UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
The Morphosyntax of Discontinuous Exponence

Permalink
https://escholarship.org/uc/item/78828731

Author
Campbell, Amy Melissa

Publication Date
2012

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

The Morphosyntax of Discontinuous Exponence

by

Amy Melissa Campbell

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Linguistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Line Mikkelsen, Chair
Professor Andrew Garrett
Professor Sharon Inkelas
Professor Johanna Nichols

Fall 2012

The Morphosyntax of Discontinuous Exponence

Copyright 2012
by
Amy Melissa Campbell

Abstract

The Morphosyntax of Discontinuous Exponence

by

Amy Melissa Campbell

Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Line Mikkelsen, Chair

This thesis offers a systematic treatment of discontinuous exponence, a pattern of inflection in which a single feature or a set of features bundled in syntax is expressed by multiple, distinct morphemes. This pattern is interesting and theoretically relevant because it represents a deviation from the expected one-to-one relationship between features and their morphological expressions. I consider cases of discontinuous exponence in verb agreement, TAM morphology, pronoun formation, and negation, showing the relationships among these various types and arguing that a unified analysis is in order.

The empirical foundation of the work is a typological survey of discontinuous exponence in the inflectional systems of 40 genetically and geographically diverse languages. This study establishes discontinuous exponence as a robust phenomenon, worthy of study in its own right, and brings to light new generalizations about the behavior of agreement features.

Working within the framework of Distributed Morphology I develop an analysis of discontinuous verb agreement that accounts for both the robustness and the noncanonicality of the phenomenon and extends naturally to other types of discontinuous exponence. My theory of Cyclic Insertion includes substantial revisions to Distributed Morphology; it rejects key assumptions such as the idea that feature insertion is feature discharge and it offers a view of vocabulary insertion that is compelled and constrained in very different ways than those assumed in the standard theory. Specifically, I assume that morphological insertion operates relative to meaning targets: insertion is motivated when it brings a form closer to its target meaning and is blocked if it cannot do so. The modifications I propose push Distributed Morphology in the direction of deriving discontinuous exponence more naturally. The noncanonicality of the phenomenon is explained with reference to greater complexity in its characteristic derivations.

I argue throughout the thesis for a view in which ϕ-features (agreement features) are bundled into sets. This view combines two independently motivated ideas – that

feature categories stand in hierarchical relations with one another and that categories themselves can be decomposed – to develop a rich, two-dimensional ϕ-set structure. Along one dimension are the fine-grained primitive features and entailments within feature categories, and on the other are hierarchical relations among the categories. These ϕ-sets have both descriptive and explanatory power; viewed as meaning targets they derive the patterns of discontinuous exponence, and within the system I propose they predict the phenomenon's cross-linguistic tendencies.

A thorough study of discontinuous exponence can illuminate much about the typology and theory of agreement. I will show that a commitment to accounting for the syntax and morphology of an agreement system – and the interface between the two modules – can lead to some very interesting insights about the necessary features of a good theory of agreement.

Contents

List of Figures **iv**

List of Tables **v**

1 Introduction **1**
1.1 Overview of Thesis . . . 2
1.2 Key Analytical Themes . . . 3
1.2.1 Complexity, Canonicality, and Frequency . . . 4
1.2.2 Meaning Targets . . . 5
1.2.3 Rich ϕ-Sets . . . 6

2 A Survey of Discontinuous Exponence **7**
2.1 Methodology . . . 8
2.1.1 Language sample . . . 8
2.1.2 Language features observed . . . 9
2.2 Patterns of Discontinuous Exponence . . . 13
2.2.1 Verb agreement . . . 15
2.2.2 Pronoun formatives . . . 19
2.2.3 TAM features . . . 23
2.2.4 Negation . . . 25
2.3 Parameters of Discontinuous Exponence . . . 26
2.3.1 Morphosyntactic type . . . 27
2.3.2 Morphological purity . . . 27
2.3.3 Morphological contiguity . . . 28
2.3.4 Referential ambiguity . . . 29
2.3.5 Summary: The typological space . . . 30

3 Noncanonicality of Discontinuous Agreement **33**
3.1 The Fusion Assumption . . . 34
3.2 Canonical Agreement . . . 41
3.2.1 Canonicality . . . 41

3.2.2 Agreement primitives . . . 42
3.2.3 Principles of canonical agreement . . . 44
3.2.4 Canonical morphology . . . 44
3.3 Discontinuous Agreement as Noncanonical Agreement . . . 45
3.4 Summary . . . 49

4 The Contribution of Discontinuous Exponence 51
4.1 Desiderata for a Theory of Discontinuous Agreement . . . 51
4.1.1 Discontinuous exponence as a single phenomenon . . . 52
4.1.2 Noncanonicality of discontinuous agreement . . . 53
4.1.3 Full expression . . . 53
4.1.4 Morpheme order in discontinuous agreement . . . 54
4.1.5 Ambiguity . . . 55
4.2 Challenges for Existing Models of Agreement . . . 56
4.2.1 Fundamentally syntactic theories . . . 56
4.2.2 Fundamentally morphological approaches . . . 61
4.3 In Favor of a Morphosyntactic Model of Discontinuous Agreement . . . 64

5 Deriving Discontinuity: Cyclic Insertion 67
5.1 Introduction . . . 67
5.2 The Structure of ϕ-Sets . . . 68
5.2.1 Relations among feature categories . . . 68
5.2.2 Structure within feature categories . . . 71
5.2.3 Two-dimensional ϕ-sets . . . 72
5.3 The Syntax of Agreement . . . 74
5.3.1 Cyclic Agree . . . 75
5.3.2 Meaning targets: ϕ-sets in syntax . . . 77
5.3.3 Summary: Output of syntax . . . 78
5.4 ϕ-Sets and the Morphology of Agreement . . . 78
5.4.1 Feature exponence . . . 81
5.4.2 Cyclic Insertion . . . 85
5.4.3 Blocking insertion . . . 99
5.5 Conclusion . . . 105

6 Applying and Extending the Theory 106
6.1 Case Study: Reanalyzing Karuk . . . 106
6.1.1 Previous analyses: Macaulay and Béjar . . . 107
6.1.2 Reanalysis: ϕ-sets and probe structure . . . 110
6.1.3 Cyclic insertion . . . 112
6.2 Consequences and Predictions of Cyclic Insertion . . . 115
6.2.1 Frequency of fused agreement morphs . . . 116

6.2.2 Frequency of subtypes of discontinuous and multiple exponence 117
6.2.3 Ordering tendencies 119
6.3 Related Phenomena 123
6.3.1 Discontinuous exponence in pronouns 123
6.3.2 Discontinuous exponence of TAM 125
6.4 Challenges for Future Work 126
6.4.1 Blocking in Cree 127
6.4.2 Problems with split probe in Georgian 128

7 Conclusions **134**

A Language Survey Results **145**

B Testing the syntactic bias of Cyclic Agree **151**

List of Figures

2.1 Morphosyntactic types of discontinuous exponence 27
2.2 Parameters of discontinuous exponence 31

4.1 Encoding values of person 60

5.1 Feature geometry for person (Harley & Ritter: 2002) 70
5.2 Encoding common values of person, number, and gender 73
5.3 Fission (Noyer, Halle, Embick & Noyer) 86
5.4 Fission (Arregi & Nevins) 86
5.5 Complex agreement node generated by Split 91

6.1 Parameters of discontinuous exponence (partial) 118

List of Tables

2.1 Languages surveyed 11
2.2 Languages surveyed, grouped by linguistic macro-area 11
2.3 Schematization of patterns 14
2.4 Basque inflection (Arregi 1999: 240) 22
2.5 Morphosyntactic types: Features, syntax, and morphology 28

3.1 German masculine nouns (Hock 1991: 211) 36
3.2 German present tense paradigm (Hock 1991: 212) 36

4.1 Order of person and number in discontinuous agreement (Trommer 2002: 89) 54
4.2 Order of person and number in discontinuous agreement (my survey) 54
4.3 Order of gender with respect to person and number in discontinuous agreement (my survey) 55
4.4 Comparing Standard Minimalism against the list of desiderata 58
4.5 Comparing Cyclic Agree against the list of desiderata 61
4.6 Comparing Distributed Optimality against the list of desiderata . . . 63
4.7 Comparing M-Case against the list of desiderata 64

5.1 Specifying person and number in two different language types 76

6.1 Karuk positive paradigm 107
6.2 Karuk optative paradigm 107
6.3 Karuk agreement morphology (Béjar 2003: 160) 109
6.4 Karuk positive and optative paradigms, first and second person . . . 110
6.5 Karuk positive and optative paradigms, first and third person 111
6.6 Karuk positive and optative paradigms, second and third person . . . 112
6.7 Karuk agreement morphology, reanalyzed 114
6.8 Number of survey patterns showing various types of discontinuity . . 119
6.9 Linear order of person and number in discontinuous agreement (my survey) 122
6.10 Nahuatl inflection 124

6.11 Basque inflection (Arregi 1999: 240) 124
6.12 Georgian transitive agreement patterns 129

A.1 Language survey results: Verb agreement 148
A.2 Language survey results: TAM 148
A.3 Language survey results: Negation 149
A.4 Language survey results: Pronouns 150

B.1 Survey languages checked for Béjar's syntactic bias 153

Acknowledgments

This thesis would not have been possible without the support of a great many friends, relatives, colleagues, and mentors. I must first thank my parents, who raised a daughter confident enough to start a graduate program and stubborn enough to finish it, and who continue to be a source of strength and inspiration to me.

In many ways, graduate school was not a natural step for me. Having been out of school and in the working world for several years, I first voiced my thoughts about pursuing graduate study in linguistics to Heather and Hans Kramer and Karen Quigg. I am grateful for their immediate encouragement, which was an important factor in my decision to enter this program.

My graduate student colleagues at UC Berkeley have been some of the most interesting and enjoyable people I've encountered, and working with them has been a pleasure. Their friendly faces in the hallways of Dwinelle and the camaraderie I experienced with them through study groups and the massive project of organizing the 2008 meeting of the Berkeley Linguistic Society made the first two years of my program so much more fun.

I owe a particular debt to my fieldwork colleagues Ramón Escamilla, Lindsey Newbold, and Justin Spence. I am deeply grateful to Victor Golla, our mentor in the study of the Hupa language, and to our wonderful friend and Hupa language consultant Verdena Parker. The time I spent working with these people on the Athabaskan language was challenging and rewarding in many ways. Through our work on text analysis I discovered the unique frustrations and satisfactions that come with a commitment to very closely analyze and explain the morphology of a language.

For financial support of my graduate studies I am extremely grateful to the National Science Foundation for a Graduate Research Fellowship, the UC Berkeley Graduate Division for summer research grants, and the Linguistic Society of America for a 2007 Summer Institute fellowship. I also thank the Hans Rausing Endangered Languages Project, the Endangered Language Fund, and the Survey of California and Other Indian Languages for fieldwork support.

As is the case with any major research project, the work presented here draws much from other researchers. I am particularly grateful to Susana Béjar and Greville Corbett for their comments on portions of this work and their willingness to answer questions about their own work. These conversations have been both interesting and productive. Portions of my research were also presented to UCB Syntax Circle, Stanford SMircle, TREND 2009, NELS 40, the Workshop on Morphological Complexity 2010, and BLS 38; I am grateful to audiences for their questions and comments on various aspects of my research.

The faculty and staff of the Linguistics department at UC Berkeley have been incredibly helpful and wonderful to work with. I thank Belén Flores and Paula Floro for administrative help with program requirements, funding applications, conference

organization, and in general too many things to list here. I am also grateful to Leanne Hinton and Andrew Garrett, respectively the past and present directors of the Survey of California and Other Indian Languages. Through my research appointments at the Survey I learned much about the difficult joy of linguistic fieldwork and I saw firsthand the positive impact it can have when done correctly.

I thank Andrew again, along with Sharon Inkelas and Johanna Nichols, for participating in my thesis committee. My conversations with each of you and our group meetings raised some extremely interesting questions and really helped to guide the evolution of this research. In particular I am deeply grateful to Line Mikkelsen for chairing this committee and for mentoring me as a student, teacher, and researcher over the past six years. Her clarity of thought, generosity with her time and her ideas, and rigorous attention to detail have showed me (and many others) what it means to be a scholar. I am extremely fortunate to have worked so closely with her, and she deserves a good deal of credit for the positive aspects of this work. I could not have asked for a better teacher or a finer advisor.

Finally, I dedicate this thesis to Caryl Shaw, who has been here through all of the twists and turns, and for whom my gratitude is beyond the measure of words.

Abbreviations

Code	Gloss
1	First person
2	Second person
3	Third person
AN	Animate gender
ASP	Aspect marker
ASSERT	Assertive
CAUS	Causative
CLS	Classifier
CUST	Customary aspect
DAT	Dative case
DIR	Direct
DPST	Definite past tense
DU	Dual number
EMPH	Emphatic
ERG	Ergative
EXCL	Exclusive
F	Feminine gender
FUT	Future tense
HAB	Habitual aspect
ICPL	Incompletive aspect
IN	Inanimate gender
INC	Inceptive
INCL	Inclusive
INCORP	Incorporated lexical element
INF	Infinitive
INV	Inverse
IPFV	Imperfective
IPST	Indefinite past tense
IRREAL	Irrealis mood
M	Masculine gender
MS	Marked scenario
N	Neuter gender
NEG	Negative
NPST	Nonpast
NSG	Nonsingular
O	Object
PASS	Passive
PART	Speech act participant
PFV	Perfective aspect
PL	Plural number
PPST	Proximate past tense
PROX	Proximate
PST	Past tense
PVB	Preverb
REAL	Realis mood
REP	Repetitive
RPST	Remote past tense
S	Subject
SG	Singular number
SPKR	Speaker
STV	Stative
THM	Thematic element
TRI	Trial number

Chapter 1

Introduction

Given the recent explosion of work on agreement within the frameworks of Minimalism and Distributed Morphology (Baker 2008; Béjar 2003; Béjar & Rezac 2009; Bobaljik 2008; Preminger 2011), the typology of agreement and certain noncanonical agreement patterns have come into theoretical focus. This dissertation is concerned with one such noncanonical pattern, discontinuous agreement, which involves a deviation from the expected one-to-one relation between (sets of) features and their morphological expressions. For instance, (1) is an example of discontinuous agreement in which multiple, coreferring agreement features – namely subject person and number – are encoded by distinct morphs. Another example of discontinuous agreement, shown in (2), involves a single feature category – here, number – splitting into more than one component value (nonsingular, which is used in both dual and plural forms in this language, and plural in the strict sense), each of which is realized by a separate morph but all of which are required to yield the intended meaning.

(1) zuek **z**-atoz-**te**
2PL **2**-come-**PL**
'You (≥2) come.'() Basque

(2) do**ː**-**yaː**-**di**-l-yo'
NEG-**PL.STRICT**-1**NSG**-CLS-love
'We (≥3) do not care for it.'() Hupa

Looking more broadly at the inflectional systems of languages with discontinuous agreement, the phenomenon can be considered as one instance of the more general phenomenon of discontinuous exponence. Example (3) shows the discontinuous marking of person and number in a Basque pronoun; compare with the similar pattern from verb agreement in (1). The pattern in (4) involves two distinct values of tense, proximate and indefinite past, marked by different affixes on a single verb; compare with the example of discontinuous agreement for number in Basque in (2). And (5)

illustrates the double-marking of mode. What these patterns have in common is that a single feature, or a set of features that can reasonably be assumed to be bundled on a single syntactic node, is split in the morphology. The chapters that follow will argue that the patterns in (1)–(2) and those in (3)–(5) can and should be given a unified analysis.

(3) **su-e**-k Boston-ea s-ixus-e-n
2-PL-PL.ABS Boston-ALL 2.ABS-go-PL-ABS-PST
'you (pl.) went to Boston' (Arregi 1999: 249) Basque

(4) su a**ː-yeː-yoː-v** vakht-as
3SG.M come-**PPST-IPST-RPST** time-DAT
'He came on time.' (Wali & Koul 1997: 225) Kashmiri

(5) tera i-**N**-p-**eNkani**
NEG.IRREAL 3.M.S-**IRREAL**-give-PASS.**IRREAL**
'He was not given (anything).' (Michael 2008: 276) Nanti

This dissertation makes three broad contributions to an understanding of inflectional morphology and, in particular, of verb agreement. First, it establishes discontinuous exponence as a robust phenomenon through a systematic study of 40 genetically and geographically diverse languages. Second, it offers new generalizations about the behavior of agreement features (person, number, and gender) in discontinuous agreement. Third, it provides a new analysis of discontinuous agreement that captures both the robustness and the noncanonicality of the phenomenon, and shows how the analysis extends naturally to other types of discontinuous exponence.

1.1 Overview of Thesis

I begin the investigation of discontinuous exponence in chapter 2, which presents a crosslinguistic survey of languages with discontinuous exponence. The 40 languages examined were selected to evenly represent the major language families and linguistic macro-areas of the world (§2.1.1). In selecting languages to include in the study I only considered those with discontinuous agreement for person and number. Most of the languages examined have additional types of discontinuous agreement and/or discontinuous exponence of TAM features.

Having established the robustness of discontinuous exponence in chapter 2, chapter 3 considers the extent to which discontinuous exponence is surprising. That is, to what extent does it deviate from the expected relationship between syntax and morphology? I consider the very pervasive assumption that certain inflectional features are most commonly fused on a single morph (§3.1) and argue that there are good reasons to accept this assumption. Then, working within the Canonicality Theory framework

of Corbett (2006), I show that discontinuous agreement is a noncanonical type of agreement, further motivating a treatment of discontinuous exponence more generally as a deviation from the norm.

Chapter 4 returns to the empirical patterns of discontinuous agreement and considers their implications for linguistic theory. I lay out several theoretical desiderata: a complete theory of agreement should capture both the robustness and the noncanonicality of the phenomenon and should account for certain crosslinguistic tendencies. The chapter concludes with an evaluation of several of the currently dominant theories of agreement against these desiderata.

In chapter 5 I propose a theory of agreement that takes Cyclic Agree (Béjar 2003; Béjar & Rezac 2009) and Distributed Morphology (Noyer 1992 and following work) as its starting point and makes substantial revisions to the latter. I will propose that the syntax creates meaning targets in the form of rich ϕ-sets, which are bundles of agreement features with a two-dimensional internal structure. The morphology aims to "hit" these meaning targets by fully encoding their features; this idea motivates a view of the vocabulary insertion operation as applying cyclically. The operation of morphological insertion relative to meaning targets also allows me to reject the standard assumption of Distributed Morphology that feature realization is feature discharge and present an alternative account of morphological blocking that depends on the principle of full meaning expression. The modifications and refinements I propose thus push Distributed Morphology in the direction of deriving discontinuous exponence much more naturally. The noncanonicality of the phenomenon is explained with reference to greater complexity in its characteristic derivations.

Finally, chapter 6 focuses on applying and extending my theory of Cyclic Insertion. I illustrate its utility by applying it to a reanalysis of agreement in Karuk. This new analysis retains the advantages of previous analyses of Karuk agreement (Macaulay 1992; Béjar 2003) while improving on both by better predicting the distribution of the inverse marker and explaining certain alternations in a principled way. It also argues strongly for the necessity of a notion of person hierarchy independent of probe structure (contra Béjar). Chapter 6 additionally lays out some of the predictions of and challenges to the theory, and shows how it can be extended to provide a unified account of related phenomena like pronoun formation and TAM morphology.

1.2 Key Analytical Themes

Three of the analytical notions mentioned above are particularly central to the development of the thesis, and it is worth giving them an expanded introduction here as they will be referred to in several places. The first is that there is an inverse relationship between the complexity of the derivation of a pattern and its canonicality or cross-linguistic frequency. The second is that morphology operates relative to

meaning targets, structured bundles of features that the morphology strives to fully express. And the third is that agreement features are bundled into sets located on syntactic nodes, and that these sets have a rich internal structure.

1.2.1 Complexity, Canonicality, and Frequency

The idea that increased derivational complexity corresponds to decreased canonicality and frequency of the generated form is foundational to the theory laid out in chapter 5. The notion of derivational complexity has been discussed in the literature at least since Chomsky's *Language and Mind* (1968), in which he noted a correlation between the number of syntactic transformations required to generate a sentence and the amount of work required to mentally process the sentence.

The question of the relationship between complexity and typological rarity has since been taken up by other researchers. In an investigation into the relative frequencies of VSO and SVO languages, Emonds attributed the rarity of the former to the fact that a particular type of rule is needed in its derivation, making VSO languages derivationally "more complicated and hence, rarer" (Emonds 1980: 44). A similar argument is advanced by Kayne in his work on the asymmetry of syntax; Kayne assumes that VSO order is derived from SVO order by leftward movement of V and that this straightforwardly explains its relative infrequency (Kayne 1994: 35–36).

Chapter 3 of this thesis discusses the pervasiveness of the Fusion Assumption (§3.1), the idea that person and number are most commonly fused on a single morph, and argued that there are many good reasons to accept it as fact. The logical extension of the Fusion Assumption is that the relationship between form and meaning is one-to-one in the default case. For verb agreement the expectation is thus that a single unit of meaning – a set of agreement features – will most commonly have a single morphological exponent. By the logic outlined above, it follows that the derivation of (rarer) discontinuous morphology should be more complex than that of (more common) fused morphology.

One way to think about morphological complexity is in terms of the number of lexical insertions required to generate a word: a morphologically complex word involves a larger number of lexical insertions and a simpler word a smaller number of insertions. This means that other things being equal, morphologically simple words express sets of features with fewer lexical items (fewer lexical insertions) and, by extension, that simpler forms are more likely to involve morphs that fuse more than one feature. This is a reasonable way to capture the inverse relation between complexity and frequency.

To generalize, I assume that the more operations a derivation involves, the more complex it is. Under current views of generative syntax and morphology, all morphological marking and syntactic structure result from operations (Merge and Adjoin

in Minimalist syntax, Vocabulary Insertion and Fission in Distributed Morphology). Thus the more morphological or syntactic operations a derivation involves, the more complex that derivation is and the less frequent the forms it produces should be across languages.

1.2.2 Meaning Targets

Chapter 5 will propose an analysis in which verb agreement is in a sense meaning-driven, in that the syntax creates meaning targets in the form of fully-featured terminal nodes, and the morphology attempts to come as close as possible to the target meaning at each cycle. This idea finds its genesis in the work of Caballero & Inkelas (to appear) who draw in turn on work by Kiparsky on blocking in inflectional paradigms. Kiparsky proposes a faithfulness constraint, EXPRESSIVENESS, which compels the output form to express all of the input meaning, other things being equal. Blocking effects in paradigms result from the tension between EXPRESSIVENESS and ECONOMY, a markedness constraint that favors simple over complex forms.

(6) Constraints active in blocking (Kiparsky 2005: 114)
   a. ECONOMY: Avoid complexity.
   b. EXPRESSIVENESS: Express meaning.

Also building on earlier work in Optimality Theory morphology (Anderson 1992; Noyer 1993; Baerman 2004), the main claim of Inkelas and Caballero is that the cyclic optimization of a word relative to its target meaning can predict occurrences of multiple exponence, just in case subsequent expressions of a feature are optimizing with respect to form or meaning. In other words, they argue that if an "inner" morphological cycle expresses a feature weakly, EXPRESSIVENESS can compel an "outer" cycle to express the same feature with a more informative morpheme.

I will return to the idea of meaning targets for agreement morphology in section 5.3.2, suggesting that they are implemented as complex agreement nodes that are valued by the agreement probes. Insertion operates relative to a meaning target; after the most informative vocabulary items is added, additional insertions are licensed if any features of the meaning target remain unexpressed. In my analysis the drive to fully express meaning thus compels the insertion of as many lexical items as are required to fully realize the features of each agreement node, giving rise to discontinuous agreement[1].

[1]In connection with the preceding discussion on derivational complexity and frequency, the non-canonicality of discontinuous exponence is captured in terms of the number of operations required to derive it.

1.2.3 Rich ϕ-Sets

It is common in linguistic analysis to specify agreement features as flat structures; for instance, a third person singular feminine argument might be specified simply as ϕ: [3, SG, F]. This type of flat structure is useful in that it captures the agreement patterns of many of the world's languages. However, I will show in chapter 5 that a more fully articulated structure will help us to account for certain noncanonical patterns of agreement.

Hierarchical structure among the agreement feature categories was suggested by Noyer (1992), who proposed a Universal Feature Hierarchy in which person features are superior to number features, which are superior to gender/class features. Noyer's hierarchy played a key role in the work of Harley (1994) and Harley & Ritter (2000, 2002), who translated his feature hierarchy effects into geometrical markedness relations, deriving patterns of dependency, contrastiveness, and markedness without needing any additional rules or filters, and defining natural classes of features for grammatical rules to target.

In her work on Context Sensitive Agreement (CSA), in which multiple arguments compete for a single morphological agreement slot, Béjar (2003) decomposed the feature categories of person and number into primitive features with entailment relations among them. She showed how a representation involving the fine structure of categories, in combination with certain assumption about the relevant syntactic operations, can predict both specific patterns of CSA and its general tendencies cross-linguistically.

