D, are not. I propose that this is because Mandarin pronouns may be coerced from D to N, while English pronouns are not as easily coerced in the same way. Here I develop a basic theory of coercion that accounts for this difference.

The claim that elements of category D can be coerced into N is not a new one; in fact, Li (1999) proposes such a mechanism for proper names to explain the different readings resulting when a name is suffixed with *-men*, as discussed in Section 3.2.2. To repeat, she claims that “a proper name can be base-generated in D to refer to a definite individual by name. In addition, it can function like a common noun, base-generated in N, and denote person(s) with the same name (*I met two Bills at the party. I like the Bill you like*) or denote person(s) with the same characteristics” (Li 1999, p. 84). Thus, when a proper name is generated in D, it directly denotes that individual – this might be seen as the “default” interpretation of a proper name. If it is generated in N, on the other hand, it has a predicative interpretation, true of individuals with that same name or the characteristics of that person – this is not the usual interpretation of names, so it can be seen as the coerced reading. Cowper and Hall (2009) propose a similar process for English pronouns, explaining why they can participate in the syntax and morphology like common nouns, creating expressions like *Is that a he or a she?* and compounds like *me-generation*. They propose that this is an instance of “a type of coercion, from featural to non-featural meaning” (p. 108). In other words, the pronouns in this case are not interpreted as simply a set of ϕ -features; instead, they denote masculinity or femininity in the case of *he* and *she*, and speaker- and listener-orientation in the case of *me* and *you*, without directly referring to individuals in the context.

I propose that this same kind of coercion is taking place when Mandarin pronouns are being modified: they are generated in N, rather than D. Evidence for this coercion comes from the construction described in Section 3.2.2, exemplified in (54), with the structure in (55), repeated below:

(64)



This structure is only possible if the pronoun is base-generated in D; recall that it is not possible with common nouns, or with proper names in their coerced interpretation. The theory sketched here predicts that as a result, pronouns in this structure cannot be coerced to N, and so they will not be modifiable at all. As seen above in (31)/(56), this prediction is borne out; the relevant example is repeated in (65c), along with examples where the adjective and Nml+Cl do not co-occur, to illustrate that it is, in fact, the co-occurrence of the two modifiers that force the ungrammaticality:

- (65)
- a. Nan-guo de ta-men mei you tu can jia huen-li.
sad DE them didn't go participate wedding
'(lit.) Sad them didn't participate in the wedding.'
 - b. Ta-men san-ge mei you tu can jia huen-li.
them three-Cl didn't go participate wedding
'Them three didn't participate in the wedding.'
 - c. *Nan-guo de ta-men san-ge mei you tu can jia huen-li.
sad DE them three-Cl didn't go participate wedding
'(intended) Sad them three didn't participate in the wedding.'

These examples show that the pronoun may co-occur with an adjective (65a), and it may co-occur with a following Nml+Cl (65b); however, it may not co-occur with both at the same time (65c). This is because, in this construction, the classifier blocks movement from N to D, so pronouns may only occur in this construction if they are base-generated in D, meaning they cannot be modified.

This analysis raises the question of why English pronouns, which (as claimed by Cowper and Hall 2009) may also be coerced in this way, are still less modifiable than those in Mandarin. I have no definite answer here, but the solution may have to do with the presence of other overt functional elements in the nominal projection. In English and other languages with overt articles, a head noun rarely surfaces without an accompanying article; this means that if a pronoun is coerced for the purposes of modification, an article is required, hence **(the) real me*. I suggest that it is this overt difference in the makeup of the nominal extended projection that determines the ease of coercion: in languages where such coercion requires the addition of overt functional material, it will be restricted to idiomatic use, while in languages where the coerced form requires no more overt material than the non-coerced form, it will be more free.²⁶ Once again, full exploration of this proposal is outside the scope of this paper, so this will remain as a suggestion here.

Thus, I have proposed that there are (at least) two factors at play in determining to what extent a language's pronouns will be modifiable: structure of the nominal fseq, and ease of coercion. And since the latter factor depends on the overt or covert nature of different functional heads, I am in fact proposing that the modifiability of pronouns is *entirely* determined by the properties of nominal projections in that language. Thus, the system proposed here predicts at least three different kinds of languages, with respect to pronoun modifiability: (i) languages where the pronouns are (by default) generated in N, so pronouns are modified freely (Japanese); (ii) languages where pronouns are generated in D, and can be coerced to N without functional material, so pronouns can be modified freely (Mandarin); and (iii) languages where pronouns are generated in D, and can only be coerced to N in the presence of functional material, so pronouns can be modified only in idiomatic usage (English). This provides a framework which allows us to predict the modifiability of a

²⁶Note that modification of pronouns in Mandarin is not entirely free. It seems that monosyllabic adjectives may not modify pronouns in this context; e.g., **xiao de tamen* 'small them', **hao de tamen* 'good them'. However, this restriction extends to non-pronominal nouns as well: **xiao de haizi* 'small child', **hao de xuesheng* 'good student' (Meng Yang, p.c.). Thus, there seems to be a phonological constraint on modification in Mandarin. An explanation for constraint is outside the scope of this paper.

language's pronouns, given only a structural analysis of that language's nominal fseq, and information on which functional heads are necessarily overtly realized in that language's definite noun phrases.

4 Conclusions and Speculation

In this paper, I proposed correlations between the fseqs of a language's noun phrases and the structures of their pronouns. Using modifiability of pronouns as a diagnostic for their syntactic category, I showed that in some languages, pronouns project to NP, and in others they project to DP, and that this is derived directly from the syntactic structure of noun phrases in the language. In addition, the Mayan languages demonstrate that pro-PP is available in languages where overt pronouns are used for purposes of contrast or emphasis, rather than argumentally. Furthermore, there are languages like Mandarin where pronouns are normally of category D, but can be coerced into being elements of category N in the presence of modification, and the possibility of such coercion also directly results from the language's nominal syntax.

In doing so, I provided arguments for a view of fseqs and extended projections wherein there is not one fixed fseq for all languages' noun phrases; instead, a given language will have a set of fseqs for noun phrases with different semantic properties (e.g., definiteness, specificity, number, etc.), and that these sets of fseqs will differ between languages. However, pronouns will have identical fseqs to non-pronominal noun phrases with similar semantic properties in that language. Thus, while the general template I follow for nominal fseqs is similar to that of Borer (2005), I allow for a much greater range of flexibility in terms of what heads will be present and absent, both within and between languages.

There are puzzles that remain to be solved, however, including the question of why the only observed pronominal categories are NP, DP and PP; here, I will only provide speculation. It is not surprising that languages will have pro-DP; after all, this is the canonical category for

argumental noun phrases, especially those that are definite, as is the case with the personal pronouns discussed in this paper. It is also not unexpected that NP will be a category for pronouns, as this is the site of substantive or lexical (as opposed to functional or inflectional) semantics in the noun phrase. Thus, one might expect the semantic capabilities of NP to be diverse; this is seen in clear cases of coercion from D to N, as discussed in Sections 3.2.2 and 3.3 and references therein, particularly Li (1999) and Cowper and Hall (2009). Finally, as P can be used to introduce adverbial elements, it is not surprising that pro-PPs may be used for non-argumental pronouns. On the other hand, the other functional categories in the nominal fseq, such as Div and #, have clear-cut semantic roles that divide out and provide quantification over the denotation introduced in NP. It is difficult to see how such a category could be extended to a pronominal denotation, since they have little to do with picking out referents in the discourse. Thus, it seems that the fact that pronouns are exclusively of the categories NP and DP results straightforwardly from the semantics associated with the different heads in the nominal extended projection.

References

- Abels, K. and A. Neeleman (2012). Linear asymmetries and the LCA. *Syntax* 15, 25–74.
- Abney, S. P. (1987). *The English Noun Phrase in its Sentential Aspect*. Ph. D. thesis, Massachusetts Institute of Technology.
- Barrie, M. and E. Mathieu (2012). Head movement and noun incorporation. *Linguistic Inquiry* 43(1), 133–142.
- Borer, H. (2005). *Structuring Sense, Vol. I: In Name Only*. Oxford: Oxford University Press.
- Bošković, Ž. (2008). What will you have, NP or DP? *Proceedings of NELS* 37, 101–114.
- Bošković, Ž. (2012). On NPs and clauses. Ms., University of Connecticut, Storrs.
- Butler, L. K. (2011). The DP-adjoined plural in Yucatec Maya and the syntax of plural marking. Ms., University of Arizona, UC Santa Cruz.
- Carlson, G. N. (1977). *Reference to Kinds in English*. Ph. D. thesis, MIT, Cambridge, MA.
- Cheng, L. L.-S. and R. Sybesma (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry* 30(4), 509–542.

- Chierchia, G. (1998). Reference to kinds across languages. *Natural Language Semantics* 6, 339–405.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagereka (Eds.), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, pp. 89–155. MIT Press.
- Chomsky, N. (2005). On phases. Ms., MIT.
- Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Corbett, G. (2000). *Number*. Cambridge, UK: Cambridge 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. (1999). What Algonquian morphology is really like: Hockett revisited. *MIT Occasional Papers in Linguistics* 17, 25–72.
- 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.
- Diesing, M. (1992). *Indefinites*. Cambridge, MA: MIT Press.
- Dubinsky, S. W. and R. Hamilton (1998). Epithets as antilogophoric pronouns. *Linguistic Inquiry* 29(4), 685–693.
- Fukui, N. (1988). Deriving the differences between English and Japanese: A case study in parametric syntax. *English Linguistics* 5, 249–270.
- Ghomeshi, J. (2003). Plural marking, indefiniteness, and the noun phrase. *Studia Linguistica* 57(2), 47–74.
- Grimshaw, J. B. (1991). Extended projection. ms., Brandeis University. Published in Grimshaw (2005).
- Grimshaw, J. B. (2005). *Words and Structure*. Stanford, CA: CSLI Publications.
- Halle, M. and A. Marantz (1993). Distributed Morphology and the pieces of inflection. In *The View from Building 20*, pp. 111–176. Cambridge, MA: MIT Press.
- Haugen, J. D. and D. Siddiqi (2013). Roots and the derivation. *Linguistic Inquiry* 44(3), 493–517.