Section 5.2.3 will combine these two independently motivated ideas – that feature categories stand in hierarchical relations with one another and that categories themselves can be decomposed – to develop a rich, two-dimensional ϕ-set structure. Along one dimension are the fine-grained primitive features and entailments within feature categories, and on the other are hierarchical relations among the categories. In later sections I will show how a ϕ-set with both dimensions, viewed as a meaning target, can both account for specific patterns of discontinuous agreement and explain some of its cross-linguistic tendencies.

Chapter 2

A Survey of Discontinuous Exponence

This chapter presents a crosslinguistic survey of languages with discontinuous exponence, emphasizing the robustness of the phenomenon across languages. Because the splitting of bundles of agreement features (taken to be person, number, and gender) is easily disgnosed, the discontinuous expression of coreferential agreement features will be considered as the central case. All of the languages considered show discontinuous agreement, though they vary in the paradigmatic pervasiveness of discontinuity. For instance, verbs in some of the surveyed languages mark person and number discontinuously across all cells of a paradigm, while in other languages discontinuities are limited to certain persons or aspectual paradigms. Discontinuous exponence involving other feature categories (namely tense, aspect, mood, and negation) will be discussed for some languages and these examples will be related to the patterns of discontinuous agreement.

Section 2.1 discusses the languages selected for examination and describes the design of the survey. Section 2.2 is concerned with the range of discontinuous patterns across languages. I will first discuss discontinuous agreement by pairs of features (discontinuity of person and number, for instance) and will then turn to discontinuous exponence of other inflectional features. Finally, in section 2.3 I take a step back, decomposing the phenomenon of discontinuous exponence into a number of discrete subtypes on the basis of a small number of parameters that follow directly from the results of the language survey.

Although discontinuous agreement patterns appear at first to show great variation across languages, I will draw out certain regularities as I proceed. Furthermore, the discontinuous expression of other inflectional features can be related to the core patterns of discontinuous agreement, as I will show in sections 2.2.2, 2.2.3, and 2.2.4. As such, this chapter establishes discontinuous exponence as a robust phenomenon, worthy of study in its own right. The generalizations captured here will form the

basis for later predictions about what discontinuous exponence tends to look like in languages that have it. If they are sound, these predictions can help us both mediate between existing theories of agreement and suggest new approaches to those problems which have not yet found satisfactory solutions.

2.1 Methodology

Before turning to the results of my study of discontinuous exponence, I discuss in section 2.1.1 the sample of languages included in the survey and report on the genetic and geographic coverage achieved. My general approach was to prioritize typological breadth over sheer number of languages in order to gain the greatest variety of discontinuous patterns. For instance, although most Athabaskan languages show discontinuous marking of person and number on the verb, I only included one of them in my sample because the patterns are largely the same across the family.

In section 2.1.2 I describe the study itself, exemplifying the kinds of data I looked for in each language source.

2.1.1 Language sample

Because discontinuous agreement is my primary object of inquiry, I selected only languages with discontinuous expression of agreement features for inclusion in the language survey[1]. Although I am aware of other types of discontinuous patterns, I did not systematically review a language's inflectional systems unless that language turned out to have discontinuous agreement. In some cases, however, a discontinuous pattern from an excluded language will be discussed as it relates to a surveyed language. For instance, Wambaya does not have discontinuous agreement but does have an interesting system of marking tense discontinuously, which will be discussed in connection with other languages in section 2.2.3.

Although little has been written about discontinuous exponence as a linguistic phenomenon, the work of three researchers provided a rich source of data and inspiration for the present study. Trommer (Trommer 2002) surveyed approximately 100 languages with subject person and number marked on the verb, and was particularly interested in the subset of languages for which person and number is marked discontinuously. Harbour (Harbour 2008) built on Trommer's work, positing an internal structure to sets of agreement features; this structure is used to derive certain facts about discontinuous patterns. And finally, Nichols (Nichols 1986) was based on 60

[1]It is worth mentioning that there are many, many languages with discontinuous agreement patterns that are not included in the present survey. I opted to keep the sample small enough to be able to look closely at each language included.

languages with rich morphology, also prerequisite for a study of discontinuous exponence. Many of the languages cited in these papers found their way into my study. I also considered additional languages known to me through my own work and my early research on the phenomenon in general.

To ensure good coverage of the genetic and areal groups of the world I consulted Nichols and Bickel's AUTOTYP genealogy and geography database, 2009 release (Nichols & Bickel 2009). For Nichols and Bickel, a linguistic *stock* is the highest-level demonstrable and reconstructible language family. Each isolate and unclassified language belongs to a separate stock. The language survey reported here includes 40 languages representing 38 different stocks. There are two Pama-Nyungan languages: Kalkatungu (major branch: Kalkatungic) and Warlpiri (major branch: Southwestern Pama-Nyungan) and two Uralic languages: Erzya Mordvin (major branch: Finno-Ugric) and Nenets (major branch: Samoyedic). I also considered linguistic *area*, simply the place where a language is spoken, in order to achieve geographic as well as genetic diversity. Table 2.1 lists each language surveyed along with its ISO639.3 code, genetic stock, and linguistic area.

Unlike linguistic area, the notion of linguistic *macro-area* is not a purely geographical one. Macro-areas are largely geographically defined, but also incorporate current knowledge and assumptions about language history and contact and genetic relationships among languages. As such, classification into macro-areas can be useful in discovering areal phenomena. Nichols and Bickel define 10 macro-areas of the world; Table 2.2 lists the languages in my survey grouped by linguistic macro-area, each of which is represented by four languages, yielding a sample that is geographically balanced with respect to macro-area.

2.1.2 Language features observed

There were two basic parts to the language survey. The first part collected general information from each language about four inflectional systems: verb agreement, pronominal formatives, TAM morphology, and negation[2]. The second part of the

Language	**ISO639.3**	**Stock**	**Area**
Aleut	ale	Eskimo-Aleut	N Coast Asia
Basque	eus	Basque	Europe
Berber (Tamazight)	tzm	Berber	N Africa
Cree (Plains)	crk	Algic	E North America
Dumi	dus	Sino-Tibetan	Indic
Fur	fvr	Fur	African Savannah
Georgian	kat	Kartvelian	Greater Mesopotamia
Halkomelem	hur	Salishan	Alaska-Oregon
Hebrew (Modern)	heb	Semitic	Greater Mesopotamia
Hixkaryana	hix	Cariban	NE South America
Huave	huv	Huave	Mesoamerica
Hupa	hup	Na-Dene	California
Iraqw	irk	Cushitic	S Africa
Juang	jun	Austroasiatic	Indic
Kalkatungu	ktg	Pama-Nyungan	S Australia
Karuk	kyh	Karuk	California
Kashmiri	kas	Indo-European	Indic
Ket	ket	Yeniseian	Inner Asia
Kiowa	kio	Kiowa-Tanoan	Basin and Plains
Kiwai (Island)	kiw	Kiwaian	S New Guinea
Lakhota	lkt	Siouan	Basin and Plains
Mapudungun	arn	Mapudungun	Andean
Maricopa	mrc	Yuman	Basin and Plains
Mayali	gup	Gunwingguan	N Australia
Mordvin (Erzya)	myv	Uralic	Inner Asia
Muna	mnb	Austronesian	Oceania
Nahuatl (Classical)	nci	Uto-Aztecan	Mesoamerica
Nanti	cox	Arawakan	NE South America
Nenets	yrk	Uralic	Inner Asia
Ngiyambaa	wyb	Pama-Nyungan	S Australia

Language	**ISO639.3**	**Stock**	**Area**
Nootka	noo	Wakashan	Alaska-Oregon
Popoluca (Texistepec)	poq	Mixe-Zoque	Mesoamerica
Quechua (Huallaga)	qub	Quechuan	Andean
Tamil	tam	Dravidian	Indic
Turkana	tuv	Nilotic	S Africa
Turkish	tur	Turkic	Greater Mesopotamia
Tzotzil	tzo	Mayan	Mesoamerica
Warlpiri	wbp	Pama-Nyungan	S Australia
Yimas	yee	Ramu-Lower Sepik	N Coast New Guinea

Table 2.1: Languages surveyed

Macro-Area	**Languages**
Africa	Berber, Fur, Iraqw, Turkana
Western and Southwestern Eurasia	Basque, Georgian, Hebrew, Turkish
Northern and Central Asia	Aleut, Ket, Mordvin, Nenets
South and Southeast Asia	Dumi, Juang, Kashmiri, Tamil
New Guinea and Oceania	Kiwai, Muna, Nimboran, Yimas
Australia	Kayardild, Mayali, Ngiyambaa, Warlpiri
Western North America	Halkomelem, Hupa, Karuk, Nootka
Eastern North America	Cree, Kiowa, Lakhota, Maricopa
Central America	Huave, Nahuatl, Popoluca, Tzotzil
South America	Hixkaryana, Mapudungun, Nanti, Quechua

Table 2.2: Languages surveyed, grouped by linguistic macro-area

survey collected detailed information specific to each discontinuous pattern.

General information

With respect to verb agreement, I determined whether agreement marking on the verb can be controlled by the subject, the direct object, and/or the indirect object (in some cases, this is contextually determined and I noted this as well). I additionally recorded which argument features are encoded on the verb: person (P), number (N), gender (G), or some combination thereof (for instance, P.N is used to schematize a marker that fuses person and number), and whether the language distinguishes between inclusive and exclusive in the first person. Finally, since the primary function of agreement is to track reference, I also recorded whether each language uses noun case to mark its grammatical relations.

Because pronominal formatives parallel agreement affixes in many langauges, I determined which agreement features are encoded on each language's independent (i.e., not bound to the verb or any other word) pronouns: P, N, G, P.N, P.G, N.G, and/or P.N.G. I marked an affirmative answer for each feature or combination of features expressed anywhere in the language, without distinguishing between different syntagmatic patterns. For instance, if a language expresses person and number discontinuously in the first and second person and fused in the third person, I marked "yes" for P, N, and P.N in the general language survey. However, this kind of contextual limitation was captured in the pattern-specific portions of the survey, which will be described below.

With respect to non-agreement features, I simply noted in the general language survey whether or not each feature can be marked on the verb or on other verbal elements, such as auxiliaries. These features include tense, aspect, mood, and negation.

[2]Although I did not systematically examine constituents other than the noun and the verb for discontinuous exponence, we do find it elsewhere. Examples (1) and (2) show adpositions inflecting to agree with subjects (Huave) and objects (Hupa); inflection of adpositions is also reported in Apalai and Tamazight Berber. Kim (2008: 223–226) notes that Huave numerals and quantity words also inflect to agree with the nouns they modify.

(1) xinan **s**-a-n-**an** ti joy
1.EX **1**-at-N-**PL** in hammock
'we (excl.) are in the hammock' (Kim 2008: 236) Huave

(2) **ya:-xo**-q'id
PL-3AO-on
'on them' Hupa

Pattern-specific information

Harbour defines discontinuous exponence as "agreement with a single argument by distinct parts of the verb" (Harbour 2008: 185). Following this definition, in the verbal agreement and pronominal systems I looked for examples where at least one agreement feature (person, number, or gender) is marked separately from one or more other overtly-marked, coreferential agreement features. Extending the notion of discontinuous exponence to the TAM and negation systems, I recorded two kinds of patterns: (i) those in which a single feature category is expressed twice in the morphology, with different values; for instance, see (32) and the immediately following examples; and (ii) patterns in which a feature (tense, aspect, mood, or negative) is fused with a root or another inflectional feature and is additionally expressed by a separate morpheme; see (37).

Each pattern was entered as a separate record. Since most languages have discontinuous exponence in more than one inflectional system or have multiple patterns possible within a system, the survey yielded a total of 115 patterns (69 verbal agreement patterns, 31 pronominal, 10 TAM, and 5 negation). For each verbal pattern, I noted the pattern's morphosyntactic type (§2.2) and the features participating in the discontinuity or discontinuities captured by the pattern. I also recorded whether the pattern gives rise to referential ambiguity (§2.3.4). For each pronominal pattern, I recorded the features involved and noted whether the pattern is similar to verbal agreement in the same language.

All patterns are included as Appendix A. The verb agreement patterns are in Table A.1, the TAM patterns in Table A.2, the negation patterns in Table A.3, and the pronoun patterns in Table A.4.

2.2 Patterns of Discontinuous Exponence

This section presents the results of the language survey described in section 2.1.2. It is organized by feature type and, for agreement features, by domain: section 2.2.1 discusses discontinuous verbal agreement, 2.2.2 deals with agreement feature discontinuities in personal pronouns, 2.2.3 looks at discontinuously expressed TAM features, and 2.2.4 discusses discontinuous negation.

Discontinuously expressed features are realized formally in many ways. Encoding strategies include multiple affixation, affix allomorphy, root allomorphy, and reduplication; some patterns involve more than one formal strategy. For instance, example (7) shows both multiple affixation (there are two affixes encoding first person) and affix allomorphy (the suffix expresses both person and number; one way to think of this is as a number suffix whose form is conditioned by person). Because there is no obvious correlation between feature category and means of formal realization, these

Features	Agreement[3]:	P = person, N = number, G = gender
	TAM:	T = tense, A = aspect, M = mood
	Negation:	Neg = negative
Domains		Aux = auxiliary
		V = verb
		Pr = pronominal stem

Table 2.3: Schematization of patterns

strategies will not be discussed systematically in the following sections.

(7) ch-**i**-tal-**otik**
ICPL-**1**-come-**1INC.PL**
'We (inclusive) are coming.' (Aissen 1987: 47) Tzotzil

In some cases it will be more convenient to refer to a schematic pattern than to an actual example. Examples will be schematized using the abbreviations given in Table 2.3. The verb form in (7), for example, is schematized as *P-V-P.N*.

In developing a definition of discontinuous exponence I will distinguish three basic subtypes of the phenomenon. *Coreferential exponence*, schematized in (8a), involves a set of features that can be expected to be bundled on a single node in the syntax (namely agreement features, person and number shown here) but that are expressed by distinct morphemes. *Combinatorial exponence* involves a single feature category in the syntax (person or tense, for instance) for which multiple, distinct values are expressed in the morphology. This is exemplified in (8b), in which the single tense category has two values realized in the morphology ($T_1 \neq T_2$). The third basic type of discontinuous exponence is simply *multiple exponence*, schematized for number in (8c).

(8) Morphosyntactic types of discontinuous exponence:
   a. Coreferential exponence
      e.g., *P-V-N*
   b. Combinatorial exponence
      e.g., *V-*T_1*-*T_2
   c. Multiple exponence
      e.g., *N-V.N*

[3]In case two different arguments are encoded on a verb, numeric subscripts will indicate which argument's features are being expressed by a particular morph. For instance, P_1-V-N_1-N_2 would represent a verb with a prefix encoding subject person, the first suffix encoding subject number, and the second suffix encoding object number.

I will begin defining the phenomenon with reference to the discontinuous expression of coreferential agreement features, the central example of discontinuous exponence. I will then extend this definition to other, related patterns. The final definition will include each of the basic morphosyntactic types of discontinuous exponence as shown in (9). These types will be discussed further in section 2.3.1 below.

(9) **Definition of discontinuous exponence**
A pattern of morphological exponence is discontinuous if one of the following situations obtains:

1. *Coreferential exponence*: A set of feature categories that can be expected to be bundled on a single node in the syntax (namely the agreement features of a single argument) are expressed by distinct morphemes
2. *Combinatorial exponence*: A feature category is realized by more than one morph, and each instantiation expresses a different value
3. *Multiple exponence*: A feature category is realized by more than one morph, and each instantiation expresses the same value

The defining trait of discontinuous exponence is that features expected to be co-located in the syntax are split in the morphology. The agreement features person, number, and gender are generally taken to be bundled on a single node in syntax[4]. Discontinuous agreement, in which coreferring agreement features are distributed across two or more morphs, is thus a good central example of discontinuous exponence and a good place to start building up a more complete definition of the phenomenon.

(10) Definition of discontinuous exponence (intermediate, to be expanded)
A pattern of morphological exponence is discontinuous if a set of feature categories that can be expected to be bundled on a single node in the syntax (namely the agreement features of a single argument) is expressed by distinct morphemes.

2.2.1 Verb agreement

By design, all of the languages surveyed show some kind of discontinuous verbal agreement. Because more than one pattern may emerge in a given language, the 40 languages give rise to 69 distinct patterns of discontinuous agreement. Strikingly, all of these patterns involve the discontinuous expression of person and number. Nine patterns (13% of all patterns) involve person-gender discontinuities, and 14 patterns (20% of all patterns) involve number-gender discontinuities. However, these facts

[4]The expectation that person, number, and gender are bundled together will be discussed at length in chapter 3.

should not be taken to suggest that gender is resistant to being expressed discontinuously; rather, my survey results indicate that gender is less likely to be expressed on the verb in general[5], which accounts for the lower number of discontinuous patterns involving gender.

Coreferential discontinuities

As discussed immediately above in connection with the definition in (10), coreferential discontinuities of agreement features represent a central, canonical type of discontinuous exponence[6]. This section examines coreferential discontinuous agreement, organized by pairs of agreement features.

P-N The most common type of discontinuous agreement involves the expression of person and number – as mentioned above, 69 patterns encode these features discontinuously. Example (11) shows the simplest (and also very common) form of person-number discontinuity: person is realized by the prefix and number by the suffix[7]. A related pattern in (12) shows person and number fused on the prefix and number additionally expressed by the suffix. In (13), the prefixes express person while number is fused with the verb root and is additionally expressed by a suffix.

(11) zuek **z**-atoz-**te**
2PL **2**-come-**PL**
'You (pl) come.' (Hualde & de Urbina 2003: 207) Basque

(12) **ì**-los-e-**tè** eèsi
2PL-go-ASP-**PL** 2PL
'You (pl) will go.' (Dimmendaal 1983: 122) Turkana

(13) **ku**-t-o-i-**kaŋ-qut-n**
2-DET-THM-PST-**2**-walk.**PL-PL**
'You (pl) walked.' (Georg 2007: 219) Ket

P-G All nine of the patterns involving person-gender discontinuities also show person-number discontinuities. The following examples show gender fused with number on a suffix in (14), and fused with the verb root in (15).

[5]Of the 40 language surveyed all 40 encode both person and number, but only 14 encode gender.

[6]Chapter 3 discusses at length what it means to be canonical and why discontinuous agreement is noncanonical. I will cite Corbett's work on Canonicality Theory, which defines a canonical pattern as one that closely matches the definition of a phenomenon (Corbett 2006: 9).

[7]As we will see in section 4.1.4, there is a strong crosslinguistic tendency for person marking to precede number marking in discontinuous agreement.

(14) **te**-kapl-**i**
2.FUT-fold-**F**.SG.FUT
'You (fem) will fold.' (Glinert 1989: 470) Hebrew

(15) **ni**-**pakamahw**-ekw-w
1-hit.**AN.O**-INV-3
'He hits me.' (Dahlstrom 1986: 29) Cree

N-G All 14 of the patterns involving number-gender discontinuities also show person-number discontinuities. Masculine gender is expressed discontinuously by a dedicated suffix in (16), in (17) animate gender is fused with the person prefix; in (18) the subject person prefix fuses masculine gender while the suffix encodes subject plurality.

(16) ŋuá-ke-sáram
bite-**DU.O**-**M.O**
'They (du, masc) will bite there.' (Anceaux 1965: 84) Nimboran

(17) widwaːd **ya**-**'**-a-ɬ-chwe'
acorn.flour **PL**-3.**AN**-CUST-CLS-make.PFV
'They would make acorn flour.' (Sapir & Golla 2001: 35) Hupa

(18) **du**-il-di-s-**n**
3**M.S**-PST-1.O-dress-**PL.S**
'They dressed me' (Georg 2007: 191) Ket

Combinatorial discontinuities

Combinatorial discontinuities in verb agreement are quite a bit less common than coreferential discontinuities; only two of the 69 discontinuous agreement patterns are combinatorial in nature. In both patterns the combinatorial feature is number, as shown below. In (19) the inner (second) prefix marks nonsingular number and the outer (first) prefix marks plural number. Without the outer prefix, this example would have a dual subject interpretation.

(19) doː-**yaː**-**di**-l-yo'
NEG-**PL**-1**NSG**-CLS-love
'We (pl) do not care for it.' (Sapir & Golla 2001: 303) Hupa

In (20), the final two prefixes both encode object number[8]. The first of the two, *ɛ́*-, encodes dual object number. The final prefix, *d*-, encodes nonsingular object number.

[8]On the surface, Kiowa agreement markers appear to fuse person and number. The example reflects the underlying form of the agreement prefixes, as analyzed by Watkins & McKenzie (1984: 115–127).

(20) nènhá**ː**dɔ́ < d-ià-**ę-d**-há**ː**dɔ́
1.S-SG.S-**DU.O-NSG.O**-shout.PFV
'I shouted to them (du).' (Watkins & McKenzie 1984: 113) Kiowa

Although there are no patterns involving combinatorial person or gender in my survey results, such examples are logically possible. For instance, one could imagine for person a language that marks first person on one morpheme and inclusive or exclusive on another. For gender, a language could mark animacy separately from masculine/feminine, for instance. It remains to be seen whether such examples are attested in language, but it would be somewhat surprising if they were not.

Combinatorial discontinuities differ from coreferential discontinuities in the following way: whereas coreferential discontinuities involve the splitting of a set of features into multiple morphemes, combinatorial discontinuities involve the splitting of one feature category into multiple values, each of which is expressed by a distinct morph. For instance, in (19) the category of number is split into the values nonsingular, marked by *di-*, and plural, marked by *yaː-*. The two types often go hand-in-hand in the domain of verb agreement; both (19) and (20) involve coreferential as well as combinatorial discontinuities. Let me then add to the definition of discontinuous exponence the combinatorial discontinuous expression of agreement features, giving the revised definition in (21).

(21) Definition of discontinuous exponence (intermediate, to be expanded)
A pattern of morphological exponence is discontinuous if one of the following situations obtains:

1. A set of feature categories that can be expected to be bundled on a single node in the syntax (namely the agreement features of a single argument) are expressed by distinct morphemes (coreferential exponence)
2. *An agreement feature category is realized by more than one morph, and each instantiation expresses a different value (combinatorial exponence)*

Multiple exponence discontinuities

Many of the patterns involving coreferential discontinuities also involve multiple exponence. For instance, example (12) above shows the coreferential discontinuous expression of person and number and also the multiple expression of number, since both affixes encode the same number value. It is also possible for agreement features to fuse with other categories, and thus the survey results contain patterns of verb agreement that show only multiple exponence, as in (22).

(22) **'**-yuu-**ksh**
1-see-**1**.PFV
'I saw it.' (Gordon 1986: 27) Maricopa

While patterns of multiple exponence do not fit the definition of discontinuous exponence so far, they are related. Recall that combinatorial exponence involves the splitting of a feature category into two or more distinct values, each of which is realized separately in the morphology. With multiple exponence a feature category is realized by multiple morphs *with the same feature value*. Multiple exponence can thus be seen as a simplification of combinatorial exponence in that the multiple morphological instantiations of a feature category all have the same value. I add to the definition of discontinuous exponence the multiple exponence of agreement features, giving the revised definition in (23).

(23) Definition of discontinuous exponence (intermediate, to be expanded)
A pattern of morphological exponence is discontinuous if one of the following situations obtains:

1. A set of feature categories that can be expected to be bundled on a single node in the syntax (namely the agreement features of a single argument) are expressed by distinct morphemes (coreferential exponence)
2. An agreement feature category is realized by more than one morph, and each instantiation expresses a different value (combinatorial exponence)
3. *An agreement feature category is realized by more than one morph, and each instantiation expresses the same value (multiple exponence)*

I have now distinguished three basic types of discontinuous exponence. Coreferential exponence is captured by point 1 of the definition in (23), combinatorial exponence by point 2 of the definition, and multiple exponence by point 3. Although the specifics are different, all three types of discontinuous exponence have in common that they represent a complication of the mapping from syntactic to morphological structure in which a single feature or set of features is multiply expressed in morphological form.

2.2.2 Pronoun formatives

The pronominal morphology of many languages contains discontinuous patterns, often similar to verb agreement patterns in the same language. 29 of the languages surveyed express agreement features discontinuously on their pronoun formatives; in 20 of these languages the pronominal patterns are similar or identical to the verb agreement patterns. Unlike verb agreement, all of the discontinuous pronominal features are of the coreferential type.

It is unsurprising that pronouns have a fundamentally different structure from verbs. I have analyzed most of the pronominal patterns as consisting of a pronominal stem fusing person plus affixes expressing other agreement features. However, I do

not consider these as examples of root allomorphy, since the pronominal stem has no meaning independent of the agreement features it bundles (compare with (13) and (15) as examples of verb root allomorphy). All examples of true root allomorphy occur in the verbal domain.

P-N As in verbal agreement, the most common type of pronominal discontinuity is person-number. 28 of the 31 total pronominal patterns involve person-number discontinuity. In the following examples, (24a) and (25a) show discontinuous person and number in the pronouns of Fur[9] and Nahuatl, respectively. Compare with (24b) and (25b), which show similar verb agreement patterns in the same languages. Among the languages I surveyed it is quite common for verb agreement and pronominal morphology to overlap in this manner.

(24) a. ìè-èN < **ìè-na**
3-PL
'they' (Jakobi 1990: 92) Fur

b. **k**-V-**na**
3.human-V-**PL**
'They V.' (Jakobi 1990: 92) Fur

(25) a. **te'**-hua-**ntin**
2-pronoun stem-**PL**
'we (pl)' (Sullivan 1988: 36) Nahuatl

b. o-**ti**-coch-**que'**
PFV-**2**-sleep-**PL**
'We (pl) slept.' (Sullivan 1988: 50) Nahuatl

P-G Six of the pronominal patterns involve discontinuously marked person and gender. Unlike in the domain of verbal agreement, some of these patterns do not additionally involve person-number discontinuities. In (26) there is both a person-number and a person-gender discontinuity, but in (27) there is only a person-gender discontinuity because singular number is not overtly marked on the pronoun.

(26) **yi**-m-**i**
3.PROX-PL-**F**
'they (fem proximate)' (Wali & Koul 1997: 196) Kashmiri

[9]The Fur forms are composed of a person-encoding prefix or pronominal stem and a number-encoding suffix. The plural suffix is limited to third person forms in both pronouns and verb agreement. For the form of the plural suffix and for an explanation of the formal dissimilarity between the third person marker in pronouns and verbs see Jakobi (1990: 92).

(27) **a-va**
3-F
'she' (Schiffman 1999: 59) Tamil

N-G There are just three pronominal patterns involving discontinuity of number and gender, all of which also mark person and number discontinuously. Example (28) involves a person-encoding pronominal stem with a prefix expressing gender and a suffix expressing person and number; (29) shows a pronominal stem expressing person with two suffixes: one expressing gender and the other expressing number.

(28) **θú-ɤáː-ləm**
F-3-**PL**
'they (fem)' (Galloway 1993: 192) Halkomelem

(29) a-**va-nga**
3-**F-PL**
'they (fem informal)' (Schiffman 1999: 59) Tamil

Relation to discontinuous verb agreement

While pronominal morphology and verb agreement both involve the exponence of agreement features, they differ in an obvious way: agreement features encoded on a pronoun are those of the pronoun itself, while agreement features encoded on a verb cross-reference one or more of the verb's arguments. Still, pronominal formation and verb agreement have much in common in terms of their surface patterns.

In many languages, pronominal morphology closely parallels verb agreement. 29 of the languages surveyed express agreement features discontinuously on their pronoun formatives; in 20 of these languages the pronominal patterns are similar or identical to verb agreement patterns in the same language. To be more specific, when I say that a pronominal pattern is similar to a pattern of verb agreement I mean that one or more of the criteria in (30) holds.

(30) Dimensions of similarity between pronominal morphology and verb agreement

1. Pronominal morphology encodes the same features as verbal morphology and the patterns of fusion are the same. For instance, person may be expressed by a dedicated morph and number by another morph in both systems, or one morph may fuse person and number while another encodes only gender in both systems.
2. The linear order of the features as expressed by pronoun formatives matches the order of features as expressed by agreement affixes on the verb.

	Independent Pronoun [10]	Agreement Affix(es)
1SG	ni	n-
1PL	gu	g-
2SG	su	s-
2PL	sue	s-...-e

Table 2.4: Basque inflection (Arregi 1999: 240)

3. Agreement features are realized with a similar phonological form in both the verbal and the pronominal domain.
4. The featural distinctions made in the pronominal domain tend to be subject to the same constraints as those made in the verbal domain. (See the immediately following discussion of Basque for an example.)

The similarity between pronouns and agreement markers is treated explicitly by Arregi (1999) for Basque. Table 2.4 shows the Basque pronominal and agreement forms, and (31) gives an example of both kinds of inflection in the same clause. Arregi argues for the same morphosyntactic analysis of both pronouns and verbal agreement affixes.

(31) **su**-**e**-k Boston-ea **s**-ixus-**e**-n
2-**PL**-PL.ABS Boston-ALL **2**.ABS-go-**PL**.ABS-PST
'You (pl) went to Boston.' (Arregi 1999: 249) Basque

Arregi analyzes *n-* as expressing 1sg, *g-* 1pl, *s-* 2, and *-e* pl in both pronouns and verb agreement. Assuming that *u* is a pronominal stem, the second person plural pronoun can be schematized as *P-Pr-N*, while a verb inflected to agree with a second person argument can be schematized as *P-V-N*. Basque then exhibits all of the characteristics in (30), including the fourth: there is fusion of person and number in the first person in both pronouns and verbs, and discontinuously expressed person and number in the second person.

In some languages, some analytical "dissection" is needed in order to draw parallels between verb agreement and pronominal formation. Still, to the extent that pronominal formatives can be analyzed as separate morphemes, the discontinuous patterns are often highly similar across the two inflectional systems. Furthermore, the same features are involved in both systems. As such, discontinuous pronominal morphology merits inclusion in a thorough study of discontinuous exponence.

[10] Basque lacks true third person pronouns; third person is marked only by verb agreement and demonstratives are optionally used for emphasis (Hualde & de Urbina 2003: 151).

2.2.3 TAM features

Discontinuous exponence of TAM features is significantly less pervasive than that of agreement features, but is found in 10 of the languages surveyed. In seven of these languages the discontinuity involves the tense category, two languages have discontinuously expressed aspect, and just one language has discontinuous mood. Patterns involve both combinatorial discontinuities and multiple exponence; each of these types is treated immediately below.

Combinatorial discontinuities

Four of the languages surveyed show combinatorial discontinuity of TAM features: Kashmiri, Kiwai, Nootka, and Warlpiri. These patterns involve multiple affixation (32), affix allomorphy (33), and affixation to an auxiliary in addition to the lexical verb (34).

In Kashmiri, the remote past tense form of an intransitive verb is built on the indefinite past tense form, which in turn is built on the proximate past tense form. Example (32) shows the stacking of three tense suffixes, all of which must be present to yield a remote past interpretation.

(32) su aː-**yeː-yoː-v** vakht-as
3SG.M come-**PPST-IPST-RPST** time-DAT
'He came on time.' (Wali & Koul 1997: 225) Kashmiri

Island Kiwai distinguishes 2 past, 1 present, and 3 future tenses. Tense is marked by a distinct form of the subject person agreement prefix, "together with combinations of prefixes, suffixes, and tense forms of the affixes which denote the number of the subject. In many verb forms, tense is signalled several times, often first in a general form... which is then followed by the indication of a specific past or future tense" (Wurm 1975: 338). This is shown in (33), in which general past tense is fused with the subject marker, which precedes the morpheme expressing definite past tense and habitual aspect.

(33) al-**g**-a-bi-**ru**-mo-uba-go-ow.al-wado-go
ASSERT-2/3.S.**PST**-HAB-TRI.S-**DPST**.HAB-NSG.S-bad-EMPH-do-REP-HAB
'They three were certainly repeatedly causing trouble as a habit.'
(Wurm 1975: 342) Kiwai

In Warlpiri, a verbal auxiliary expresses tense and agreement marking, and the lexical verb can also be marked for tense. Example (34) shows two different markings for tense: the verb is marked for nonpast tense, and the auxiliary is marked more specifically for present tense.

(34) kuyu **ka**-rlipa paka-rni-nja ya-**ni**
meat **PRES**-1INCL.PL kill-INF go-**NPST**
'We (inclusive) are going along killing game.' (Nash 1980: 44) Warlpiri

Wambaya[11], an Australian language unrelated to Warlpiri, also marks tense on both the verb and an auxiliary. Work by Nordlinger (1995) reports that the two elements may carry different specifications for tense; this is shown in (35), which is similar to (34) in that both involve a general tense specification (zero-marked nonpast) in addition to a more specific specification (here, future).

(35) bard-**ba** irri-∅
run-**FUT** 3PL.S-**NPST**
'They will run.' (Nordlinger 1995: 228) Wambaya

Furthermore, Nordlinger reports that imperative mood[12] is expressed by marking the verb with the future tense suffix and the auxiliary with the non-future suffix as in (36). This example differs from the previous in that here, two apparently contradictory tense specifications in combination yield a meaning not predictable from either value on its own. Tense in Wambaya is thus a good example of combinatorial exponence.

(36) jiyaj-**ba** girri-ng-**a** manganyma!
give-**FUT** 2PL.A-1.OBJ-**NFUT** food.ACC
'Give (pl) me some food!' (Nordlinger 1995: 229) Wambaya

Multiple exponence discontinuities

In the other six languages with TAM discontinuities (Dumi, Hupa, Ket, Mayali, Nahuatl, and Nanti), the patterns are of the multiple exponence type. Here verb root allomorphy (37) and affix allomorpy (38) are used as strategies for feature realization.

(37) da-sitej/q-i-**(t)in-a**
3F.S-INCORP-3F.O-**PST-BECOME.PST**
'She woke her up.' (Georg 2007: 221) Ket

(38) tera i-**N**-p-**eNkani**
NEG.IRREAL 3.M.S-**IRREAL**-give-PASS.**IRREAL**
'He was not given (anything).' (Michael 2008: 276) Nanti

[11]Wambaya is not included in my survey, but is discussed here in connection with example (34).
[12]Unlike imperative mood, irrealis and hypothetical mood in Wambaya are overtly marked on the auxiliary.

Relation to canonical discontinuous exponence

Although they do not show the canonical, coreferential type of discontinuous pattern, discontinuous TAM patterns have a clear relationship to the verbal agreement patterns presented in the coreferential and multiple exponence portions of section 2.2.1. Coreferential agreement and TAM morphology both involve a single feature category for which multiple values are possible, with different values marked by different morphs. Compare example (19), in which the number category is realized by two distinct values, with (33), in which the tense category is realized by two distinct values. Discontinuous patterns of the multiple exponence type are also found in both the verb agreement and TAM systems.

The definition of discontinuous exponence already includes combinatorial and multiple exponence in the agreement system. All that is needed to account for discontinuous TAM patterns, then, is to expand the second and third points of the definition in (23) (referring to combinatorial exponence and multiple exponence, respectively) from agreement features to inflectional features more generally. This gives the final definition that was initially presented in (9), repeated here as (39).

(39) Definition of discontinuous exponence
A pattern of morphological exponence is discontinuous if one of the following situations obtains:

1. Coreferential exponence: A set of feature categories that can be expected to be bundled on a single node in the syntax (namely the agreement features of a single argument) are expressed by distinct morphemes
2. Combinatorial exponence: A feature category is realized by more than one morph, and each instantiation expresses a different value
3. Multiple exponence: A feature category is realized by more than one morph, and each instantiation expresses the same value

2.2.4 Negation

Five of the languages surveyed show discontinuous expression of negation: Berber, Dumi, Karuk, Maricopa, and Turkish. This category may be under-reported, however, because not all grammars document negation fully (or at all), so it may be the case that affixes that express negation secondarily may not be represented here.

The formal realizations of discontinuous negation patterns include multiple affixation (40), allomorphy of an agreement (41) or non-agreement (42) affix, and root allomorphy (43).

(40) **waly**-'-tpuy-**ma**-k
NEG-1-kill-**NEG**-REAL
'I didn't kill him.' (Gordon 1986: 72) Maricopa

(41) **pu**-'áhoː-**p**
NEG-walk-PL.**NEG**
'They don't walk.' (Bright 1957: 67) Karuk

(42) aŋ-a tom khɔlɔ ŋɔ kir-ni **mə**-tsaːpt-u-**nəʔ**
1-ERG that all EMPH carry-INF **NEG**.PST-be.able-1SG.S.3.O-**NEG**
'I wasn't able to carry all of that.' (van Driem 1993: 124) Dumi

(43) **ur**-i-**bdi**
NEG-3SG.M-begin.**NEG**
'He did not begin.' (Abdel-Massih 1971: 173) Berber

Relation to canonical discontinuous exponence

Because negation involves a single feature rather than a set of features, discontinuous patterns of negation will never be of the coreferential type. Similarly, since it is not possible to have multiple, distinct but logically possible values for this feature category[13], discontinuous patterns of negation will never involve coreferential exponence. However, discontinuous negation patterns of the multiple exponence type can be found. Compare (22), in which a single value for person is expressed by two different morphs, and (38), in which a single mood value is multiply expressed, with (40–43), in which a single negation value (namely, negative), is expressed by multiple morphs.

Note that it is not necessary to further expand the definition to accommodate discontinuous negation, since the third point of (39), which deals with multiple exponence, is already generalized with respect to the type of feature.

2.3 Parameters of Discontinuous Exponence

One of the main findings of the typological study is that discontinuous exponence is a very diverse phenomenon. Within the categories defined by the three morphosyntactic types and the four domains surveyed, there is still some variation in how patterns are formally expressed. This diversity means that a systematic study of discontinuous exponence is a significant undertaking. In this section I aim to further decompose the problem by dividing discontinuous exponence into subtypes, which will allow me to more precisely characterize each pattern.

[13]It is difficult to imagine that pure negation, i.e. negation not encoding a secondary function like modality or evidentiality, could be anything other than a privative or binary feature.

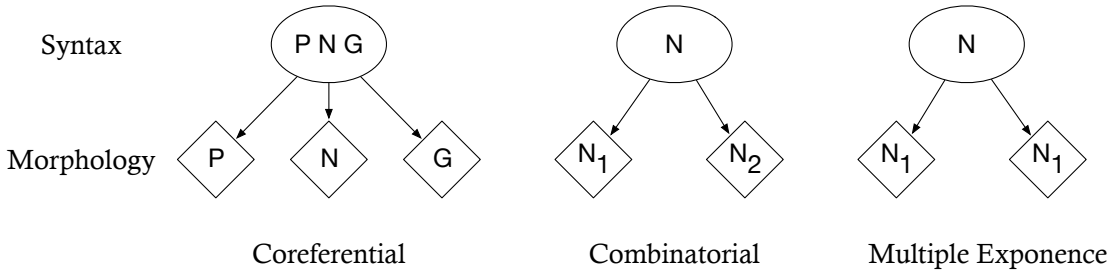


Figure 2.1: Morphosyntactic types of discontinuous exponence

2.3.1 Morphosyntactic type

Throughout section 2.2, three basic types of discontinuous exponence were discussed: coreferential, combinatorial, and multiple exponence. These types are morphosyntactic in that they describe the relationship between syntactic features and their morphological realizations. What they all have in common is a deviation from the expected one-to-one relationship between features and morphological expressions; recall the definition of canonical (coreferential) discontinuous exponence as the realization of one or more features from a single syntactic head by separate morphemes. I discussed this as the splitting of a set of features.

Combinatorial exponence involves splitting of a different nature; one feature category is split into more than one distinct component value, each of which is realized by a separate morpheme. The multiple feature values combine to yield the complete (possibly complex) value of the category. Finally, in a multiple exponence pattern a feature category is realized by multiple morphs, but each morph realizes the same value. Figure 2.1 diagrams the syntax-morphology mapping for each morphosyntactic type of discontinuous exponence.

Table 2.5 shows the participating feature categories and the characteristic syntax and morphology of each type of discontinuous exponence. Because agreement features are the only ones assumed to be bundled in the syntax, they are the only categories to participate in coreferential exponence. Combinatorial exponence involves features whose values can combine semantically, agreement and TAM features. And multiple exponence can involve any inflectional category, including privative or binary features.

2.3.2 Morphological purity

The notion of purity, introduced by Harbour (2008), distinguishes patterns like *P-V-N* (44) from those like *V-P-P.N* (45). Both examples involve a person-number discontinuity: (44) expresses person with the prefix and number discontinuously via

DE Type	Features	Syntax	Morphology
Coreferential	PNG	Feature set	Multiple feature categories distributed across morphemes
Combinatorial	PNG, T(AM)	Feature category	Multiple feature values distributed across morphemes
ME	PNG, TAM, Neg	Feature category	One feature value expressed by multiple morphemes

Table 2.5: Morphosyntactic types: Features, syntax, and morphology

the suffix, and (45) expresses person with the first suffix and number via the second suffix. However, in (45) both agreement suffixes encode person, and so only (44) represents a pure discontinuity.

(44) da=**v**-cer-**t**
PVB-**1**-write-**PL**
'We will write it.' (Aronson 1990: 44) Georgian

(45) a lo-**t**-**á'**
move-**2**.III-**2PL**.III
'You (pl) move.' (Mous 1993: 161) Iraqw

A pure discontinuity can be thought of as resulting from a cleanly split feature set. If a set of agreement features in the syntax splits and each feature maps to at most one morpheme, the discontinuity is pure. In (44), person and number each map to one morph. In (45), though, person maps to two morphs and so the discontinuity is impure. Note that multiple features may map to one morph without creating impurity, as shown in (46) in which number and gender are fused on one suffix. In other words, pure coreference may involve a many-to-one mapping from syntax to morphology, but not a one-to-many mapping.

(46) **te**-kapl-**i**
2.FUT-fold-**F.SG**.FUT
'You (fem) will fold.' (Glinert 1989: 470) Hebrew

2.3.3 Morphological contiguity

Contiguity refers simply to whether the morphemes that represent features discontinuously are linearly adjacent (contiguous) or not (non-contiguous). Example (16),

repeated here as (47), shows the contiguous discontinuous exponence of person and number/gender. Example (11), repeated here as (48), shows the non-contiguous discontinuous exponence of person and number.

(47) widwa:d **ya-'**-a-ɬ-chwe'
acorn.flour **PL-3.AN**-CUST-CLS-make.PFV
'They would make acorn flour.' (Sapir & Golla 2001: 35) Hupa

(48) zuek **z**-atoz-**te**
2PL **2**-come-**PL**
'You (pl) come.' (Hualde & de Urbina 2003: 207) Basque

This distinction cross-cuts all other parameters discussed here. Either a pure or impure discontinuity of any of the morphosyntactic types may be realized in a contiguous or non-contiguous fashion. This parameter refers only to surface form, so it is strictly morphological.

2.3.4 Referential ambiguity

A final parameter in the characterization of a pattern of discontinuous exponence is whether it can give rise to ambiguous interpretations[14]. Ambiguity sometimes arises when an agreement feature splits away from the other features in its set, as demonstrated by the various interpretations impossible for the following examples.

(49) a:= **ya:-xo-**ɬ**-ch'i**-de:-ne'
thus= **PL-3.A.O**-COM-**3.A.S**-THM-say
'he said to them' / 'they said to him' / 'they said to them'
(Sapir & Golla 2001) Hupa

(50) mo-g-k'al-i-t < mo-**v-g**-k'al-i-**t**
PVB-**1.S-2.O**-kill-THM-**PL**
'I kill you (pl.)' / 'we kill you (sg.)' / 'we kill you (pl.)'
(Hewitt 1995) Georgian

(51) **i**-kamoso-**hig**-ak-e-**ri**
3M.S-visit-**PL**-PFV-REALIS-**3M.O**
'he visited them' / 'they visited him' / 'they visited them'
(Michael 2008) Nanti

The only examples I have found of such ambiguity have the following properties. First, they involve the expression of number discontinuously from the person (and

[14]What I am calling referential ambiguity was described as "promiscuous number marking" by Leer (1991) for several indigenous languages of North America.

sometimes gender) features of the corresponding argument. Second, they code more than one argument on the verb. In these patterns, the separately marked number feature can be interpreted to mark the number of one argument or the other, or sometimes both.

However, while these properties might be necessary for ambiguity to arise, they are not sufficient. Many properties with the same general form do not have various possible interpretations, and there are two strategies that I have observed languages to employ for preventing ambiguity. One such strategy is to fuse case with the number morpheme (as in Basque; see for instance (31) in which the second person prefix fuses absolutive case), thereby explicitly indicating which argument's number is being encoded. In other languages, number affixes appear in a morphological position dedicated to the expression of the number of just one argument, as in Maricopa which has a prefix that encodes object number only.

2.3.5 Summary: The typological space

The parameters discussed in this section are illustrated in Figure 2.2, where each node represents one type of discontinuous exponence. Each leaf node can be further separated according to exponential contiguity (§2.3.3). A schematized example is given for each terminal node.

This tree shows that parametric splits are only possible for certain types of patterns. The distinction between pure and impure discontinuities refers to whether a set of agreement features splits cleanly or not (§2.3.2). Because it refers to feature sets, it is limited to the coreferential exponence type. The distinction between ambiguous and unambiguous reference is limited as well, to purely expressed coreferential patterns. The limitation to coreferential patterns follows from the fact that non-agreement features cannot be coreferential in the same way as agreement features, so their discontinuous marking cannot be construed as a splitting of feature sets, which is what gives rise to ambiguity. The limitation to pure coreferential discontinuities follows logically because impurely discontinuously expressed agreement features by definition appear elsewhere in the domain, fused with other agreement features; these other features provide the context necessary to resolve potential ambiguity.

The next chapter will consider more deeply the notion of canonicality in relation to discontinuous exponence. I will argue that discontinuous agreement is a noncanonical form of agreement, but it is still possible to consider the relative canonicality of different subtypes of discontinuous agreement. More generally, when discontinuous exponence is observed it makes sense to ask how canonical a particular pattern is in the context of the phenomenon as a whole. Throughout this chapter I have referred to coreferential discontinuities as the most canonical kind of discontinuous pattern; other types of discontinuous exponence (coreferential and multiple exponence) were discussed as they relate to coreferential patterns. Before ending this chapter I would

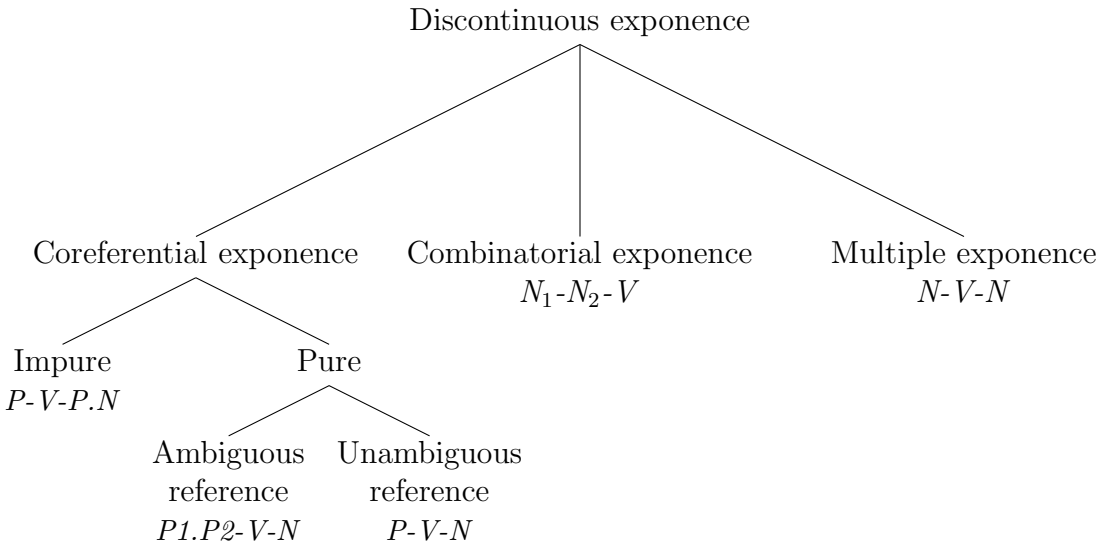


Figure 2.2: Parameters of discontinuous exponence

like to be more explicit about why coreferential discontinuities are the most canonical. What it means to be 'canonical' will be discussed in section 3.2.1; that section presents a working definition of canonical agreement, but for now it will suffice to say that a canonical example is one that most closely matches the definition of a phenomenon.

At a high level, discontinuous exponence involves a one-to-many relationship between syntactic and morphological elements. I discussed the splitting of atomic sets of agreement features in the syntax into multiple morphs encoding subsets of these features (coreferential exponence) and the splitting of a single feature category into multiple morphs encoding different values (combinatorial exponence) or the same value (multiple exponence). These are schematized in Figure 2.1. I suggest that the one-to-many relation in coreferential agreement is the most canonical of the three, because the split is so easily identifiable[15]: there is good reason to assume that agreement features are bundled in the syntax and it is easy to see that coreferential agreement features are expressed by separate morphemes in examples like (11)–(18).

Figure 2.2 shows that within coreferential discontinuous exponence there are two subtype distinctions. Considering first the pure/impure distinction, the principles of identifiability and simplicity point to pure discontinuities as the most canonical. Pure patterns like P-V-N, as in example (44), involve a cleanly split set of agreement features while impure patterns like P-V-PN in (45) have a more complicated morphosyntactic mapping in which the person feature maps to both morphs.

The final parameter to consider is the distinction between ambiguous and unam-

[15]Identifiability is a basic property of canonical agreement, and will be discussed in the next chapter.

biguous reference. Agreement involving a single controller is more canonical than agreement involving more than one controller (Corbett 2006: 18), and ambiguous discontinuities necessarily involve multiple controllers, so unambiguous discontinuities are more likely to be canonical. In other words, discontinuous agreement involving a single controller (more canonical) is necessarily unambiguous. Furthermore, the principles of simplicity and identifiability dictate that forms in which the reference is unambiguously resolvable are more canonical.