- Huang, C.-T. J. and M. Ochi (2012). Remarks on classifiers and nominal structure in East Asian. To appear in a volume tentatively titled *Peaches and Plums*, by Academia Sinica, Taiwan.
- Iljic, R. (1994). Quantification in Mandarin Chinese: Two markers of plurality. *Linguistics* 32, 91–116.
- Kratzer, A. (2009). Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2), 187–237.
- Li, Y.-H. A. (1999). Plurality in a classifier language. *Journal of East Asian Linguistics* 8, 75–99.
- Longobardi, G. (1994). Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25(4), 609–665.
- Lyons, C. (1999). *Definiteness*. Cambridge, UK: Cambridge University Press.
- Mathieu, E. (2012a). Flavors of division. *Linguistic Inquiry* 43(4), 650–679.
- Mathieu, E. (2012b). Many a plural. Ms., University of Ottawa, to appear in *Weak Referentiality*, ed. A. Aguilar-Guevara, B. Le Bruyn, and J. Zwarts, John Benjamins.
- Matushansky, O. (2006). Head movement in linguistic theory. *Linguistic Inquiry* 37(1), 69–109.
- Melchin, P. B. (2013). Nominal projections and pronoun types. MA major research paper, University of Ottawa.
- Melchin, P. B. and R. Truswell (2013). Noun phrases and non-projecting heads. Paper presented at the Canadian Linguistics Association, 2 June 2013, University of Victoria.
- Nakanishi, K. and S. Tomioka (2004). Japanese plurals are exceptional. *Journal of East Asian Linguistics* 13, 113–140.
- Neeleman, A. and K. Szendrői (2007). Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38(4), 671–714.
- Nemoto, N. (2005). On mass denotations of bare nouns in Japanese and Korean. *Language* 43(2), 383–413.
- Noguchi, T. (1997). Two types of pronouns and variable binding. *Language* 73, 773–797.
- Partee, B. (1987). Noun phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh, and M. Stokhof (Eds.), *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, pp. 115–141. Dordrecht, The Netherlands: Foris Publications Holland.
- Postal, P. (1966). On so-called pronouns in English. In D. Reibel and S. Schane (Eds.), *Modern Studies in English*, pp. 201–223. Englewood Cliffs, NJ: Prentice-Hall.

- Richardson, K. (2001). What secondary predicates in Russian tell us about the link between tense, aspect, and case. *ZAS Papers in Linguistics* 26.
- Ritter, E. (1991). Two functional categories in noun phrases: evidence from Modern Hebrew. *Syntax and Semantics* 25, 37–62.
- Ritter, E. (1995). On the syntactic category of pronouns and agreement. *Natural Language & Linguistic Theory* 13(3), 335–397.
- Roberts, I. (2010). *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals*. Cambridge, MA: MIT Press.
- Runić, J. (2011). Clitic doubling in non-standard Serbian and Slovenian dialects. General paper, University of Connecticut.
- Saito, M., T.-H. J. Lin, and K. Murasugi (2008). N'-ellipsis and the structure of noun phrases in Chinese and Japanese. *Journal of East Asian Linguistics* 17, 247–271.
- Scott, G.-J. (2002). Stacked adjectival modification and the structure of nominal phrases. In G. Cinque (Ed.), *Functional Structure in DP and IP: The Cartography of Syntactic Structures*, pp. 91–120. Oxford: Oxford University Press.
- Shklovsky, K. (2012). *Tseltal Clause Structure*. Ph. D. thesis, MIT.
- Svenonius, P. (2012). Spanning. Ms., CASTL, University of Tromsø.
- Travis, L. (1984). *Parameters and Effects of Word-Order Variation*. Ph. D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
- Ueda, Y. and T. Haraguchi (2008). Plurality in Japanese and Chinese. *Nanzan Linguistics: Special Issue 3 2*, 229–242.
- Valentine, J. R. (2001). *Nishnaabemwin Reference Grammar*. Toronto, ON: University of Toronto Press.
- van Riemsdijk, H. (1998). Categorial feature magnetism: The endocentricity and distribution of projections. *The Journal of Comparative Germanic Linguistics* 2(1), 1–48.
- Watanabe, A. (2006). Functional projections of nominals in Japanese: Syntax of classifiers. *Natural Language and Linguistic Theory* 24, 241–306.
- Watanabe, A. (2008). The structure of DP. In S. Miyagawa (Ed.), *The Oxford Handbook of Japanese Linguistics*, pp. 513–540. Oxford: Oxford University Press.
- Watanabe, A. (2010). Notes on nominal ellipsis and the nature of *no* and classifiers in Japanese. *Journal of East Asian Linguistics* 19, 61–74.
- Wiltschko, M. (2008). The syntax of non-inflectional plural marking. *Natural Language and Linguistic Theory* 26, 639–694.

- 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.