Therefore, by the principles of identifiability and simplicity, and following the work of Corbett (2006), I claim that the most canonical example of discontinuous exponence is a pattern of the coreferential type, characterized by a pure discontinuity of agreement features that does not yield a referentially ambiguous meaning. However, any theory of discontinuous exponence, and thus any complete theory of inflectional morphology, must account for *all* of the subtypes of discontinuous exponence, including the contextual restrictions on certain distinctions.

Chapter 3

Noncanonicality of Discontinuous Agreement

Chapter 2 presented discontinuous exponence as a robust phenomenon with internal complexity that is pervasive across languages and feature types. I said that at its core, discontinuous exponence represents a deviation from the expected relationship between syntax and morphology in that it involves a one-to-many relation between (sets of) morphosyntactic features on the one hand and morphs on the other. However, in order for this to be meaningful we must have a clearer picture of what the expectations are. That is, I must contend with an important question: How surprising is discontinuous exponence? I must say more about when features are expected to be fused on a single morph, and whether they are ever expected to be realized discontinuously.

Ideally, an investigation of these issues would draw on the results of a crosslinguistic study of the frequency of discontinuous exponence in the inflectional morphology of a large number of languages; however, such a study has not yet been done. I discuss in section 3.1 the very pervasive assumption that certain inflectional features are most commonly fused on a single morph. Despite its apparently widespread acceptance in the theoretical and typological literature, this idea has not yet received adequate empirical support. Gathering the data necessary for such a study will be an important part of future work on discontinuous exponence, but is unfortunately far beyond the scope of the present work.

The final sections of this chapter thus outline another possible approach to positioning discontinuous exponence as an unexpected morphosyntactic pattern. I will concentrate on discontinuous agreement or, more specifically, the coreferential discontinuous expression of agreement features (section 2.2.1). This will allow us to limit the scope of the discussion and draw on the existing literature in a more focused way. The discussion here is not intended to be conclusive, but rather exploratory and suggestive.

In section 3.2 I summarize the work of Corbett (2006) on Canonicality Theory, a potentially useful framework for considering the extent to which certain morphological forms and morphosyntactic patterns match the expectations[1]. Corbett takes a broad look at agreement and related phenomena in many languages, aiming to distinguish canonical from non-canonical agreement. I will to some extent attempt to synthesize his claims, drawing out certain general principles that underlie his work.

Section 3.3 examines discontinuous agreement from a canonicality perspective. Although Corbett (2006) does not address discontinuous agreement directly, I will show that his observations about noncanonical agreement phenomena can be brought to bear on the matter, particularly when one takes into account the general principles that motivate his main claims. In this respect the present work makes a contribution to Canonicality Theory.

Finally, section 3.4 will conclude that canonicality theory supports the fusion assumption and that it is reasonable to consider discontinuous agreement as a form of noncanonical agreement, motivating a treatment of discontinuous exponence more generally as a deviation from the expected relationship between syntax and morphology.

3.1 The Fusion Assumption

As was just noted, there is an extremely pervasive assumption that agreement features are by default expressed cumulatively[2]. This assumption extends beyond the theoretical literature, cropping up in historical and typologically-oriented research and descriptive work as well. In this section I review representative examples of each type and argue that the fusion assumption, while apparently well-accepted, has not yet been empirically supported.

Historical work

Before examining the fusion assumption in current linguistic theory, I briefly discuss three key ideas from the literature on historical linguistics that bear on the matter: (i) the early structuralist view of morphemes as Saussurean signs; (ii) the question of whether language change favors simple or complex forms; and (iii) the diachronic source of agreement markers.

Early work in the American structuralist tradition, in particular that of Bloomfield (1984, reprinted from his 1933 work), held that words can be decomposed into

[1] Although I focus on Corbett's work on agreement, he has also applied Canonicality Theory to other domains such as suppletion (Corbett 2007) and derivational morphology (Corbett 2010).

[2] In this chapter I use the terms "cumulative expression" and "fused" or "fusional morphology" to refer to the situation where more than one agreement feature is expressed by a single morph. Cumulative and fusional exponence thus stand in contrast to discontinuous exponence.

constituent parts, the meanings of which are subparts of the meanings of the words in which they occur. For Bloomfield, a morpheme is "a linguistic form which bears no partial phonetic-semantic resemblance to any other form" (Bloomfield 1984: 161); in other words, a morpheme is the smallest analyzable unit of form and its concomitant meaning[3]. Although this idea was later challenged by Aronoff (1976), Anderson (1992), and Noyer (1992), among others, it is still extremely influential to the way morphological analysis is done.

The relevance this notion to a discussion of discontinuous agreement can be seen in light of another pervasive idea, discussed in more detail in the context of theoretical literature below, namely that the agreement features person, number, and gender bundle together semantically to form a single, atomic meaning. Taking seriously both ideas (that morphemes are minimal pairings of form and meaning and that agreement features are semantically bundled), the meaning of an agreement morph should be a fully-valued set of agreement features. Therefore, agreement morphology should take the form of a single morph encoding all of the overtly-expressed features of a single argument. In other words, discontinuous agreement is unexpected[4].

With regard to the second key idea listed above, the claim that language change favors complex forms was articulated by Kuryłowicz as his first 'law' of analogy[5]:

(52) Kuryłowicz' first 'law' (Hock 1991: 211)
A bipartite marker tends to replace an isofunctional morpheme consisting of only one of these elements, i.e. a complex marker replaces a simple marker.

Table 3.1 demonstrates this tendency for German masculine nouns. The expected reflexes of the plural form *boum-a* in Table 3.1 row (a) is the unattested **Baum-e*, where plural is marked solely by the ending -e, in row (b). Instead, the correct form is *Bäum-e* as in row (c), where the plural is marked by both the plural ending and umlaut in the root. Here the more complex, bipartite pattern is generalized at the expense of the simple pattern.

Although many such examples can be found, Hock notes that examples showing the reverse direction of change, favoring the generalization of simple markers, are also common (Hock 1991: 211–212). In Table 3.2 row (a), the first person singular form *gib-u* is shown to be double-marked for subject agreement in OHG by a vowel

[3]This is often referred to in the literature as a "Saussurean sign", which Anderson points out is a "particularly limited view of the sign relation, as compared with that maintained by de Saussure himself" Anderson (1992: 49, and reference therein).

[4]As an aside, it seems that if we take these ideas very seriously then portmanteau agreement, in which a single morph encodes the features of more than one argument, is also unexpected.

[5]Here I follow Hock's convention of enclosing the word 'law' in single quotes when discussing Kuryłowicz. Hock describes these 'laws' as being in fact introspectively-based generalizations resulting from Kuryłowicz' extensive work on analogical change, rather than categorical or statistical facts supported by typological study (Hock 1991: 210).

a. OHG	gast	:	gest-i	'guest(s)'
	boum	:	boum-a	'tree(s)'
b. NHG (expected)	Gast	:	Gäst-e	
	Baum	:	***Baum-e**	
c. NHG (attested)	Gast	:	Gäst-e	
	Baum	:	**Bäum-e**	

Table 3.1: German masculine nouns (Hock 1991: 211)

			'go, drive'		'give'
a. OHG	sg.	1	far-**u**	:	g**i**b-**u**
		2	fer-ist	:	gib-ist
		3	fer-it	:	gib-it
	pl.	1	far-am	:	geb-am
		2	far-at	:	geb-at
		3	far-ant	:	geb-ant
b. NHG	sg.	1	fahr-**e**	:	geb-**e**
		2	fähr-st	:	gib-st
		3	fähr-t	:	gib-t
	pl.	1	fahr-en	:	geb-en
		2	fahr-t	:	geb-t
		3	fahr-en	:	geb-en

Table 3.2: German present tense paradigm (Hock 1991: 212)

alternation in the root (see the plural forms for comparison) and the *-u* ending, while in row (b) a simpler pattern has been generalized by analogy with words like 'go, drive', so that the NHG first person singular form *geb-e* is marked only by the *-e* ending[6].

Working on a similar topic but with a very different approach, Mańczak proposed several hypotheses about the general direction of analogical change. Hock takes one of his proposals, the fourth, to be particularly relevant to Kuryłowicz' first 'law':

(53) Mańczak's fourth tendency (Hock 1991: 231)
∅-endings are more frequently replaced by full ones than vice-versa.

Hock notes that the contribution of Mańczak's observation is that it "demonstrates that the tendency toward more overt marking is not restricted to Kuryłowicz'

[6]The more complex pattern is retained in the second and third person singular in NHG, where the 'go, drive' paradigm also has bipartite markers.

claimed preference for bipartite over simple markers." In fact, Hock seems to be skeptical of Kuryłowicz' 'law' as formulated, particularly given examples like (3.2)[7]; rather he seems to take it as one possible consequence of a more fundamental trend toward increased formal marking. In the end, he concludes that at the heart of the matter lies a tension between expressivity and formal simplicity, as manifested by "the tendency toward more overt marking" on one hand and "the tendency to eliminate or reduce (morphologically unimportant) alternations" on the other (Hock 1991: 234–237). Thus, both (52) as interpreted by Hock and (53) are compatible with a view of language change as favoring expressive but, all else being equal, simple forms[8]. This answers the concern raised by (52) that language change would tend to give rise to bipartite (discontinuous) markers, which would be at odds with the Fusion Assumption.

A final relevant notion from the historical literature is the idea that personal pronouns are a common diachronic source of agreement morphology (Hopper & Traugott 2003: 131). In contrast, Marianne Mithun cites personal communication with Talmy Givón, who reports that a discontinuous plural marker in Tolowa[9], an extinct Pacific Coast Athabaskan language, has evolved from a word meaning 'all' (Mithun 1999: 68). If the former case is more common than the latter, that is if pronouns are a more common diachronic source of agreement markers than quantificational elements, then agreement morphs might be expected to encode multiple agreement features[10] because a language's agreement markers would be more likely to preserve all of the contrasts expressed by the personal pronouns of the language.

To sum up the discussion so far, while the fusion assumption is not overtly present in the historical literature the observations pointed out here are consistent with a view of discontinuous agreement as a special type of form-meaning relationship that is diachronically less stable and less likely to arise than cumulative agreement. If this is right, then agreement morphs are expected to fuse agreement features most commonly, since language change is more likely to create cumulative markers and to

[7]Hock points out that examples (3.1) and (3.2) also show that (52) does not even hold across a given language. In light of the empirical focus of this thesis it is interesting to note that verb agreement is the counterexample; it remains to be seen whether this is purely coincidental or whether it is a more general fact about verb agreement.

[8]Note also that examples like (3.1), which are adduced as the most convincing examples of (52), involve vowel allomorphy in the root. It would be interesting to ask whether examples of bipartite markers involving multiple affixation are also frequent, since the addition of an extra affix would seem to pose a greater disruption to formal simplicity than a mere vowel alternation.

[9]Tolowa is closely related to Hupa, which is included in the language survey reported in chapter 2. The plural marker referenced here is very likely cognate with the discontinuous plural marker in Hupa, which in turn is cognate with elements in other Athabaskan language that are more quantificational or distributive in meaning.

[10]Or in the case of a language with analyzable pronouns, the morphs encoding the agreement features should be continuous.

turn discontinuous markers into cumulative markers than vice versa.

Theoretical work

In this section I examine two strands of theoretical research that relate to the fusion assumption; the first is the idea that coreferential agreement features constitute a logical set that is atomic in syntax. The organization of the natural class of phi features (the morphosyntactic agreement features, generally taken to include person, number, gender, and sometimes case) into logical sets is widely assumed. The idea that phi sets sets exist and that they have internal hierarchical organization has been argued for explicitly by many researchers, including Bonet (1991), Noyer (1992), and Harley & Ritter (2002).

It is a classical assumption within the Minimalist Program that these sets have syntactic reality. Agreement is taken variously as the spellout of an Agr-node specified for phi features (Chomsky 1991; Chomsky 1993), a pronoun-like set of interpretable phi features (Chomsky 1995b), or a set of valued uninterpretable features on a separate terminal such as T^0 (Chomsky 2000). What these analyses share is the bundling of phi features into sets with an atomic representation in syntax. This is a very natural move, considering that phi features describe the real-world properties of their referents and that a collection of coreferential phi features by definition describes a single entity. Just as lexical items can in theory be fully described by their component semantic features (for instance, the lexical item mare is characterized by the features female, equine, etc.), pronouns and agreement markers can be fully described by their component phi features.

The logical unity of agreement features is thus reflected by their syntactic unity in the various versions of Minimalist theory. Harbour (2008) maintains this idea within the context of discontinuous exponence. Harbour posits that phi sets are themselves syntactic categories with internal syntactic structure. Even in cases of discontinuous agreement, Harbour argues that the agreement features share a morphosyntactic locus despite their discontinuous surface order.

The second strand of theoretical work to be discussed here relates to the Bloomfieldian view of the morpheme discussed above. Discontinuous exponence has long been acknowledged as a problem for theories that adopt a view of the morpheme similar to Bloomfield's, those that assume a one-to-one mapping between elements of form and meaning at the sub-word level (Matthews 1974; Aronoff 1976; Carstairs 1987; Anderson 1992). Multiple exponence and discontinuous exponence were first referenced in the literature by Matthews (1974), whose term "extended exponence" I take as a cover term for certain subtypes of both phenomena. Matthews introduces extended exponence as the logical opposite of "cumulative exponence," which he defines as the realization of two or more inflectional features ("morphosyntactic categories") by one morpheme.

> The reverse of the cumulative case would be one in which a category, if positively identified at all, would have exponents in each of two or more distinct positions. But it would be very hard to exemplify a pattern as precise and general as this.
> –Matthews (1974: 149)

Matthews ultimately (and incorrectly) dismisses extended exponence as a phenomenon that is active only at the level of particular words and does not show the same paradigmatic regularity as cumulative exponence. The notion of extended exponence is taken up again by Carstairs (1987), who probes the relationship between morphosyntactic properties and their inflectional realizations. Assuming Matthews' view of morphosyntactic categories and properties, Carstairs suggests that the most *basic* (characterized by uniformity and transparency) and *expected* pattern is a one-to-one relationship between properties and inflectional morphemes. Variations from this pattern are classified according to cardinality and the surface relationship of the morphological exponents (syntagmatic or paradigmatic). Potential deviant patterns are listed in (54).

(54) Logically possible deviations from one-to-one patterning (Carstairs 1987: 14)

1. One property to many exponents syntagmatically
2. One property to many exponents paradigmatically
3. Many properties to one exponent syntagmatically
4. Many properties to one exponent paradigmatically

Carstairs' deviation 3 thus corresponds to Matthews' notion of cumulative exponence, and deviation 1 to his extended exponence, which overlaps substantially with my notion of discontinuous exponence. In work based on his 1992 dissertation, Noyer makes the related observation that "the relation between minimal syntactic positions... and positions-of-exponence... is not one-to-one *except in the default instance*" (Noyer 1997: xxxviii, emphasis mine). He terms the phenomenon in which morphological rules (or analogous devices) seem to operate across rule blocks (or strata, morphosyntactic heads, etc.) "discontinuous bleeding."

Importantly, the work of Matthews, Carstairs, and Noyer shares the view that there is a one-to-one association between form and meaning at the sub-word level in the general case; in the context of agreement, this implies a general expectation that agreement morphemes express agreement features cumulatively. Although their terminology differs, each of these authors acknowledges that discontinuous exponence exists and they agree that it represents a deviation from the more common situation of bidirectionally unique pairings of form and meaning at the level of the morpheme[11].

[11] Another notable and thorough critique of Bloomfield's view of the morpheme is presented by

Typological work

The fusion assumption extends to the typological literature, as well. For instance, in their chapter on inflectional morphology Balthasar Bickel and Johanna Nichols note that number agreement "is systematically marked in the great majority of languages having person agreement on the verb," and that it "often shares formatives or at least paradigms and position slots with person" (Bickel & Nichols 2007: 231).

In addition to statements like this there is at least one posited linguistic universal, namely number 401 in the Konstanz Universals Archive[12]:

(55) Konstanz universal 401 (Moravcsik 1994)
If in a language any inflectional categories are expressed cumulatively and the language does have person-number inflection, person-number will be among the cumulatively expressed distinctions.

These are just two examples of a very common type, however I have not been able to find any published empirical work to adequately support such claims. To the contrary, I have encountered just two potentially suggestive surveys and in both cases the results indicate a higher frequency of discontinuous agreement than the literature surveyed above would lead one to expect. Although I believe there is a real empirical basis for the fusion assumption, and that these studies might be slightly misleading with respect to the specific question at issue, I would be remiss not to mention them here.

First, chapter 35 of the World Atlas of Language Structures deals with plurality in independent personal pronouns. Michael Daniel reports that in 111 of the 261 languages examined, independent subject pronouns specify P and N discontinuously– a whopping 42.5%. Of these 111 languages, 69 express person and number via an impure continuity (26% of all languages considered), and 42 via a pure discontinuity (16% of all languages considered)[13]. An obvious problem with applying Daniel's results to this discussion is that he considers only pronoun structure, and not verb agreement. This is compounded by the fact that Daniel considers languages that are morphologically isolating and that therefore do not have verb agreement at all (e.g., Chinese). A final issue is that, in Daniel's words, "to keep the number of distinct types reasonable, the classification in this chapter always goes with the first person." In the results of the language survey presented in chapter 2 person limitations are

Anderson (1992), who builds on the claims of Aronoff (1976) that in the general case form pairs with meaning only at the level of complete words. This differs sharply from the other works presented here, which retain the idea of default bidirectional uniqueness at the morpheme level with discontinuous exponence and portmanteau agreement as special cases.

[12]Available online at http://typo.uni-konstanz.de/archive/intro/index.php.

[13]Of the remaining 150 languages, 139 express person and number cumulatively in personal pronouns, nine languages have pronouns that do not express number, and two languages do not have independent subject pronouns.

noted for 11 languages, and in 9 of those languages discontinuous agreement is limited to second person, third person, or both. This suggests that first person agreement markers may be the most likely to express person and number cumulatively.

The first (and only, to date) systematic study of one type of discontinuous exponence is reported by Trommer (2002), who surveyed approximately 100 languages with subject agreement affixes on the verb, 58 of which showed discontinuous marking of subject person and number. Again, this is a higher frequency of discontinuous agreement than is expected if the fusion assumption is sound. In this case selectional bias probably accounts for the surprising results: Trommer was not aiming to test the actual frequency of discontinuous patterns but rather to examine certain properties of discontinuous patterns. As such, he purposefully selected a large number of languages that show discontinuous agreement (Trommer 2002: 288).

To summarize, I have argued that there is broad and deep support for the fusion assumption in the historical, theoretical, and typological literature, but have noted also that there is a lack of robust data about the frequency of fused vs. discontinuous exponence. In the absence of such data I turn to Canonicality Theory, which offers a framework for grounding our expectations about morphological typology.

3.2 Canonical Agreement

This section introduces Canonicality Theory as developed by Corbett (2006). I begin with a terminological discussion, defining canonicality itself in 3.2.1 and considering the primitives of canonical agreement in section 3.2.2. In 3.2.4 I ask what canonical agreement morphology looks like and discuss several less-canonical morphological phenomena.

3.2.1 Canonicality

Canonicality is an intuitive but somewhat abstract notion, and as such it is difficult to pin down definitionally. Consider the following passage:

> "...I shall adopt a canonical approach. This means that I shall take definitions to their logical end point and build a theoretical space of possibilities. Only then do I ask how this space is populated. It follows that canonical instances, which are the best and clearest examples, those most closely matching the 'canon', may well not be the most frequent. They may indeed be extremely rare. However, they fix a point from which occurring phenomena can be calibrated."
>
> –Corbett (2006: 9)

Corbett thus distinguishes canonicality from frequency, establishing the usefulness of canonicality theory to the discussion of discontinuous agreement as a deviation from what agreement is expected to look like. We need not concern ourselves with the frequency with which agreement morphs express features discontinuously across languages, but rather we can ask to what extent discontinuous agreement resembles canonical agreement in general. It is also possible to evaluate the canonicality of particular (types of) discontinuous patterns.

Notice that Corbett defines canonicality in a fairly subtle way: a canonical pattern is one that *most closely* matches the definition of a phenomenon. In other words, if a particular phenomenon is typified by a particular constellation of characteristic properties then canonical instances of that phenomenon are those that exhibit all or most of those properties, while less-canonical instances exhibit fewer of the typical properties. For example, agreement typically involves a controller that expresses its features overtly, a target with bound, obligatory agreement morphs, and covariance of features on the controller and the target (Corbett 2006: 9). Less-canonical agreement might then involve optionality of agreement morphs on the target, or lack of overt feature expression on the controller.

Although this link is not explicitly discussed by (Corbett 2006), canonicality theory shares many of the underpinnings of prototype theory as developed by Rosch (1973 and following works) and others. A key difference between the two is that prototype theory does not reject frequency as a defining characteristic of prototypicality (Rosch 1978), while canonicality theory explicitly does so. According to prototype theory categories are not rigidly defined or strictly bounded, but rather are characterized by a central example and graded membership. A canonical example of a phenomenon is thus like the central, prototypical member of the conceptual category defined by the phenomenon, while a less-canonical example is like a peripheral example that only partially resembles the central member.

3.2.2 Agreement primitives

The primitives of agreement, by which I mean the most basic elements of the agreement relation, are taken by Corbett to be the controller, the target, the domain of agreement, agreement features, and any special conditions on the application of agreement. Each of these will be discussed in turn below.

The **controller** is the element in the domain of agreement that "triggers" agreement. it is the goal within the influential probe-goal model (Chomsky 2000 and following). Canonical controllers are overtly present within the domain, express their agreement features overtly, and control a consistent pattern of agreement across features and regardless of their own parts of speech.

The **target** is the element in the domain of agreement that, through the relation of agreement, winds up expressing some or all of the agreement features of the con-

troller. In the canonical case, agreement morphology is morphologically bound to the target, is obligatory and morphologically regular (§3.2.4), and is productive. Targets canonically show agreement with a single controller only; in less canonical cases of agreement, a target may agree with more than one controller or may "choose" among more than one potential controller.

Domains of agreement are used by Corbett to capture more information about the relation between target and controller; for instance, in a particular language verbs may be specified to agree with their subjects within the clausal domain. In this example the domain not only limits the scope of agreement to the clause, but it also adds information about which of the NPs within the clause (subjects) can trigger agreement. Domains are canonically local; long distance agreement is one example of agreement involving a noncanonical domain. They are also canonically asymmetric (for instance, the noun's features are typically marked on the verb, not vice versa).

Agreement features which, following Corbett (2006: 125), I take to be person, number, and gender, are inherent to the controller and end up being expressed on the target through the agreement relation. Features are canonically lexical (by which Corbett means formally assigned, not semantically based), with values that are deterministic and match across all expressions.

Conditions on agreement may limit the application of agreement to certain contexts. For instance, first and second person subjects may trigger agreement in a particular language, while third person subjects fail to do so. In the canonical case agreement is not subject to any special conditions.

The definition in (56) summarizes the preceding discussion and, based on Corbett, fixes a reference point against which observed agreement patterns can be compared. I will refer back to this definition in the following sections to discuss less-canonical agreement phenomena as deviations from this canonical picture.

(56) **Definition of canonical agreement (Corbett 2006: 8–26)**
Canonical agreement shows the following properties:

1. *Controller*: The controller is present, expresses its agreement features overtly, and controls a consistent agreement pattern regardless of its features or part of speech.
2. *Target*: The target shows bound agreement morphemes that are obligatory, morphologically regular (discussed further in my §3.2.4), and productive. The target agrees deterministically with a single controller, even if the controller is not overtly present. The target's part of speech is irrelevant.
3. *Domain*: The domain is asymmetric, local, and is one member of a set of domains.

4. *Features*: Features are lexical (based on formal assignment). Their values are deterministic and match across all expressions.
5. *Conditions*: There are no special conditions on agreement.

All of these facts are formulated by Corbett (2006) as a set of 20 criteria of canonical agreement. Each of these criteria specifies a "direction" of canonicality; for instance, condition C-15 states "local domain > non-local domain" meaning that is is more canonical for the domain to be local. There are three general principles, two described by Corbett and a third that I will suggest below, that can be used to motivate the direction of canonicality in these criteria, making the distinction of canonical agreement patterns less arbitrary and more principled.

3.2.3 Principles of canonical agreement

Redundancy Canonical agreement is redundant rather than informative (Corbett 2006: 11). This principle motivates, for instance, the canonicality of overt controllers (56.1), obligatory agreement morphology on the target (56.2), and feature values that match across all expressions (56.4).

Simplicity Canonical agreement is syntactically simple (Corbett 2006: 12). This principle is reflected in the canonicality of a consistent pattern of agreement (56.1), the fact that the canonical target's part of speech is irrelevant (56.2), the canonicality of local domains (56.3), and the preference for no special conditions on agreement (56.5).

Identifiability To Corbett's redundancy and simplicity I would add a third general principle: identifiability. By this I mean that it is easy to identify which formal segments are marking agreement and which features are being expressed, and that it is possible to unambiguously determine which argument(s) the agreement morphology cross-references. This principle is intended to motivate certain aspects of canonical agreement that cannot easily be accounted for by redundancy or simplicity, such as the regularity of agreement morphology on the target (56.2), the preference for deterministic feature values (56.4), and many of the properties of canonical morphology to be discussed immediately below.

3.2.4 Canonical morphology

Canonical agreement is realized by canonical inflectional morphology (Corbett 2006: 27, 72). In the typical case, inflectional morphology is tightly phonologically **bound** to a host; agreement morphology bound to the target is thus more canonical than agreement realized by an independent word, for instance. Within bound forms, affixes

are more canonical than clitics. I suggest that the principle of identifiability structures these observations: bound, affixal morphology is the most identifiable as agreement because it always appears on the target, versus clitic agreement, which may be more promiscuous in its attachment to a host.

Canonical agreement morphology is also **obligatory**. Identifiability is at work here, too, for obvious reasons: it is not possible to identify agreement morphology as such, or to determine what the controller is and which of the controller's features are involved in the agreement relation, unless the agreement is overtly present. The obligatory nature of canonical agreement (in combination with the preference for overt controllers) also reflects the principle of redundancy, which requires that the same information be overtly expressed on both the controller and the target.

In addition to being bound and obligatory, **regular** morphology is more canonical than irregular. Examples of irregular agreement morphology include suppletive, reduced, or periphrastic agreement or multiple exponence of agreement features[14]. The principle of identifiability is again evident in the canonicality of regular morphology: suppletive agreement morphs can be difficult to identify as expressing the same features as their regular counterparts and with reduced agreement there is overt expression of some, but not all, agreement features. The principles of simplicity is relevant as well: periphrastic agreement and cases of multiple exponence can be syntactically more complex than their morphologically regular counterparts.

3.3 Discontinuous Agreement as Noncanonical Agreement

Although, as I have emphasized throughout this chapter, the fusion assumption (§3.1) has not been empirically well-supported, I do believe that is assumed for good reasons. Canonicality theory (§3.2) offers a rich framework for thinking through in a detailed way why this should be so, without focusing on (as yet unavailable) frequency data. Corbett (2006) does not address discontinuous exponence as such, and so this chapter will compare the range of discontinuous agreement patterns discussed in the preceding chapter against the definitional criteria for, and morphological characteristics of, canonical agreement in general.

My aim here is simply to work through the ways in which discontinuous agreement patterns with noncanonical agreement. Although the claims made here will still need to be examined empirically, canonicality theory offers one possible way to ground the fusion assumption[15]. By supporting the expectation that agreement features are

[14]For a complete discussion of noncanonical agreement morphology see chapter 4 of Corbett (2006). Here I mention only the phenomena that will be relevant to the discussion of discontinuous agreement in section 3.3.

[15]Though note that that the correctness of canonicality theory itself is orthogonal to the correct-

expressed cumulatively in the default case, canonicality theory also bolsters one of the main claims of the next chapter: Discontinuous exponence results from a derivation that, other things being equal, is more complex than one giving rise to cumulative agreement.

Definitional evidence

Recall the definition of canonical agreement presented in section 3.2.2 and repeated here as (57).

(57) **Definition of canonical agreement (Corbett 2006: 8–26)**
Canonical agreement shows the following properties:

1. *Controller*: The controller is present, expresses its agreement features overtly, and controls a consistent agreement pattern regardless of its features or part of speech.
2. *Target*: The target shows bound agreement morphemes that are obligatory, morphologically regular (discussed further in my §3.2.4), and productive. The target agrees deterministically with a single controller, even if the controller is not overtly present. The target's part of speech is irrelevant.
3. *Domain*: The domain is asymmetric, local, and is one member of a set of domains.
4. *Features*: Features are lexical (based on formal assignment). Their values are deterministic and match across all expressions.
5. *Conditions*: There are no special conditions on agreement.

The discontinuous agreement patterns in the data set discussed in chapter 2 show both canonical and non-canonical controllers: there are both overt and missing controllers, and some of the controllers express their agreement features overtly while others do not. The domains of agreement are overall canonical in that they are asymmetric (the noun's features are marked on the verb, not vice versa), local, and clausal (a canonical domain type). With respect to the remaining agreement primitives, on the other hand, there is a strong tendency in the direction of noncanonicality.

Target Property (57.2) requires that agreement morphemes on the target be "morphologically regular". More specifically, canonical agreement morphemes are affixal and non-suppletive. In chapter 2 I pointed out various strategies languages use to realize coreferential features discontinuously, with root and affix suppletion being quite

ness of the fusion assumption. If this approach turns out to be untenable, the fusion assumption may be supported in a number of other ways.

common. For instance, see (12), repeated here as (58), in which number is expressed by a dedicated suffix and is additionally fused on the person-encoding prefix, and (13), repeated here as (59), in which number is expressed by a suffix and is also fused with the verb root, which takes a special form in the context of a plural subject. The frequency of suppletion in the agreement systems of the languages surveyed in chapter 2 supports the idea that discontinuous agreement is noncanonical.

(58) **ì**-los-e-**tè** eèsi
2PL-go-ASP-**PL** 2PL
'You (pl) will go.' (Dimmendaal 1983: 122) Turkana

(59) **ku**-t-o-i-**kaŋ-qut-n**
2-DET-THM-PST-**2**-walk.**PL-PL**
'You (pl) walked.' (Georg 2007: 219) Ket

Features Canonical agreement property (57.4) requires that feature values match across all expressions. Combinatorial discontinuous agreement examples thus pattern with noncanonical agreement because a single feature category is expressed by two distinct component values, both of which much be known in order to yield the correct interpretation. In example (19), repeated here as (60), the category of number is expressed twice: the person-encoding prefix *di-* also encodes nonsingular number (dual or plural) and the prefix *ya:-* encodes strictly plural number (not dual).

(60) do:-**ya:-di**-l-yo'
NEG-**PL**-1**NSG**-CLS-love
'We (pl) do not care for it.' (Sapir & Golla 2001: 303) Hupa

The two expressions of number in this example fail to match because they encode two distinct values for the same feature. I will take up the question of how examples like this might arise in section 5.4.1.

Conditions Finally, (57.5) states that conditions on agreement are noncanonical. In fact, although agreement conditions were not explicitly treated in chapter 2 and although they are presumably not inherent to the phenomenon of discontinuous exponence itself, they are extremely common throughout the discontinuous agreement data examined. Within a language, discontinuous patterns are often limited to certain person values or certain paradigms. For instance, one type of person-number discontinuity in Dumi verb agreement is conditioned on person. In (61a) the third person plural prefix *ham-* marks number discontinuously from the non-first person subject marking suffix *-a*. The third person singular form in (61b) lacks overt number marking. For comparison, note that the second person plural form in (61c) lacks the

discontinuous number-encoding prefix; Van Driem analyzes this form as being zero-marked for non-first person subject and instead showing a suffix *-ini* that fuses person (non-first) and number (plural) marking. Discontinuous number marking with *ham-* is thus limited to third person Dumi forms. Such limitations are extremely common throughout the results of the typological survey in chapter 2.

(61) Discontinuous agreement conditioned on person in Dumi

a. **ham**-phɨk-t-a
3PL.S-rise-NPST-2/3.S
'They get up.' (van Driem 1993: 163)

b. phɨk-t-**a**
rise-NPST-**2/3.S**
'He/she gets up.' (van Driem 1993: 162)

c. a-phɨk-t-∅-**ini**
MS-rise-NPST-2/3.S-**2/3PL**
'You (pl) get up.' (van Driem 1993: 162)

Thus from Corbett's definition alone it is clear that there are many important ways in which discontinuous agreement differs from canonical agreement. The non-canonicality of targets, features, and the high number of systems with agreement conditions in the survey data from chapter 2 support a treatment of discontinuous agreement as noncanonical agreement.

Morphological evidence

Another way to evaluate the canonicality of discontinuous agreement is to look only at surface patterns. In this section I consider two morphological phenomena considered by Corbett (2006) to be noncanonical, and show that each of them is characteristic of discontinuous agreement.

Multiple exponence Corbett (2006: 77–78) takes multiple exponence to be less canonical than single exponence, whether a single feature is realized multiple times by the same form or by different forms. In chapter 2, section 2.2, I showed that patterns of multiple exponence involving not only agreement features, but also TAM and negation, are common. In particular I noted that multiple exponence is a common secondary pattern within coreferential discontinuous exponence, as seen for instance in example (22).

Interestingly, some of the examples that Corbett includes under the rubric of multiple exponence fall into my definition of discontinuous agreement. With respect to example (62) Corbett points out the multiple exponence of third person plural

agreement. However, the form *te=i* also shows the impure discontinuous exponence of person and number, both of which are expressed by the prefix while the root fuses only number.

(62) te te=i t-i
3PL 3PL=fall.PL 3PL-go.3PL
'They fell over.' (Corbett 2006: 77) Skou

More generally, Corbett considers as multiple exponence the situation in which agreement with a single controller is found in more than one morphological slot on the target, even if the different agreement slots encode different features Corbett (2006: 78)[16]. As an example Corbett cites Maltese imperfective verbs, which agree with their subject prefixally for person and suffixally for number, which I would schematize as described in section 2.2 as *P-V-N*). Because multiple exponence is noncanonical, and because at least one major subtype of discontinuous agreement is taken by Corbett to be multiple exponence, there is direct evidence for the noncanonicality of discontinuous exponence.

Suppletion Suppletive morphology is also treated as noncanonical by Corbett (2006: 86), presumably because it disrupts simplicity (when the conditioning environment for the suppletive allomorph(s) is syntactic) and identifiability (§3.2.3). In particular, Corbett views suppletion triggered by contextual (including agreement) features as unexpected. As was noted above with respect to target canonicality, the discontinuous agreement data set shows a great amount of verb and affix allomorphy conditioned by agreement (and other inflectional) features. In this respect also, discontinuous agreement patterns with noncanonical morphology.

3.4 Summary

This chapter has argued that discontinuous exponence represents a deviation from the expected relationship between syntax and morphology, in that it involves a one-to-many relation between (sets of) morphosyntactic features on the one hand and morphs on the other. This claim presupposes that in the default case the relation between features or sets of related features and the morphs that realize them is in fact one-to-one. The fact that this presupposition, termed the "fusion assumption" and discussed in section 3.1, is held widely and with good reason has been the focus of the discussion.

[16]Notice that his classification of this type of pattern reflects the fusion assumption. Corbett assumes here that the agreement features of a single argument are bundled together, so that cross-referencing a single feature also cross-references the whole bundle; otherwise I would not take this to be multiple exponence.

As was noted in the chapter introduction, an ideal way to go about testing the fusion assumption would be to complete a major crosslinguistic survey of inflectional morphology, focusing on a closed set of inflectional features and asking with what frequency across (and perhaps within) languages the various combinations of features are expressed cumulatively versus discontinuously. Until such a study can be completed, I take as evidence in support of the fusion assumption its intuitive naturalness, its pervasive adoption in theoretical and typological literature, and the fact that canonicality theory supports a view of discontinuous agreement as deviant.

While perhaps none of the points of evidence considered in this chapter is fully satisfying on its own, the fact that several types of evidence converge in the ways discussed here is strongly suggestive. If we thus accept as a working hypothesis that discontinuous agreement stands as an exception to default agreement, in which agreement features are fused on a single morph, and if we recall that the discontinuous exponence of other types of features has much in common formally with that of agreement features (chapter 2, section 2.2), then we can reasonably conclude that although discontinuous exponence is a robust phenomenon across languages, it is still a deviation from the norm.

In the next chapter I will consider some of the necessary aspects of a theory of discontinuous agreement, arguing that the deviant nature of the phenomenon suggests that its derivation should be in some way more complex than the derivation of canonical agreement.

Chapter 4

The Contribution of Discontinuous Exponence

The central question of this chapter is what discontinuous exponence can illuminate about the necessary descriptive and explanatory properties of a complete theory of agreement. Section 4.1 lays out several theoretical desiderata: a good theory should capture both the robustness and the noncanonicality of the phenomenon and should account for certain empirical generalizations. In section 4.2 I examine several of the currently dominant theories of agreement and consider how they fare against these desiderata. Finally, section 4.3 argues that discontinuous exponence is an essentially morphosyntactic phenomenon that cannot be fully accounted for by a purely syntactic theory.

4.1 Desiderata for a Theory of Discontinuous Agreement

Chapter 2 established the robustness of discontinuous exponence as a phenomenon and Chapter 3 demonstrated its noncanonicality. A complete theory of agreement, then, should offer a unified treatment of the different types of discontinuous exponence (§4.1.1). Additionally, it should both account naturally for inflectional discontinuity and explain why it is not the default pattern (§4.1.2).

Languages with discontinuous agreement show us that, at least in some cases, expression of information is more important than economy of form. A good theory should commit to a set of factors that can override economy (§4.1.3).

Finally, the results of the typological study show several crosslinguistic tendencies in discontinuous exponence. This section focuses on two: that it is possible to make strong generalizations about the relative order of agreement morphs (§4.1.4) and that certain types of discontinuous exponence can give rise to referential ambiguity

(§4.1.5).

4.1.1 Discontinuous exponence as a single phenomenon

The language study in chapter 2 showed that discontinuous morphology can be found in several different domains; I adduced examples of discontinuous verb agreement, TAM morphology, and pronoun formation. There are several good reasons to give discontinuous exponence a unified treatment across these domains. First, the morphosyntactic types of discontinuity cross-cut these domains: coreferential, combinatorial, and multiple exponence are found in both verb agreement and pronoun formation. Even more convincing, within a language the order of the features as expressed by pronoun formatives tends to parallel the order of features as expressed by agreement affixes on the verb. In some cases, features are realized with the same phonological form in both the verbal and the pronominal domain.

Although coreferential discontinuous exponence is limited to the set of agreement features, and thus for obvious reasons is not found in the domain of TAM morphology, it is the case that both the combinatorial and the multiple exponence types extend into this domain. Recall from chapter 2 examples (19) and (33), repeated here as (63) and (64). In (63) number is expressed discontinuously by two morphs: a general one that encodes nonsingularity and a more specific one that encodes strict plurality (three or more participants). In (64) tense is expressed discontinuously by two morphs: a general one that encodes general past tense and a more specific one that encodes definite past.

(63) do:-**ya:**-**di**-l-yo'
NEG-**PL**-1**NSG**-CLS-love
'We (pl) do not care for it.' (Sapir & Golla 2001: 303) Hupa

(64) al-**g**-a-bi-**ru**-mo-uba-go-ow.al-wado-go
ASSERT-2/3.S.**PST**-HAB-TRI.S-**DPST**.HAB-NSG.S-bad-EMPH-do-REP-HAB
'They three were certainly repeatedly causing trouble as a habit.'
(Wurm 1975: 342) Kiwai

The fact that the patterns of discontinuous exponence in pronoun formation and TAM inflection so closely parallel those in verb agreement argues strongly for a unified account of the phenomenon. Whatever analysis and theoretical mechanisms I propose for discontinuous verb agreement should therefore extend naturally to these related data. They should help explain pronoun formation, so they should not be too tightly tied to the verbal domain. They should be applicable to TAM morphology, so they should allow for the participation of non-agreement features. These qualities will also help make a theory of agreement flexible enough to account for potential other types of discontinuous exponence that have not yet been studied.

4.1.2 Noncanonicality of discontinuous agreement

Chapter 3 discussed the pervasiveness of the Fusion Assumption (§3.1), the idea that person and number are most commonly fused on a single morph, and argued that there are many good reasons to accept it as fact. The logical extension of the Fusion Assumption is that the relationship between form and meaning is one-to-one in the default case. For verb agreement I thus expect that in general a single unit of meaning (a set of agreement features) will have a single morphological exponent (an agreement marker). At the same time, the existence of discontinuous exponence proves that this relationship may be manipulated.

In this way discontinuous exponence can be understood as a disruption of the one-to-one relation between syntactic feature sets and morphological feature expressions. The expectation is that fused agreement is more basic, more canonical, and more frequent than discontinuous agreement. By the same reasoning forms with single discontinuities are expected to be more basic, canonical, and frequent than those with multiple discontinuities. A good theory should straightforwardly derive the default patterns while still allowing for the noncanonical patterns and, ideally, grounding their noncanonicality in some general way.

4.1.3 Full expression

Given that exponence is fused in the default case, I assume that the lexical inventory of a typical language includes agreement markers that fuse multiple features. However, in a language in which the lexical inventory contains agreement markers that express only a subset of agreement features there are two (synchronic) alternatives for encoding agreement. One option is to pick the best marker, the one that encodes the most features matching the cross-referenced argument. Another is to pick multiple markers, which in aggregate encode the cross-referenced argument more fully. The later strategy clearly gives rise to discontinuous agreement. The fundamental difference between these alternatives is expressiveness: The first option sacrifices full expression for the sake of simplicity while the second prioritizes full expression, even at the cost of added complexity.

The existence of all of the patterns in all of the languages studied in chapter 2 supports the idea that there exists in language a drive to fully express meaning, even if it takes multiple morphs to do so. A good theory of discontinuous agreement should capture this tendency toward full expression and connect it to the derivation of discontinuous patterns.

	Both prefix	**Both suffix**	**Mixed**	**All**
P>N	9 (90%)	22 (73.3%)	39 (97.5%)	70 (87.5%)
N>P	1 (10%)	8 (26.7%)	1 (2.5%)	10 (12.5%)

Table 4.1: Order of person and number in discontinuous agreement (Trommer 2002: 89)

Order	**Number**
P>N	50 (70.4%)
N>P	16 (22.5%)
Both/neither	5 (7%)

Table 4.2: Order of person and number in discontinuous agreement (my survey)

4.1.4 Morpheme order in discontinuous agreement

Section 3.1 mentioned the work of Trommer (2002), who surveyed approximately 100 languages with subject agreement affixes on the verb, 58 of which showed discontinuous marking of person and number. Focusing on affix order, Trommer found a strong cross-linguistic tendency in such languages for person marking to precede number marking.

The 58 languages in Trommer's survey showed 80 different ordering patterns. Of these patterns, 12.5% involved prefixed person and number, 37.5% showed suffixed person and number, and the remaining 50% showed a mixed pattern with one prefix and one suffix. Table 4.1 summarizes the ordering of person and number marking in all 80 patterns, which include both pure and impure discontinuities. The first row tabulates patterns in which person precedes number, and the second row patterns in which number precedes person. The dominance of the first row shows the tendency of person marking to precede number marking; I call this "Trommer's Generalization," as shown in (65).

(65) Trommer's Generalization
Person marking precedes number marking

In terms of linear order, my survey results are consistent with Trommer's. Table 4.2 shows that person marking most commonly precedes number marking, though the tendency is not as strong (70.4% in my survey vs. 87.5% in Trommer's). Unfortunately the 17 survey patterns in which gender is marked do not reveal any strong tendencies in the relative order of gender on one hand and person or number on the other (Table 4.3).

Order	**Number**
G>P	1 (16.7%)
P>G	3 (50%)
Both/neither	2 (33.3%)
G>N	5 (38.5%)
N>G	7 (53.8%)
Both/neither	1 (7.7%)

Table 4.3: Order of gender with respect to person and number in discontinuous agreement (my survey)

A descriptively adequate theory of agreement must account for the strong tendency for person marking to precede number marking in discontinuous agreement.

4.1.5 Ambiguity

It was noted in Chapter 2 that when an agreement feature splits away from the other features in its set, semantic ambiguity may arise. Examples were given in (49–51), repeated here as (66–68).

(66) a꞉= **ya꞉-xo**-ɬ-**ch'i**-de꞉-ne'
thus= **PL-3.A.O**-COM-**3.A.S**-THM-say
'he said to them' / 'they said to him' / 'they said to them'
(Sapir & Golla 2001) Hupa

(67) mo-g-k'al-i-t < mo-**v**-**g**-k'al-i-**t**
PVB-**1.S-2.O**-kill-THM-**PL**
'I kill you (pl.)' / 'we kill you (sg.)' / 'we kill you (pl.)'
(Hewitt 1995) Georgian

(68) **i**-kamoso-**hig**-ak-e-**ri**
3M.S-visit-**PL**-PFV-REALIS-**3M.O**
'he visited them' / 'they visited him' / 'they visited them'
(Michael 2008) Nanti

Recall from section 2.3.4 that all of the ambiguous examples in the results of my language survey involve the the pure discontinuous expression of person and number. They also occur in verbs that cross-reference more than one argument. In these patterns, the separately marked number feature can be interpreted to mark the number of one argument or the other, or sometimes both.

However, there are patterns in the survey that do show a pure person-number discontinuity in verbs referencing multiple arguments but do not yield various possible

interpretations. There are two strategies that these languages employ that function to resolve ambiguity. One strategy is to fuse case with the number morpheme, thereby explicitly indicating which argument's number is being encoded. Other languages use dedicated ordering to disambiguate; for instance, Maricopa has a prefix that encodes the number of the verb's object only.

A theory of agreement should be able to explain how referential ambiguity arises and why it is limited to the particular context described above.

4.2 Challenges for Existing Models of Agreement

This section assesses several current theories of agreement, considering how they fare against the desiderata laid out in the preceding section. To summarize, in order to fully account for discontinuous exponence a theory should be able to do all of the following.

1. Allow a unified treatment of discontinuous agreement, pronoun formation, and TAM inflection (§4.1.1)
2. Straightforwardly account for discontinuous morphology while explaining its noncanonicality (§4.1.2)
3. Capture the drive to fully express meaning (§4.1.3)
4. Derive ordering tendencies (§4.1.4)
5. Explain the source of referential ambiguity (§4.1.5)

A theory of discontinuous agreement should use generalizable mechanisms and operations that can apply in other domains of inflection. Ideally the derivation of discontinuous agreement should be more costly in some way than the derivation of fused agreement. The principle of parsimony then dictates that discontinuity will arise only when it serves some greater purpose, grounding both the noncanonicality of discontinuous agreement and the tension between economy and expressiveness.

4.2.1 Fundamentally syntactic theories

Standard Minimalism

The Minimalist Program (Chomsky 1993 and following work) is a framework for syntactic analysis that aims to pare down phrase structure and syntactic derivations to just the elements that are required for grammaticality. Operations are driven to match uninterpretable features (those that do not contribute to the meaning of

the utterance) with interpretable features (those that contribute to meaning, such as past tense or plural number). The principle of Full Interpretation dictates that all uninterpretable features must be checked at the conclusion of the derivation; if this condition does not obtain, the derivation fails (Adger 2003: 85). The realization of these features is, for the most part, left to the morphology.

Under Minimalism, two nodes enter into an agreement relation when they are in the proper structural configuration and when an uninterpretable features on one node (the "probe") matches and checks or values an interpretable feature on the other (the "goal"). The probe will match with the structurally closest goal that bears the relevant feature; for instance, a probe high in the tree will agree with a subject with matching features, but in case the subject's features do not match it will skip over the subject and probe the object.

Minimalism fares poorly against the desiderata in (4.1), mainly because it a theory of syntax and it leaves aside most questions of morphology. Although Agree is a general operation that is used to value uninterpretable TAM features as well as agreement features, there is no way to account for discontinuous expression within this framework and thus no way to unite discontinuous agreement with other kinds of discontinuous exponence. Because Minimalism has little to say about morphology, it also does not explain the existence or the noncanonicality of discontinuous agreement and it cannot derive the strong tendencies in the order of agreement markers across languages.

There is some overlap between the Minimalist principle of Full Interpretation and the theoretical goal of full feature expression. Probes will always be fully valued in a converging operation and so all that remains is to ensure that all probe features are realized in the morphology.

With respect to ambiguity, the specific features of the lexical items that enter into a derivation and the order in which they are merged can help us to understand how ambiguity can arise[1], but Minimalism cannot connect ambiguity to a particular type of morphological discontinuity.

Cyclic Agree

The theory of Cyclic Agree, developed by Béjar (2003) and Béjar & Rezac (2009), proposes that context sensitive agreement, which occurs when multiple arguments compete for the control of a single morphological slot, can be derived by a finely articulated probe together with certain assumptions about the syntax of the Agree operation. Each language has a characteristic way of specifying the probe, which derives variation in context sensitive patterns while capturing the commonalities across

[1]For instance, a number probe high in the tree may be valued by either the subject or object, whichever is the closest node with the matching feature. In this way, multiple underlying structures can give rise to nonspecific argument plurality.

1.	Unified treatment of agreement, pronouns, TAM	X
2.	Explain both the existence and noncanonicality of DA	X
3.	Capture full expression	*
4.	Derive ordering tendencies	X
5.	Explain referential ambiguity	X

X = no; * = yes, with extensions

Table 4.4: Comparing Standard Minimalism against the list of desiderata

languages. Like the Minimalist Program, Cyclic Agree treats agreement as an essentially syntactic phenomenon.

The empirical basis of this research is agreement that is sensitive to a person hierarchy. Specifically, Béjar and Rezac focus on languages in which multiple arguments compete for a single morphological expression as in Basque, in which the object generally controls verb agreement (69a) but the subject may control agreement if it outranks the object (69b).

(69) Basque: Person hierarchy sensitivity

a. ikusi **z**-in-t-u-da-n
seen **2**-X-PL-have-1-PST
'I saw **you**'

b. ikusi **n**-u-en
seen **1**-have-PST
'**I** saw him'

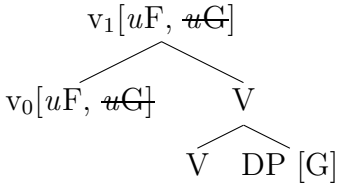
Cyclic Agree adopts the syntactic framework of Bare Phrase Structure, in which derivations are built from the bottom up, the lexical items determine the structure (there is no prescribed phrase structure as in X-bar theory), and there is no distinction between a head and a terminal (Chomsky 1995a). Béjar and Rezac carry forth the assumptions that selector labels project upon Merge and that labels are more or less equivalent to lexical items. For instance, in example (70) V selects DP and so it projects to create the upper V node. As a result, after projecting the probe is in the right configuration with the subject to trigger Agree[2].

For instance, a probe located on v_0 will probe the object upon merging with V as schematized in (70), in which F and G are abstract features and uF and uG the corresponding uninterpretable, unvalued abstract features. The DP object of the verb is specified for G, the probe is searching for a G value, and the object's G value is

[2]Béjar and Rezac point out that one could alternatively model the projecting probe in terms of v raising to a higher head like T.

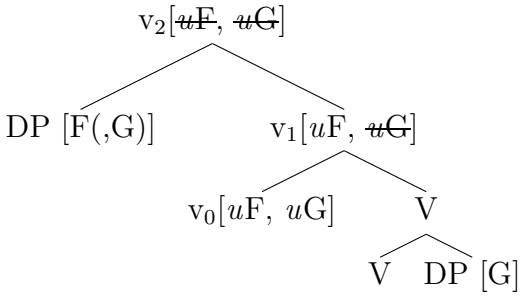
thus copied to the probe. The object has partially valued the probe, as shown by the strikethrough of the uninterpretable uG feature. Because the probe's uF feature remains unvalued, the entire probe projects as v_1.

(70) Agreement with object on the first projection of v



The fact that the probe is only partially valued licenses an additional cycle of agreement, in which the subject is probed as shown in (71). The DP subject is also valued for G (although this is not relevant, since the probe's G feature has already been valued by the lower argument) and is additionally valued for feature F. The subject values the remaining features of the probe – its F value is copied to the probe – and agreement is complete. The probe is fully valued, as shown by the strikethrough of uF and uG on the highest probe projection, v_2.

(71) Agreement with subject on the second projection of v



A key element of Béjar and Rezac's work is the idea that feature categories in agreement probes have fine-grained internal structure. For instance, a person probe that does not privilege any particular value would be specified as [uP], meaning that it probes for any value of person. Feature values are encoded as shown in Figure 4.1; note that first, second, and third person are all specified with P, hence all match a [uP] probe. A value lower in a feature structure entails the values higher in the structure: in order to be a speaker, one must be a speech act participant; in order to be a participant, one must be a person. Because of these entailments, the feature set encoding a less-specified value (e.g. third person) is a proper subset of the set encoding a more specified value (e.g. first or second person)[3].

[3]The encodings shown here are for a language in which first person is the most highly specified, but this isn't the case in all languages. As such, the fine feature structure of arguments and probes may vary cross-linguistically. For instance, in a language in which second person is the most highly

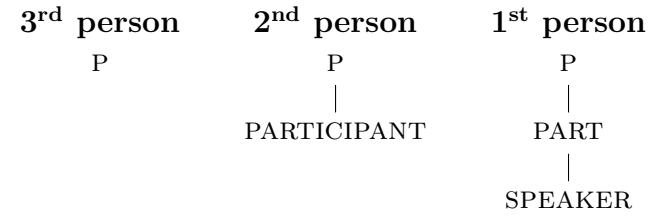


Figure 4.1: Encoding values of person

In a context sensitive language, for instance one in which agreement with a local (first or second) person is privileged, the probe would be specified as [*u*P, *u*PARTICIPANT]. This would mean that the argument nearest the probe would control agreement if it were first or second person[4]. If the nearest probe were third person, though, another cycle of Agree could additionally value the probe with features from another argument. In this way, a single probe can agree with more than one argument, deriving context sensitive agreement. In the hypothetical examples described above, the object values the probe as much as it can and the subject ends up controlling agreement only if it can add to the probe's value. The fact that different languages have different characteristic probe structures explains much of the crosslinguistic variation in agreement, for instance the preference for subject or object to control agreement, or the possibility of context agreement sensitive to a language-specific person hierarchy.

Cyclic Agree is primarily concerned with the bundling and distribution of agreement features in the syntax, and does not directly address the morphological realization of those features. However, Béjar (2003) does briefly address the issue of discontinuous agreement, proposing that particular syntactic configurations favor the insertion of markers that fuse person and number features. Specifically, she suggests that if the person and number probes are valued on the same functional projection (in other words, by the same argument) the result is a single vocabulary insertion site that is valued for both person and number: "a natural candidate for a vocabulary insertion rule that makes reference to both sets of features" (Béjar 2003: 160–161). This is an elegant and appealing idea, but based on a quick evaluation against the results of my typological study it does not seem to adequately predict the shape of a

specified, first person would be encoded as [P [PARTICIPANT]] and second person as [P [PARTICIPANT [ADDRESSEE]]]. I will return to the idea of fine-grained intra-category structure in section 5.2.2, incorporating it into my proposal for the structure of ϕ-sets (§5.2.3), and will return to the idea of language-specific probes in section 5.3.1.

[4]Under Béjar's view, the Agree operations copy whole feature structures. Thus, if a [*u*P, *u*PARTICIPANT] probe agrees with a [P [PARTICIPANT [SPEAKER]]] argument, the [speaker] value is copied to the probe along with the rest of the features and triggers the proper agreement marking.

1.	Unified treatment of agreement, pronouns, TAM	X
2.	Explain both the existence and noncanonicality of DA	X
3.	Capture full expression	✓
4.	Derive ordering tendencies	X
5.	Explain referential ambiguity	✓

✓= yes; X = no

Table 4.5: Comparing Cyclic Agree against the list of desiderata

language's agreement markers (see Appendix B).

Even if the idea that valuing probe features on a single projection creates a bias toward fusion turns out to have some truth, it still has mixed results when evaluated against the goals developed in section 4.1. The environment creating this bias is entirely syntactic and is a product of the Agree operation, applicable only to agreement features, and so it cannot help us to unify the treatment of discontinuous verb agreement, pronoun formation, and TAM morphology. However, the valuation of a probe on a two distinct projections is clearly more derivationally costly than valuation on a single projection; this fact could be leveraged to explain the noncanonicality of discontinuous agreement, particularly if there is any external evidence that probes are typically able to be fully valued by a single argument. What the theory fails to explain is discontinuous agreement in languages without context sensitivity, and there are such languages present in my survey of discontinuous exponence. In these languages, multiple cycles of Agree are not predicted and so discontinuity would need to explained in other terms.

There is some commonality between Cyclic Agree and the idea of full expression. Recall that just in case the first argument fails to fully value the probe, another cycle of Agree is licensed and another argument can contribute to the probe's value. This mechanism could easily be grounded in a more general notion of full expression.

Because Cyclic Agree is not directly concerned with the morphological realization of agreement features it does not have much to say about ordering tendencies. However, the idea that a number probe may behave independently of a person probe and may potentially be valued by more than one argument would be quite useful in deriving referential ambiguity.

4.2.2 Fundamentally morphological approaches

Distributed Optimality

Motivated in part by his observations about the crosslinguistic tendencies in agreement marker order (§4.1.4), Trommer proposes his theory of Distributed Optimality,

a "constraint-based, modular version of Distributed Morphology" (Trommer 2002: 285). Trommer follows the assumption of Distributed Morphology that morphological operations apply directly to the output of syntax, altering morphosyntactic structure and adding phonological content. His proposal differs from standard Distributed Morphology in the idea that each step of syntax and morphology operates according to the principles of Optimality Theory (Prince & Smolensky 1993; McCarthy & Prince 1994), in which inputs are mapped to outputs using ranked, violable constraints to surface the most optimal output for a given input.

Trommer assumes that nodes may split in the morphology, and that this split is what gives rise to discontinuous agreement. This notion leads nicely to a view of discontinuous exponence as more derivationally complex than fused exponence and it could be leveraged to explain the noncanonicality of discontinuous exponence. However, Distributed Optimality carries over the assumption of Distributed Morphology that agreement has a special status in the grammar: unlike lexical heads, which are already present in the syntactic module, agreement nodes are not inserted until the morphology. This distinction leads Trommer to assume that only agreement morphemes can split during vocabulary insertion (Trommer 2002: 286, 297), which limits the ability of the theory to unify the treatment of agreement, pronoun formation, and TAM morphology.

The fact that agreement morphemes split to allow multiple insertion of agreement markers fits nicely with the idea of full expression of meaning: morphs will be inserted as long as they match the node's features. However, it is not entirely clear what Trommer takes to be the cause. Does the drive to express meaning trigger splitting and thus license multiple insertion? Or does multiple insertion just apply mechanically, triggered by the presence in the derivation of matching lexical items, so that full expression is a side effect? Still, with minimal effort this theory could straightforwardly implement full expression and connect it to the patterns of discontinuous agreement.

The complementary distribution of fused and discontinuous agreement led Trommer to assume that each kind of agreement comes from a single head. Because he also assumes that the features of a head are unordered, he is led to conclude that there can't be a syntactic explanation for the person-before-number ordering tendency (Trommer 2002: 286). He instead analyzes the tendency in terms of a pair of alignment constraints; one assigns a violation for every vocabulary item that intervenes between the person marker and the left edge of the word and the other assigns a violation for every vocabulary item that intervenes between the number marker and the right edge of the word.

M-Case

Bobaljik's theory of M-Case proposes that agreement is a purely morphological process, and not a (narrowly) syntactic one. Specifically, he argues that verb agreement

1.	Unified treatment of agreement, pronouns, TAM	X
2.	Explain both the existence and noncanonicality of DA	✓
3.	Capture full expression	✓
4.	Derive ordering tendencies	✓
5.	Explain referential ambiguity	X

✓= yes; X = no

Table 4.6: Comparing Distributed Optimality against the list of desiderata

is controlled by the highest accessible NP in the clausal domain[5]. The proposal is based on languages in which case and grammatical function are not aligned, languages in which multiple arguments are accessible for agreement, and languages with long distance agreement.

The core of the argument is that morphological case (m-case) is distinct from syntactic licensing[6]. Bobaljik presents evidence from mismatches between case and grammatical function in Icelandic nominative objects that suggests agreement is only sensitive to the output of m-case assignment, and not to grammatical function. He further argues that m-case assignment follows syntax, by common assumption and following references cited. Therefore, he claims, agreement must be a post-syntactic operation because agreement depends on m-case and m-case follows syntax (Bobaljik 2008: 300–302).

The three primary types of m-case (following Marantz (1991) are lexically governed case (OBL), dependent case (ACC, ERG), and unmarked/default case (NOM, ABS). The m-case assigned to an argument factors into agreement in that it determines accessibility in one of two ways: In "Type 1" languages, only unmarked case is accessible. In "Type 2" languages, both unmarked and dependent case is accessible. Thus, if the highest argument receives unmarked case then it will control agreement, but if the highest argument's m-case is lexically governed then the lower argument may control agreement. The fact that the highest argument typically controls agreement maintains the apparent subject-orientation of single agreement systems and the fact that inaccessible NPs are invisible for the purpose of selecting the agreement controller derives context sensitivity – meaning that person hierarchies play no role in the competition of NPs for control of agreement – and apparent intervention effects.

Although this theory attempts to derive agreement strictly in the morphology, it

[5]This is only argued for languages in which the verb agrees with a single argument. Bobaljik suggests this could be extended to multiple agreement languages (Bobaljik 2008: 311, fn. 17), but it is not clear to me how this would work. The fact that person and number agreement behave differently in some languages (for instance, in Georgian the subject is the preferred number controller while the object is the preferred person controller) seems particularly problematic for this kind of extension.

[6]This is supported with evidence from Icelandic quirky case assignment (Bobaljik 2008: 298).

1.	Unified treatment of agreement, pronouns, TAM	X
2.	Explain both the existence and noncanonicality of DA	X
3.	Capture full expression	X
4.	Derive ordering tendencies	X
5.	Explain referential ambiguity	X

✓= yes; X = no

Table 4.7: Comparing M-Case against the list of desiderata

is in a sense more like the syntactic theories in that its primary focus is the establishment of target-controller relationships. As such, it does not address the issue of the exponence of agreement features. On its own, it therefore has little bearing on discontinuous agreement. There are no mechanisms proposed to derive discontinuity, so of course there is nothing to extend to the domains of pronoun formation and TAM morphology and there are no predictions about the ordering of agreement feature markers. The existence and noncanonicality of discontinuous agreement and the principle of full feature expression would need to be derived by mechanisms external to case assignment, so they are not accounted for by anything in this theory.

Finally, the theory in its current form only derives agreement with a single argument. It thus cannot explain examples of referential ambiguity, which crucially involve agreement with more than one argument.

4.3 In Favor of a Morphosyntactic Model of Discontinuous Agreement

The theories reviewed in section 4.2 were, for the most part, not developed to account for discontinuous exponence and are not easily able to do so in their current forms. A standard Minimalist treatment of agreement fell short on all of the theoretical desiderata developed in section 4.1, mainly because the sharp distinction between syntax and morphology inherent in this approach affords no way to talk about the relationship between syntactic arrangements and morphological realizations. Likewise, the M-Case approach aims mainly to show a new way to structure the relations between agreement controllers and targets. As it is limited to single agreement languages and does not consider how features are expressed in the morphology, it has little to say about discontinuous exponence.

Cyclic Agree also focuses on the syntax of agreement, but provides a potential view into discontinuous morphology by decomposing probes into very fine structures and allowing them to be valued by different syntactic heads. This in effect pushes the

theory closer to the syntax-morphology interface and leads Béjar to wonder whether certain syntactic arrangements favor fused versus split agreement markers. Although her particular suggestion is too closely tied to the phenomenon of context sensitivity to be generally applicable to discontinuous exponence, the question of what constitutes an appropriate syntactic context for the insertion of discontinuous morphology is extremely important and I will spend a good deal of time addressing it in the next chapter.

Another important contribution of Cyclic Agree is the idea that coreferring agreement features can behave independently of one another. In chapter 6 I will point out some problems with the strict separation of person and number probes. I take these issues to support the analysis I will develop in the next chapter, in which the coreferring agreement features are bundled in syntax but may break apart in the morphology.

The theory of Distributed Optimality has the explicit goal of explaining one particular pattern of discontinuous exponence, namely the strong tendency for person agreement to precede number agreement when those features are marked separately in the morphology. As such, it held up well against several of the criteria developed in section 4.1. Because it assumes that agreement features are bundled in the syntax and split by an operation in the morphology, it can easily explain both the existence and the noncanonicality of discontinuous agreement. However, Trommer assumes that in this regard agreement has a special status and unique properties, making it difficult to extend the analysis to pronoun formation and TAM inflection.

The alignment constraints of Distributed Optimality do correctly derive the ordering tendency with regard to split person and number agreement, but because there is no independent evidence or other grounding for the constraints the analysis is not fully satisfying. The next two chapters will develop an account that instead relies on hierarchical relations inside the agreement node to account for the person-before-number generalization.

The syntactic models reviewed here focus on setting up probe-goal relations, which are necessary to correctly derive the patterns of control in agreement but do not help us to understand the morphology of agreement. Any model of agreement that is primarily syntactic cannot effectively derive discontinuous agreement or explain why it is noncanonical. Furthermore, syntactic models tend to afford agreement a special status and assume that agreement operations have unique properties. This may well be true in the syntax but I have already noted that the patterns of discontinuous exponence generalize across domains of inflection, suggesting that it is best derived in the morphology.

Looking forward, what is needed is a theory of agreement that both retains the insights of the syntactic models and pushes further by committing to fully account for the morphological realization of agreement. It will also need to ground the principle of full expression in some way and to derive discontinuous exponence and capture its

noncanonicality using operations not limited to a particular domain of inflection or set of features.

Chapter 5

Deriving Discontinuity: Cyclic Insertion

5.1 Introduction

This chapter develops a theory that aims both to naturally derive discontinuous exponence and to explain its noncanonicality. The syntax is a straightforward adoption of Béjar and Rezac's theory of Cyclic Agree (introduced in §4.2.1 above, to be further discussed in §5.3.1). The real work will be done in the morphology, which takes as its starting point the framework of Distributed Morphology (Noyer 1992; Embick & Noyer 2007). I propose substantial refinements to the latter theory and introduce some necessary analytical devices.

Specifically, I propose that ϕ-sets with rich internal structure are at the heart of both the syntax and the morphology of agreement (§5.2.3, 5.3.2, and 5.4). In making this argument I pick up on Béjar's (2003) notion of feature category decomposition in syntax and push it even further, showing how fine-grained feature structure has explanatory power in morphology as well. These rich ϕ-sets function as meaning targets that the morphology aims to hit, motivating a view of the Vocabulary Insertion operation as applying cyclically.

I also decompose the insertion operation into the sub-operations Rewrite and Split, the latter of which replaces Fission but has some key differences (§5.4.2). Insertion order is constrained by the same forces that determine hierarchical relations in ϕ-sets.

Underlying the discussion that follows is the general principle of Expressiveness, or the notion that the grammar operates not only to generate well-formed words and utterances but also to express target meanings as fully as possible. This principle follows from the intuition that the essential function of language is to communicate meaning and it has several consequences in the analysis presented below, most importantly the distinction between strong and weak exponence (§5.4.1) and the motivation

of multiple insertion when a target meaning is not fully expressed (§5.4.2).

5.2 The Structure of ϕ-Sets

It is common in linguistic analysis to specify agreement features as flat structures; for instance, a third person singular feminine argument might be specified simply as ϕ: [3, SG, F]. This type of flat structure is useful in that it captures the agreement patterns of many of the world's languages. However, a more fully articulated structure will be necessary in order to account for certain noncanonical patterns of agreement.

Hierarchical structure among the agreement feature categories was suggested by Noyer (1992), who proposed a Universal Feature Hierarchy in which person features are superior to number features, which are superior to gender/class features. Noyer's hierarchy played a key role in the work of Harley (1994) and Harley & Ritter (2000, 2002), who translated his feature hierarchy effects into geometrical markedness relations, deriving patterns of dependency, contrastiveness, and markedness without needing any additional rules or filters, and defining natural classes of features for grammatical rules to target. Section 5.2.1 discusses the patterns of DA in light of the feature hierarchy.

Recall from section 4.2.1 that Béjar's work on Context Sensitive Agreement (CSA), in which multiple arguments compete for a single morphological agreement slot, decomposed the feature categories of person and number into primitive features with entailment relations among them. She showed how a representation involving the fine structure of categories, in combination with certain assumption about the relevant syntactic operations, can predict both specific patterns of CSA and its general tendencies cross-linguistically. In 5.2.2 I push Béjar's approach into the morphology, demonstrating how fine-grained feature category decomposition also helps account for the patterns of discontinuous exponence.

Section 5.2.3 combines these two independently motivated ideas – that feature categories stand in hierarchical relations with one another and that categories themselves can be decomposed – to develop a rich, two-dimensional ϕ-set structure. Along one dimension are the fine-grained primitive features and entailments within feature categories, and on the other are hierarchical relations among the categories. In later sections I will demonstrate how a ϕ-set with both dimensions, viewed as a meaning target (§5.3.2), can both account for specific patterns of DA and explain its cross-linguistic tendencies (§5.4).

5.2.1 Relations among feature categories

At least since the work of Greenberg (1963) it has been noted that systematic relations hold among the agreement feature categories cross-linguistically. The dominance of

number over gender is established by several of Greenberg's implicational generalizations: 36, "If a language has the category of gender, it always has the category of number," 37, "A language never has more gender categories in nonsingular numbers than in the singular," and 45, "If there are any gender distinctions in the plural of the pronoun, there are some gender distinctions in the singular also". Later researchers noted a similar dominance of person over number as evidenced by patterns of neutralization of number in certain persons, for instance.

Noyer encodes these dominance relations in his Feature Hierarchy Hypothesis (Noyer 1992: 45 and passim), which states that there exists a universal hierarchy of morphosyntactic features (72) that constrains both the kinds of morphological rules that can exist and the order in which rules can apply. In combination with his assumptions about morphological operations, this hierarchy makes valid cross-linguistic predictions about morpheme insertion order, the direction of neutralization of feature distinctions, and, to a certain extent, the surface order of affixes.

(72) Universal hierarchy of morphosyntactic features (partial) (Noyer 1992)
person features > number features > gender features

With respect to Noyer's feature hierarchy it is important to note that it does not reflect any real organization of ϕ-features, which Noyer treats as an unstructured bundle, but rather a set of external constraints on how the grammar may interact with those features[1]. Harley (1994) and Harley & Ritter (2000, 2002, henceforth H&R) rejected the view of ϕ-feature sets as unstructured bundles, transforming some of the constraints encoded by Noyer's hierarchy into real structure represented as the morphological feature geometry in Figure 5.1 for referring expressions (agreement and other pronominal elements).

The feature geometry in Figure 5.1 encodes the dominance of number over gender directly by including gender (in H&R's terminology, "class") as a subnode of "individuation", which is used to represent number features. Since lower nodes are morphologically more marked than higher nodes, a referring expression must be specified for number in order to be specified for gender. The dominance of person over number is not, however, encoded in this feature geometry precisely because it represents morphological markedness only.

Another strand of research on the internal structure of ϕ-sets began with the work of Trommer (2002), discussed in section 4.2.2 above, who surveyed about 100 languages with subject agreement and observed that where the agreement prefix is split into separate marking of person and number there is a robust tendency for person marking to precede number marking. Harbour (2008) picks up on Trommer's generalization, arguing that an internal ϕ-set structure in which person dominates

[1]For instance, a rule banning a particular combination of person and number features must be specified to delete number, which is lower on the feature hierarchy, rather than person.

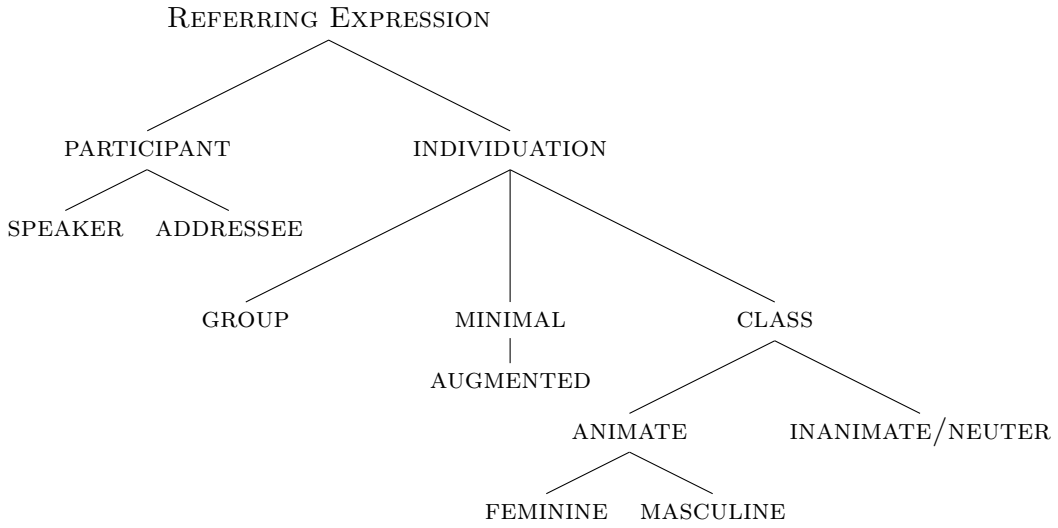
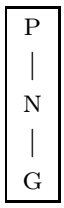


Figure 5.1: Feature geometry for person (Harley & Ritter: 2002)

number can, in combination with certain assumptions about morphological spellout, derive this ordering tendency. Like the work of H&R, this proposal uses hierarchical structure among agreement features to encode a portion of Noyer's feature hiearchy. However, Harbour departs from the work of H&R in an important way: in H&R's structure dominance is strictly morphological and represents relative unmarkedness; in Harbour's structure dominance is syntactic and encodes greater semantic abstractness (Harbour 2008: 195)[2].

I follow both H&R and Harbour in assuming that the well-established dominance of person over number and number over gender is encoded as structure internal to agreement feature bundles, and I adopt Harbour's particular view of dominance as encoding relative semantic abstractness. This structure forms the first of two dimensions in my ϕ-set representation and is shown on the vertical dimension as in (73), in which a feature higher in the structure dominates a feature lower in the structure.

(73) ϕ-set structure: Relations among feature categories



[2]Harbour supports his view of ϕ-set structure by pointing out the failure of H&R's structure to account for impure discontinuous agreement (Harbour 2008: 195–197).

5.2.2 Structure within feature categories

The second dimension of ϕ-set structure, within-category structure, represents the decomposition of agreement feature categories into a small set of primitive features. Noyer proposed to decompose the categories of person and number using the features given in (74). The various permutations of this feature set can encode all possible values of person used in language; for instance, third person is represented as [-I, -you, -participant], second person as [-I, + you, +participant], first person exclusive as [+I, -you, +participant], and first person inclusive as [+I, + you, +participant].

(74) a. Person values (Noyer 1992: 107–141)
- [± I]
- [± you]
- [± participant]

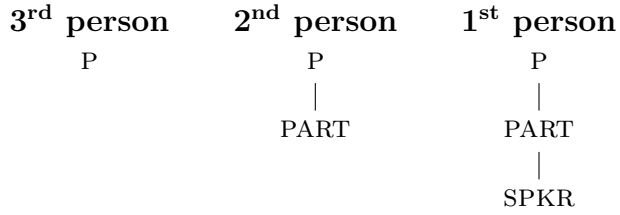
b. Number values (Noyer 1992: 141–196)
- [± singular]
- [dual]
- [trial]
- [quadral]
- [± augmented]

Harley (1994) and Harley & Ritter (2000, 2002) probe further the idea of feature category decomposition, suggesting that primitive features stand in a hierarchical feature geometry for referring expressions (Figure 5.1). These works translate the Feature hierarchy effects noted by Noyer into geometrical markedness relations, deriving patterns of dependency relations, contrastiveness, and markedness without needing any additional rules or filters, and defining natural classes of features for grammatical rules to target.

More recently, Béjar (Béjar 2003; Béjar & Rezac 2009) has adapted the feature geometry of Harley and Ritter, aiming to capture both underspecification within feature categories and the entailment relations between features, in her theory of Cyclic Agree (§4.2.1). Béjar proposes a structure in which the representation of each agreement feature category includes a root node (P, N, or G), which corresponds to an underspecified value of the category. For instance, the person root node [P] in the absence of any additional features is interpreted as the least specified person value: third person.

Béjar follows Noyer and Harley & Ritter in assuming the privative features [PART] (participant), [SPKR] (speaker), and [ADD] (addressee) (Béjar 2003: 44–50). A set of person features with PART as the highest level of specification would be interpreted as second person, and a set with all three features encodes first person.

(75) Encoding values of person (Béjar 2003)



Notice with respect to (75) that the feature set encoding a less specified value of person is a proper subset of the set encoding a more specified value of person[3]. In this way Béjar captures the logical relations among the different values of feature categories. For instance, for a referring expression to denote a speaker (first person) means that it also denotes a speech act participant; in feature notation [P [PART [SPKR]]] entails [P [PART]].

Béjar leverages her within-category structure to explain why a single probe can be valued by two arguments, thus deriving patterns of Context Sensitive Agreement in which multiple arguments compete for the control of a single morphological slot (I will return to her analysis in section 5.3.1). In my theory, within-category structure gives ϕ-sets greater descriptive and explanatory power as well. For instance, this structure grounds the distinction between strong and weak exponence and thus derives some of the patterns of Combinatorial DA (§5.4.1).

I assume the structures shown in Figure 5.2 for common values of person, number, and gender.

5.2.3 Two-dimensional ϕ-sets

The preceding sections have supported a view of sets of agreement features in which dominance relations hold between categories: gender is subordinate to number, and number is subordinate to person (§5.2.1). I have also followed Béjar in assuming that feature categories have a fine-grained structure with entailment relations among the individual features (§5.2.2). Combining these two dimensions of structure yields a complete ϕ-set representation as demonstrated in (76) for first person dual feminine.

[3]It is important to note that these entailment structures may vary crosslinguistically; the structures given in (75) are for a language in which first person is the most highly specified value, but this is not always the case. For a language in which second person is more highly specified the entailment structure for first person would be [P [PART]] and for second person [P [PART [ADD]]].

a. **Person**

1. Third

P

2. Second

P
|
PART

3. First

P
|
PART
|
SPKR

b. **Number**

1. Singular

N

2. Nonsingular

N
|
NSG

3. Dual

N
|
NSG
|
DU

c. **Gender**

1. Inanimate

G

2. Animate

G
|
ANIM

3. Feminine

G
|
ANIM
|
F

Figure 5.2: Encoding common values of person, number, and gender

(76) Feature structure for first person (Figure 5.2 a.3) dual (5.2 b.3) feminine (5.2 c.3)

P	—	PART	—	SPKR
|				
N	—	NSG	—	DU
|				
G	—	ANIM	—	F

The between-category structure is represented on the vertical dimension and within-category structure on the horizontal dimension. I follow Noyer in assuming that the same features are active in all components of the grammar: they are positioned by the syntax, they drive operations in the morphology, and they are interpreted in the semantics (Noyer 1992: 106). The ϕ-set structure that I have argued for in this section is thus intended to remain consistent across modules.

Previous work has established the universality (or at least strong cross-linguistic tendency) of some aspects of ϕ-set structure and the language-specificity of other aspects. On the horizontal (within-category) dimension, structure is always present because the feature categories can always be decomposed into primitive features with entailment relations among them (Harley & Ritter 2002; Béjar 2003). However, the exact form of the within-category structure may vary and a particular value may be

decomposed in different ways in different languages. For instance, in some languages first person is more highly specified than second person, but in other languages the reverse is true (see footnote 3 in chapter 4).

On the other hand, Béjar (2003) suggests that the vertical (between-category) dimension of structure is not always present. A language may not include all feature categories (e.g., many languages do not have gender agreement). Very rarely a language even divides the agreement features of a single argument across multiple heads (e.g., Béjar locates the person and number probes on two separate syntactic heads in Georgian). However, when between-category structure is present its form is invariant, as is well established by the robust cross-linguistic evidence for Noyer's Universal Feature Hierarchy.

The remaining sections in this chapter will demonstrate the power of these richly-structured ϕ-sets in the morphology. I will propose in section 5.3.2 that they serve as meaning targets that compel and constrain morphological operations. Section 5.4 discusses these operations in more detail, showing how ϕ-sets motivate a distinction between strong and weak exponence and support a cyclic view of insertion, just as they support a cyclic view of the syntactic operation Agree for Béjar (2003).

5.3 The Syntax of Agreement

I follow the standard assumption that relations between agreement controllers and targets are established in the syntax, adopting the particular implementation of Béjar (2003) and Béjar & Rezac (2009) introduced in section 4.2.1. Consider the following examples from Basque; in (77a) the direct object controls the preverbal agreement slot, while in (77b) the subject controls the same slot.

(77) Basque CSA (Béjar & Rezac 2009)

a. ikusi **z**-in-t-u-da-n
seen **2**-X-PL-have-1-PST
'I saw **you**.'

b. ikusi **n**-u-en
seen **1**-have-PST
'**I** saw him.'

Based on Basque and other languages with related patterns, Béjar and Rezac propose that Context Sensitive Agreement (CSA), which occurs when multiple arguments compete for the control of a single morphological agreement slot, can be derived by an articulated probe together with certain assumptions about the syntax of the Agree operation, as discussed in section 4.2.1. The fact that the articulated probe may be specified in different ways derives cross-linguistic variation in CSA patterns.

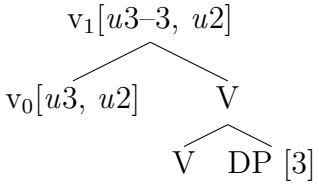
5.3.1 Cyclic Agree

Béjar and Rezac's theory of Cyclic Agree, makes three key assumptions. First, the feature categories on agreement probes have fine-grained structure. Second, there exist universal conditions on the Agree operation that allow a single agreement probe to be valued by two different arguments, giving rise to the phenomenon of CSA. And third, variation in CSA patterns across languages derives from the characteristic way in which agreement features are bundled and distributed in the syntax of each language. The first assumption was discussed in section 5.2.2; the latter two will be examined in turn below.

Valuing the probe

Following the analysis of Béjar & Rezac (2009), locality conditions derive a cross-linguistic preference for agreement with the internal argument. If, however, the internal argument cannot fully value the probe, a second cycle of agreement is licensed in which the probe may be further valued by the external argument. Agreement between a second person probe[4] and a third person internal argument is demonstrated in (78).

(78) Agreement with the internal argument on the first projection of v



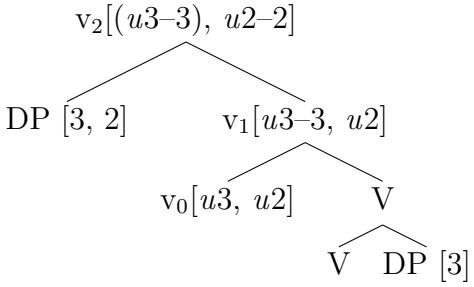
Notice that when the v head merges with the verb phrase its uninterpretable third person feature $u3$ is valued by the internal argument of the verb, as indicated by the notation "$u3$–3" in the highest-level projection of v. The internal argument only partially valued the v probe, however; the uninterpretable feature $u2$ remains unvalued, licensing a second cycle of Agree in which the v head projects to probe the external argument, as shown in (79).

[4]Béjar and Rezac use the notation [$u3$, $u2$] to represent a second person value. This is intended to represent a decomposed person category in which the PART feature (interpreted as second person in the absence of a SPKR feature, as discussed in section (5.2.2), and so represented here as $u2$) entails the root P feature (interpreted as third person in the absence of other person features, and so represented here as $u3$).

	Second Person	**First Person**
Type 1	[$u3$, $u2$]	[$u3$, $u2$, $u1$]
Type 2	[$u3$, $u1$, $u2$]	[$u3$, $u1$]

Table 5.1: Specifying person and number in two different language types

(79) Agreement with the external argument on the second projection of v



When the probe merges with the second person external argument its uninterpretable second person feature $u2$ is valued. The $u3$ feature and its corresponding value are enclosed in parentheses to indicate that the valuation happened on a lower projection of v. The probe is now fully valued, and agreement is complete.

The fine-grained structure of articulated ϕ-probes means that a single probe can agree with two arguments, as demonstrated above. The internal argument values the probe as fully as it can, but if the external argument can value the probe more fully than the internal argument it will end up controlling agreement.

Characteristic Probes

Alongside the universal conditions on the agreement operation, Béjar and Rezac propose a small number of possibilities for the language-specific specification of interpretable and uninterpretable phi features. For instance, the grammars of some languages treat first person as the most highly specified value of person; this is represented in the Type 1 row in Table 5.1. Other languages treat second person as more highly specified; this is shown in the Type 2 row.

These differences in specification give rise to some of the cross-linguistic differences in what have been described as Person Hierarchy effects[5]. For instance, in a Type 1 language a first person internal argument will fully value a second person probe but a second person argument will not fully value a first person probe; in the latter case, the probe will project and a first person external argument will end up controlling

[5]Béjar makes the strong claim that apparent Person Hierarchy effects can ultimately be reduced to a language's characteristic feature entailments and probe specification. I will show that this does not always hold true in a case study of agreement in Karuk in section 6.1 below

agreement. The reverse is true in a Type 2 language. The effect is that first person seems to outrank second person in Type 1 languages, while second person outranks first person in Type 2 languages.

Béjar (2003: 139ff) also observes that different languages distribute probes in the syntax in different ways. Low-ϕ languages exhibit a preference for the internal argument to control agreement for all feature categories; all probes are located on *v*. Split-ϕ languages locate one or more probes on v and one or more probes on T. For instance, in Georgian the internal argument is the preferred person controller (person probe on *v*) and the external argument is the preferred number controller (number probe on T). Double-ϕ languages have full sets of probes on both T and v to encode agreement with both subject and object, and Triple-ϕ languages include three full sets of probes.

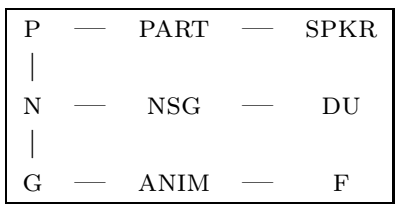
5.3.2 Meaning targets: ϕ-sets in syntax

I assume a meaning-driven view of grammar, in which the syntax creates meaning targets and the morphology attempts to come as close as possible to the target meaning at each cycle. This idea finds its genesis in the work of Caballero & Inkelas (to appear) on multiple exponence. Their main claim is that the cyclic optimization of a word relative to its meaning target can predict occurrences of multiple exponence, just in case subsequent expressions of a feature are optimizing with respect to form or meaning. In other words, they argue that if an "inner" morphological cycle expresses a feature weakly, an "outer" cycle will, if possible, express the same feature with a more informative morpheme.

Specifically, I propose that the meaning targets for agreement morphology are the agreement probes proposed in section 5.3.1. These originate on the *v* and T heads and, as discussed, receive their feature values from DP arguments via the Agree operation. The feature values of each probe are then copied onto agreement nodes early in the morphology; I will return to this process in section 5.4.

Each meaning target for agreement morphology takes the form of a complex ϕ-set (§5.2). For instance, recall the structure for a ϕ-set encoding first person dual feminine in (76), repeated here as (80).

(80) Feature structure for first person dual feminine

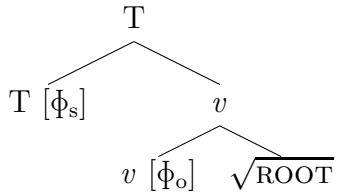


The morphology will strive to hit its meaning target by expressing every element of this complex structure. For instance, an agreement marker encoding first person will fully express the category of person, because it explicitly realizes the feature SPKR and implicitly realizes the features entailed by SPKR: PART and P. An agreement marker encoding nonsingular number, on the other hand, will only partially express the category of number because it realizes the feature NSG and its entailed feature N, but fails to realize the DU number feature.

5.3.3 Summary: Output of syntax

To summarize, the syntax determines the relations between the controllers and the targets of agreement according to the language-specific characteristic probe structure and distribution of probes on syntactic heads. The output of syntax is a tree as schematized in (81), in which ϕ_s represents the valued probe(s) on T (generally the features of the subject, unless the probe was valued on a second cycle of Agree) and ϕ_o represents the valued probes on v (generally the features of the object).

(81)



Each terminal node in (81) functions as a meaning target in the morphology. The ϕnodes have a two-dimensional internal structure, as discussed in section 5.2.3. For instance, in a clause with a first person singular feminine subject and a third person plural masculine object, the ϕ nodes would be specified as shown in (82).

(82) a. Internal structure of ϕ_s

P	—	PART	—	SPKR
|				
N				
|				
G	—	ANIM	—	F

b. Internal structure of ϕ_o

P				
|				
N	—	NSG	—	PL
|				
G	—	ANIM	—	M

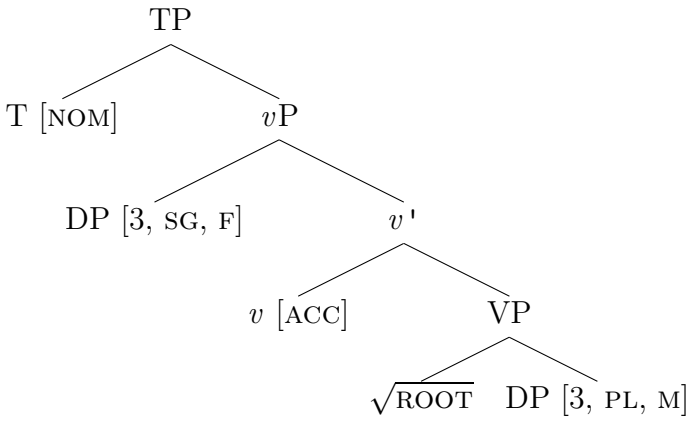
5.4 ϕ-Sets and the Morphology of Agreement

Certain patterns of agreement such as Context Sensitive Agreement (§4.2.1, §5.3) can be explained with reference to syntax alone. On the other hand, chapter 4 discussed the ways in which the syntactic theories reviewed therein are unable to make accurate predictions about DA. Section 4.3 explicitly argued that a purely syntactic

account of agreement cannot derive DA. This section treats DA as a morphosyntactic phenomenon, proposing a refined, internally complex version of vocabulary insertion that, in concert with the ϕ-set structure proposed above, derives both the specific patterns and the general tendencies of DA.

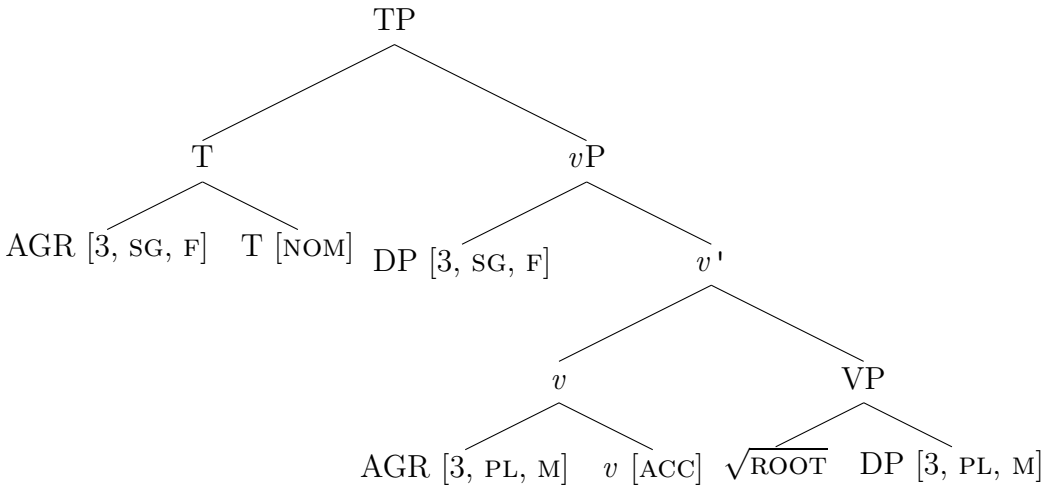
I will work within the framework of Distributed Morphology (DM), which proposes that "a single generative system is responsible both for word structure and phrase structure" (Embick & Noyer 2007: 290). Morphological operations apply directly to the output of syntax, altering morphosyntactic structure and adding phonological content. DM assumes two types of syntactic terminals: (i) abstract morphemes, which are bundles of grammatical features that do not contain any phonological specification; and (ii) roots, which are sequences of phonological features without any syntactic or semantic features. Example (83) shows the output of the syntactic derivation of a simple transitive sentence in which the T, *v*, and DP nodes are abstract morphemes.

(83) Transitive clause, terminal syntactic structure



I will assume the standard DM process of node insertion, which applies early in the morphology to adjoin morphological agreement nodes (AGR) to all syntactic Case-assigning nodes; that is, T and *v* in (84). Agreement features are copied from the argument nodes to the agreement nodes, creating an isomorphism between agreement in the morphology and Case assignment in the syntax. The application of this rule to (83) results in the morphosyntactic structure shown in (84). For the sake of simplicity, the Case features on T and *v* will henceforth not be shown.

(84) Transitive clause after Node Insertion



Another standard DM operation is Vocabulary Insertion, in which Vocabulary Items (VIs) are added to abstract morphemes, inserting phonological content to realize each node's feature bundle. The following sections will deal primarily with the question of exactly how this operation works. I will argue that insertion is motivated by the general grammatical principle of expressiveness, viewed here as the drive to hit a meaning target (§5.3.2). Section 5.4.1 builds on the standard framework by formalizing the notion of expressiveness, showing how a distinction between weak and strong exponence follows quite naturally from the ϕ-set structure proposed in section 5.2.3. I also depart from standard DM in rejecting the assumption that feature exponence necessarily involves feature discharge, and propose an alternative way to construe the blocking effects traditionally captured by feature discharge.

In section 5.4.2 I turn to the details of the vocabulary insertion operation. I show how the proposed ϕ-set structure conditions insertion order and argue that insertion applies cyclically, relative to a particular meaning target. I also propose that Insert is an internally complex operation composed of (at least) the sub-operations Rewrite and Split; the former introduces phonological content into abstract morphological nodes, and the latter creates additional positions of exponence under certain circumstances[6].

Finally, 5.4.3 deals with two types of blocking in DA. In node-level blocking a more specific VI can prevent the insertion of a more general one. Word-level blocking, on the other hand, involves preventing insertion across morphosyntactic nodes. The two types of blocking are given a unified explanation in terms of expressiveness relative to a meaning target.

[6]The Split operation proposed here refines and replaces the standard DM operation Fission (Embick & Noyer 2007: 314–318).

5.4.1 Feature exponence

I now turn to the notions of feature realization and exponence strength, both of which are central to the theory being developed. The sections herein view vocabulary insertion as being compelled by expressiveness relative to a meaning target and constrained by the content and internal structure of that meaning target, formalized as the complex ϕ-set containing the features of the agreement node being vocabularized.

Feature realization, not feature discharge

It was noted above that vocabulary insertion realizes an abstract morpheme by filling a morphological position with the phonological string that best realizes its morphosyntactic features (Embick & Noyer 2007: 297–299). A Vocabulary Item (VI), then, is a pairing of phonological and morphosyntactic information of the form given in (85), in which the left-hand side represents the exponent that is inserted in the context of the set of features given on the right-hand side.

(85) A schematic Vocabulary Item
xyz ↔ [F1, F2, F3]

It is generally accepted that only one exponent can be inserted into a terminal node in the default case (Embick & Noyer 2007: 298)[7]. The insertion operation is thus typically constrained to apply just once at a given node, and insertion is construed as a competition among all VIs that realize feature sets that match (are subsets of) the abstract morpheme. Once a VI is inserted, the features of the abstract morpheme realized by that VI are "discharged", or deactivated, and may not condition subsequent insertions. As an analytical device, feature discharge has been used to rule out redundant exponence. For instance, if a language offers a default set of agreement markers alongside a special set, used only in negative forms, say, then insertion of a negative agreement marker discharges the node's agreement features and thereby prevents insertion of the default agreement marker[8].

If feature exponence is feature discharge, then certain types of allomorphy are not predicted. For instance, in (86) the prefix and the suffix both apparently code second person so the insertion of the (more specific) suffix should discharge the person feature, leaving nothing to condition the insertion of the prefix. In order to account for this type of allomorphy, Noyer (1992: 14–18) introduced the notion of principal vs. secondary exponents.

[7]Embick and Noyer note that there are operations that allow for the insertion of multiple exponents into a single node; I will return to this point below.

[8]An appeal to feature discharge is only needed in the special cases in which multiple VIs are inserted into a single node, of course; otherwise the most specific VI rule would apply and insertion would end.

(86) t-aktub-iina
2-write-2SG.F
'You (f sg) write.' (Noyer 1992: 14) Classical Arabic

The idea is that person is a "principal exponent" only in the prefix, the insertion of which discharges the node's person feature permanently. After a feature is discharged it is available only as a "secondary exponent" to condition allomorphy in other VIs as, for instance, the contextual second person feature in the suffix[9]. Noyer captures the principal/secondary distinction by designating any secondary exponents as insertion context, as shown in (87). Person is a principal exponent in (87a) only; in (87b), person is in the set of features following the forward slash and is thus a secondary exponent.

(87) Principal vs. secondary exponence
(Noyer 1992: 15; notation changed for parallelism with the current work)
a. *t-* ↔ [2]
b. *-iina* ↔ [F] / [2]

One problem with this approach to allomorphy is that the distinction between primary and secondary exponence is not a principled one, and is in fact difficult to make in certain patterns. For instance, in a pattern like *P.N-V-N.G*, in which both the prefix and the suffix encode number and one other agreement feature, there are three possibilities for diagnosing number exponence: (i) N is the principal exponent of the prefix and a secondary exponent of the suffix, (ii) N is the principal exponent of the suffix and a secondary exponent of the affix, or (iii) N is a secondary exponent of both affixes. The pattern offers us no way to determine which analysis is the best.

Furthermore, it seems strange that a morphological rule would add phonological content to an abstract morpheme while treating the features of that morpheme as context; in other words, a single abstract entity (namely, a single ϕ-set) should not be both the target of a rule and the conditioning environment for that rule.

I take a different perspective on feature exponence by suggesting that as a morphological operation it is driven by the need to express the features of each node as fully as possible[10]. The insertion of a VI realizes the features of that VI, but does not discharge, delete, or render them inactive in any way. The general principle is that each VI must bring the node closer to its meaning target by realizing some as-yet-unrealized feature; that is, *insertion must make some semantic contribution*[11].

[9]This same distinction is discussed by Julien (2002) as exponence vs. "reflection" and is used to analyze away DA, a phenomenon that is problematic for Julien's theory.

[10]I will deal in more detail with the question of what it means to express a node's features fully in the section immediately below.

[11]This principle is sufficient for present purposes, but will need to be expanded in order to derive multiple exponence.

This offers a different way to rule out redundant morphology: once a VI has been inserted it will not be reinserted, not because the conditioning features have been discharged but because subsequent insertion of the same VI would not add anything new to the form. In other words, in order for multiple VIs to be inserted into a single node, their insertion rules must have distinct right-hand sides. For instance, the exponents in (88a) and (88b) may be inserted into a single node even though they both express F1, but (88a) will not be inserted twice into a single node. The restriction on redundancy is thus moved away from the level of individual features and to the level of feature bundles (VIs).

(88) a. *abc* ↔ [F1, F2]
b. *def* ↔ [F1, F3]

Once vocabulary insertion is viewed in this light, the principal/secondary distinction among coreferring agreement markers becomes unnecessary[12]. In agreement, exponence is exponence and the grammar must try to realize as many features of the target meaning as possible. This idea overlaps with the work of Béjar (2003) and Preminger (2011), who move away from an explanation of the obligatoriness of agreement in terms of feature checking. For Béjar and Preminger, the grammar must *try* to value uninterpretable features via Agree, but the derivation does not necessarily crash if those features cannot be checked. In other words, both in my approach to feature realization and in Béjar and Preminger's approach to syntactic agreement the emphasis is shifted away from the consequences of operations to the obligatoriness of the operations themselves.

Furthermore, a meaning-driven approach to exponence obviates an explicit restriction against the insertion of multiple VIs into a single node. In general there will be only a single insertion, because the situation in which multiple VIs match a single node – and they all make a semantic contribution – is rare. However, DA is quite naturally derived in languages with the right kind of lexicon.

Exponence strength

Section 5.2 discussed the internal structure of phi sets, proposing a structure in which person dominates number, which dominates gender, and in which these categories themselves have finer-grained structure based on feature entailments. For instance, consider the following example of number agreement in Hupa, an Athabaskan language traditionally spoken in Northern California.

[12]I do not necessarily wish to rule out non-local insertion contexts, which refer to the features of an abstract morpheme than the insertion site. For instance, a subject agreement VI may refer to tense as a contextual feature, in which case it might be desirable to say that the marker secondarily expresses tense.

(89) na:-**ya:**-s-**di**-l-to'n
PVB-**PL**-PFV-1**NSG**S-THM-jump
'we (PL) danced' (Sapir & Golla 2001)

In this language, all first person nonsingular arguments are marked with a nonsingular number marker (the inner number marker *di-*, where inner means closest to the verb root) while strict plural first person arguments, whose number must be greater than or equal to 3, are marked with an additional number marker (the outer number marker *ya:-*). A full ϕ-set encoding first person plural feminine[13] in Hupa is given in (90).

(90) ϕ-set structure for first person strict plural feminine in Hupa

P	—	PART	—	SPKR
|				
N	—	NSG	—	PL
|				
G	—	ANIM	—	F

The full meaning of this ϕ-set is yielded by the rightmost features; these represent the most narrowly precise specification of each feature category. It is in general these feature values that will trigger the insertion of particular VIs, but occasionally an insertion is conditioned by a more general, entailed value of one of the feature categories.

Returning to the observation that number is marked twice in (89), note that the outer prefix (*ya:-*) realizes the most precise specification of number – strict plural – while the inner prefix (*di-*) realizes the more general value of nonsingular. I assume that the first person plural agreement VIs in Hupa are as specified in (91). Leaving the implementation details aside until section 5.4.2, I turn now to the question of *why* these two VIs are inserted into a single agreement node.

(91) *ya:-* ↔ [PL]
di- ↔ [1, NSG]

In their work on multiple exponence, Caballero & Inkelas (to appear) note that one of its diachronic sources is "the weakening of an exponent to the point where a new, outer layer of morphology is required in order to robustly expone the relevant property." Within their Optimality Theory-based framework, then, multiple exponence is predicted when an inner agreement marker expresses an agreement feature

[13]Gender is not grammatically marked for first or second person arguments in Hupa, but it is marked for third person arguments so I assume that it is represented in the characteristic ϕ-set of the language, though nothing in my analysis hinges on this assumption.

weakly and an outer marker expresses the same feature with greater strength. For instance, a weak exponent might be an inner affix that does not mark a particular feature value transparently enough and so combines with a more transparent, productive outer affix to improve the form's parsability (Caballero & Inkelas to appear: 22–23).

The notion of exponence strength can be formalized with reference to the internal structure of the agreement node's ϕ-set, which represents the meaning target for vocabulary insertion at this node. Insertion of the inner agreement marker *di-* in (89), which encodes nonsingularity, encodes number weakly because NSG is an entailed feature in the meaning target in (90).

(92) Weak exponence = partial expression of the category's meaning target

a. Partial meaning target (number category only)

N	—	**NSG**	—	PL

b. *di-* ↔ [1, **NSG**]

Insertion of the outer agreement marker *ya:-*, on the other hand, encodes number strongly because the feature PL is not entailed by any other feature in the meaning target. This insertion fully expresses the category of number for this agreement node.

(93) Strong exponence = full expression of the category's meaning target

a. Partial meaning target (number category only)

N	—	NSG	—	**PL**

b. *ya:-* ↔ [**PL**]

The idea of exponence strength thus helps us to understand how certain patterns of DA are licensed. Just in case an earlier insertion realizes a feature category weakly, a later insertion is motivated if it expresses the same feature category with greater strength. In addition, a distinction between weak and strong exponence follows naturally from the ϕ-set structure proposed in section 5.2.3 because each feature category is decomposed into multiple feature values with entailment relations among them. A VI weakly expresses a feature category if it realizes an entailed member of that category's internal structure in the meaning target.

5.4.2 Cyclic Insertion

Perhaps the most striking difference between DA and more canonical forms of agreement (chapter 3) is that in DA the features of a single agreement node are realized by multiple morphs. The standard DM operation by which multiple exponents are inserted into a single terminal node is Fission, which has been defined in the literature in different ways. As conceived by Noyer (1992), Halle (1997), and Embick & Noyer



Figure 5.3: Fission (Noyer, Halle, Embick & Noyer)

$$\mathrm{C}\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ \dots \\ F_m \end{bmatrix} \rightarrow \begin{bmatrix} F_1 \\ F_3 \\ \dots \\ F_m \end{bmatrix} \begin{bmatrix} F_2 \\ F_3 \\ \dots \\ F_m \end{bmatrix}$$

Figure 5.4: Fission (Arregi & Nevins)

(2007), certain morphemes are specified as able to undergo fission. When insertion into such a node discharges only some of the node's features, the remaining features split off to form a subsidiary morpheme, into which another item can be inserted. For instance, the first insertion in Figure 5.3 discharges F1 and F2 leaving unexpressed F3, which triggers Fission and the subsequent insertion of another VI.

Arregi & Nevins (2012) offer another take on Fission. For them, fission between features F1 and F2 splits a morpheme of category C containing F1 and F2 as shown in Figure 5.4.

This happens prior to insertion and is the effect of a language-specific rule that references certain feature combinations. For instance, rule (94) in Basque splits features in second and third person plural clitics:

(94) Basque Plural Fission: Clitic, [-author], [-singular][14]

In this section I will propose the morphological operation Split which, based on evidence from the patterns of DA, substantially refines and replaces Fission. In the present theory, Split is conceived as a sub-part of the vocabulary insertion operation that is obligatorily invoked under certain circumstances. Recall that Béjar's theory of Cyclic Agreement (§4.2.1, §5.3.1) assumes that when a goal fails to fully value a probe in syntactic agreement, a second cycle of the syntactic operation Agree is licensed. I pursue an analogous idea in the morphology: if an insertion fails to fully express a node's meaning target, a second cycle of the morphological operation Insert is licensed. The intuition at the heart of both ideas is that if the grammar doesn't get it fully right the first time, it can and must try again.

[14]These feature specifications are Arregi & Nevins'. The person value [-author] corresponds to my [P – PART] and the number value [-singular] corresponds to my [N – NSG].

Specifically, I propose to decompose the insertion operation as shown in (95). In the first step, Rewrite, a Vocabulary Item is added to the node (I will turn immediately below to a discussion of the order in which VIs are added). If the first step expresses the node's meaning target as fully as possible then the Insert operation terminates, otherwise the second step splits off a new position of exponence, licensing a subsequent application of Insert.

(95) Insert

a. Rewrite

b. If target meaning not fully expressed, Split

Conditioning insertion order

I turn now to a consideration of the order in which agreement morphs are inserted. Although the languages I surveyed in chapter 2 involve a number of distinct ordering challenges and a complete account of morpheme order is beyond the scope of the present work, it is possible within the framework of my analysis to capture and make predictions about the relative ordering of agreement markers. The hierarchical arrangement of features in ϕ-sets will ensure that the feature-realizing morphs are inserted in a particular order (person first, then number, then gender), which is reflected as a layered structure within the complex agreement node (person on top, number in the middle, gender at the bottom). I will show in section 6.2.3 how this structure biases toward certain linear orderings of agreement markers.

Noyer (1992) noted a strong cross-linguistic tendency for agreement features higher on his Feature Hierarchy (72) to be inserted before features lower on the hierarchy. His hypothesis is that there are two conditions on insertion order.

(96) Noyer's Spell-Out Ordering Hypothesis (Noyer 1992: 44–45)

a. Panini's Principle: If one rule's structural description is contained in the other's, the rule with the more specific structural description applies first.

b. Feature Hierarchy: If the structural descriptions are disjoint or overlapping, then the rule referring to the hierarchically higher feature applies first.

Because the Feature Hierarchy is reified in my theory as dominance relations among feature categories in the complex ϕ-set, insertion order can be explained in terms of those relations. Information higher in the Feature Hierarchy (and likewise higher in ϕ-set structure) is more "central"; exponents of higher properties will never lose out to lower properties in the competition for expression (Noyer 1992: 45).

The general principle is to encode the most important information first. Specifically, vocabulary insertion proceeds top-to-bottom through the ϕ-set. Within each

category, the best-matching VI (the maximally specified VI that is not more specified than the meaning target) wins the competition for expression. For instance, recall from section 5.4.1 the example of DA in Hupa, repeated here as (112), the corresponding AGR node's ϕ-set, repeated here as (98), and the matching VIs, repeated here as (99).

(97) naː-yaː-s-di-l-to'n
PVB-PL-PFV-1NSGS-THM-jump
'we (PL) danced' (Sapir & Golla 2001)

(98) ϕ-set structure for first person strict plural feminine in Hupa

P	—	PART	—	SPKR
|				
N	—	NSG	—	PL
|				
G	—	ANIM	—	F

(99) *yaː-* ↔ [PL]
di- ↔ [1, NSG]

Searching from top to bottom in the ϕ-set, notice that the matching VI that encodes the structurally highest category is *di-*, the first person nonsingular subject marker. Because this VI realizes the information that is, by assumption, the most important, it will be inserted first by condition (96b). In effect, this condition rules out insertion of the plural *yaː-* before the first person nonsingular *di-*.

Alongside the VIs in (99) exists the first person singular subject marker, which is unspecified for number, shown in (100).

(100) *wh-* ↔ [1]

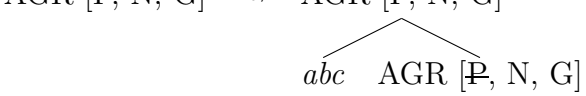
Notice that this VI also matches the ϕ-set structure in (98). However, condition (96a) ensures that *di-* will be inserted before *wh-* because it is the more highly specified VI. Section 5.4.3 below deals with the question of how the subsequent insertion of *wh-* is blocked.

Split

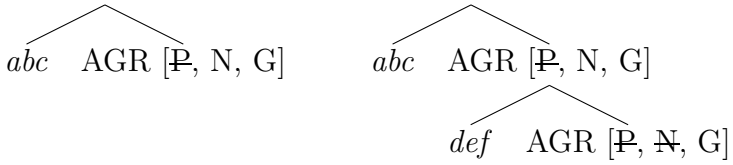
I now turn to a discussion of the implementation of Split, differentiating it from the more standard Fission operation, which it is intended to replace. Taking seriously the idea that the morphological grammar is driven to fully express meaning targets (§5.3.2), I propose that Split is invoked by Insert just in case the item being inserted does not expone all of the features of an agreement node. This operation is similar

to Fission (Embick & Noyer 2007: 314–318) in that it creates a new position of exponence, but it differs in several ways to be discussed below. The basic function of the operation is to split a node with n positions of exponence into one with $n+1$ positions of exponence as schematized for an agreement node in (101)[15], in which each squiggly arrow indicates an application of Insert with Split.

(101) a. First Split triggered by insertion of P
VI 1: *abc* ↔ [P]



b. Second Split triggered by insertion of N
VI 2: *def* ↔ [N]



Each terminal node in (101) filled by a string indicates a position of exponence that has been realized by the Insert operation[16]. The first insertion, shown in (101a), adds the string *abc*, which realizes the appropriate person feature. Because the number and gender features of the node are left unrealized, the insertion operation invokes Split to create the structure on the right-hand side of (101a). The agreement node now has two positions of exponence, one of which is filled by the inserted string *abc* and one of which is empty. The person feature is crossed out in the terminal ϕ-set, indicating that it has been realized (but not that it has been discharged, as discussed above).

The process is repeated for an insertion realizing the number feature in (101b). Note that after each application of Split, the empty position of exponence is still specified for all of the features of the agreement node. This insures that all subsequent insertions have access to the full set of features specified on the original node; although these features have already been realized (indicated by strikethrough), they are still available to condition insertion.

I mentioned above that Split is similar to Fission in some ways and is intended to replace that operation. Fission, as conceived by Noyer (1992), Halle (1997), and

[15]In order to simplify the trees, the examples in this section collapse the fine-grained structure within feature categories in ϕ-sets. This should be understood as shorthand for the full ϕ-set structure discussed in section 5.2.

[16]The details of the Insert operation will be discussed in the section immediately below.

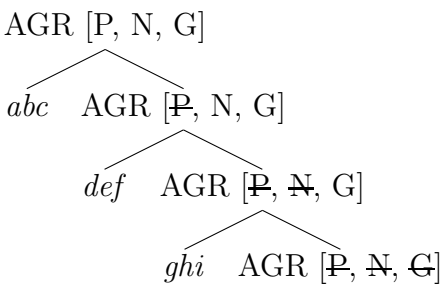
Embick & Noyer (2007), is licensed when certain morphemes are specified to split into multiple positions of exponence. When insertion discharges only some of the features of a node specified for Fission, the remaining features split off to form a subsidiary morpheme into which another item can be inserted. This was schematized in Figure 5.3 above. I will refer to this view of Fission as "Fission-NHE" (where the abbreviation NHE refers to the first authors of the above-cited works in chronological order). As conceived by Arregi & Nevins (2012), Fission rules are language-specific and apply early in the morphology, prior to Vocabulary Insertion, to split nodes with certain feature combinations into multiple positions of exponence. Their view was schematized in Figure 5.4 above, and I will henceforth refer to it as "Fission-AN".

I have chosen to call my version of the Fission-like operation by a new name, Split, to highlight the fact that it differs from both Fission-NHE and Fission-AN in several important ways. The remainder of this section explains three major conceptual differences between Split and Fission, the latter of which has to do with DA patterns that are not straightforwardly captured by previous accounts of the Fission operation.

The first conceptual difference between Fission and Split is that in the present theory Split is a subpart of Insert. This is in the spirit of Fission-NHE, which assumes that insertion automatically triggers Fission if the inserted VI fails to fully realize the node's features. I likewise assume that Split applies automatically when Insert does not fully realize the node, but I reject the idea from Fission-NHE that a morpheme must be specified as able to split in order for Fission to occur. I propose that Split applies anywhere there is partial feature realization – the question of whether there is subsequent insertion depends on the features of the node and on the lexicon of the particular language. Incorporating Split into the insertion operation is a radical departure from Fission-AN, which assumes that Fission applies to alter morphological structure before vocabulary insertion.

The second difference has to do with the way that Split builds structure. Fission-NHE assumes that a flat structure results from Fission, specifically that "Vocabulary Items accrete on the sister of the fissioned morpheme until all Vocabulary Items which can be inserted have been" (Noyer 1992). Fission-AN does not attempt to define the possible structures that can be generated by Fission, but since the rule itself is language-specific so, I assume, is the resultant structure. Arguing against both approaches are strong cross-linguistic ordering tendencies among agreement markers in DA. For instance, a robust generalization was observed by Trommer (2002) that when person and number are marked by distinct morphs, person marking precedes number marking.

I assume, following Harbour (2008), that these cross-linguistic regularities derive from the consistent, hierarchical internal structure of ϕ-sets. Harbour proposes a particular view of the linearization operation that, in combination with a ϕ-set in which person dominates number, derives the P-before-N tendency. I take this idea one step further, suggesting that it is the internal structure of the agreement node



abc ↔ P
def ↔ N
ghi ↔ G

Figure 5.5: Complex agreement node generated by Split

(which reflects the internal structure of the ϕ-set) that creates a morphosyntactic bias toward a particular linear order, just as morphosyntactic structure partially determines constitutent order in the sentence.

The Split operation adds a morphosyntactic layer with each application, as shown in (101), and thus creates a complex agreement node with hierarchical internal structure, schematized in Figure 5.5. Recall from the previous section that the order of insertion of agreement VIs is determined by the internal structure of ϕ-sets: person exponence comes first, followed by number exponence then gender exponence. The fact that Split is invoked by Insert thus has an important consequence: the internal structure of agreement nodes will also, in general, be consistent with ϕ-set structure because morphemes vocabularized earlier are hierarchically superior to morphemes vocabularized later. The top layer(s) of the complex agreement node will be realized by person-encoding strings, the middle layer by number-encoding strings, and the bottom layer by gender-encoding strings. What this means for morphological structure is that the person-realizing VI is always superior to the number-realizing VI in the complex AGR node.

As a consequence and depending on the particular implementation of the linearization operation, the hierarchical relations among the layers of the ϕ-set are reflected as linear order tendencies in the generated form. I will return to this point in chapter 6.

The third major conceptual difference between Split and Fission has to do with the way that features are propagated from the original, splitting node to the new terminal nodes. Fission-NHE (Figure 5.3) assumes that only the unrealized features of the parent node are inherited by any fissioned child nodes. Fission-AN (Figure 5.4), on the other hand, argues that the two features stipulated to split slot into different child nodes, but that all other features of the parent node are inherited by both child nodes.

Unfortunately, neither view of Fission works well for impure patterns of DA like the one shown in (102), given my assumptions about insertion order. Under Fission-

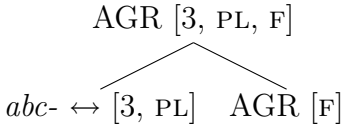
NHE features are discharged as soon as they are expressed, meaning that they are no longer available to trigger the insertion of VIs. Feature discharge is problematic for patterns like (102) because once the exponent fusing person and number is inserted, no feature remains to trigger the insertion of the discontinuous number exponent.

(102) Impure DA
P.N-V-N

Let me be more explicit about why Fission-NHE fails on this pattern. Examples of VI insertion rules for the two affixes are given in (103), and in (104) I consider the Fission operation of Noyer, et al. By assumption, the person-encoding VI 1 (which also encodes number) is inserted first, leaving undischarged only the gender feature. Because this feature cannot trigger the insertion of VI 2, this formulation fails to generate the correct output form.

(103) VI 1: *abc-* ↔ [3, PL]
VI 2: *-xyz* ↔ [PL]

(104) Fission-NHE (to be rejected)

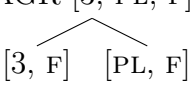


Under Arregi & Nevins' view, the clean split between features is problematic for impure DA. In a language with a person-number discontinuity, the morphological fission rule would split an agreement node [PNG] into two nodes, [PG] and [NG]. Since neither of the output nodes is specified for both person and number, insertion of the prefix in (102), which fuses person and number, is not straightforward.

I can again use the example VIs in (103) to demonstrate Arregi & Nevins' Fission operation, which represents a rule that applies prior to Vocabulary Insertion. Since this discontinuity involves a split between person and number I appeal to the morphological rule in (105a) to generate the structure shown in (105b). Because neither of the output terminal nodes is specified for both person and number, there is no insertion site for VI 1.

(105) a. $\text{AGR} \begin{bmatrix} P \\ N \\ G \end{bmatrix} \rightarrow \begin{bmatrix} P \\ G \end{bmatrix} \begin{bmatrix} N \\ G \end{bmatrix}$

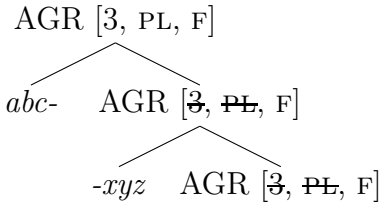
b. Fission-AN (to be rejected)



Both of the Fission systems could be made to generate a pattern like (102) by appealing to local insertion contexts, but my theory disallows this for the reasons discussed in section 5.4.1. Additionally, an assumption that an insertion context can reference the morphological features of another node would have to be added to Arregi & Nevins' analysis.

Example (106) shows the structure generated by the Split operation under the analysis currently being developed. By assumption, the person-encoding VI 1 will be inserted first, triggering an application of Split. Next, the number-encoding VI 2 is inserted. Split is triggered by this insertion, as well, but assuming that the language doesn't include a VI that expresses feminine gender the lowest AGR node will remain unmatched.

(106) Split



Patterns like (102) show us that Split is descriptively superior to both of the views of Fission discussed above. Because the features on a terminal node resulting from Split will always be a superset of those on a terminal node resulting from Fission (compare the terminal AGR node in (106) against those in (104) and (105b), recalling that struck-through features in my formalism are not discharged and can still condition insertion), Split can condition the insertion of any VI that Fission could. In addition, Split can easily derive patterns that could not be captured straightforwardly by Fission-NHE or Fission-AN.

The careful reader will notice something odd about how the present theory handles (106). I proposed above that Split is triggered by Insert in case there are features on the node that have not yet been realized. After the first insertion in (106), which expones person and number, only the gender feature remains unexpressed. Split creates a new position of exponence with all of the features of the parent node, into which is inserted a string that again realizes number. Although nothing in the theory rules this out, it's odd that an unexpressed gender feature should drive the subsequent insertion of a number feature.

In their work on multiple exponence, Inkelas & Caballero offer a promising line of attack on this problem by taking a more nuanced view of the distinction between weak and strong exponence. For instance, a weak exponent might be an inner affix that does not mark a particular feature value transparently enough and so combines with a more transparent, productive outer affix to improve the form's parsability (Caballero & Inkelas to appear: 22–23). Alternatively, some apparently superfluous insertions

may make a non-semantic contribution to the word, increasing their well-formedness by spelling out structural positions that are required for independent reasons such as improving phonological stem shape or moving a stem toward wordhood (Caballero & Inkelas to appear: 12–16).

By this logic, although both vocabulary items inserted in (106) appear to realize the same value of number, it may be the case that *abc-* is not fully transparent, productive, or parsable, and that this subtler form of weak exponence is what licenses the subsequent number-realizing insertion. Or we may find, given additional evidence from the language, that the second number insertion fills some position that must be spelled out for grammaticality.

In summary, I have proposed that Split is triggered in (at least) three different situations; these are listed in (107). By appealing to the notions of expressiveness and exponence strength I have fully captured situations (107a) and (107b); the exact factors at play in situation (107c) are complex, non-semantic, and may vary from language to language.

(107) What compels Split?

a. Different VIs express different feature categories of the target meaning; earlier insertions leave some categories unrealized

b. Different VIs express a feature category of the target meaning with different strengths; earlier insertions express categories weakly (some impure DA patterns)

c. Other: language-specific well-formedness factors, etc. (remaining impure DA patterns and ME)

How Cyclic Insertion generates different types of discontinuity

I suggested above that if the first vocabulary insertion at a node fails to fully express that nodes meaning target, a second cycle of the morphological operation Insert is licensed. I also proposed the decomposed Insert operation in (95), in which the first step introduces the best-matching VI into the abstract node and the second step, if applicable, creates a new position of exponence, thereby licensing the next application of Insert.

Let's look at a simple example. Recall from chapter 2 that coreferential DE patterns arise when a set of feature categories that can be expected to be bundled on a single node in the syntax (namely the agreement features person, number, and gender) is expressed by distinct morphemes. A pure coreferential discontinuity can be thought of as resulting from a cleanly split feature set. If a set of agreement features in the syntax splits and each feature maps to at most one morpheme[17], the

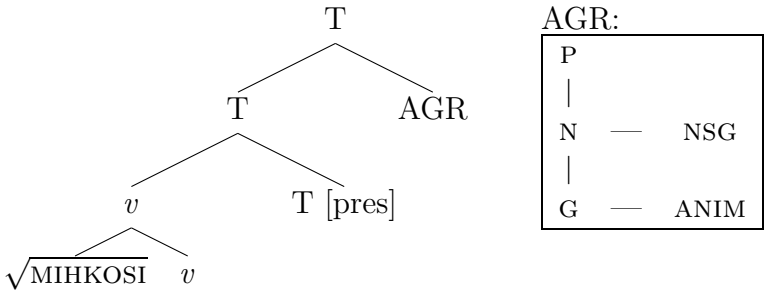
[17]Note that multiple features may map to one morph without creating impurity. In other words,

discontinuity is pure as in (108), in which the inner suffix marks person and the outer suffix marks number and gender.

(108) mihkosi-w-ak
be.red-3-PL.ANIM
'they (anim) are red' (Dahlstrom 1986: 16) Cree

The full meaning target for this form is shown in (109). The terminal nodes in the tree are abstract morphemes, and the internal structure of the AGR node is specified at the right.

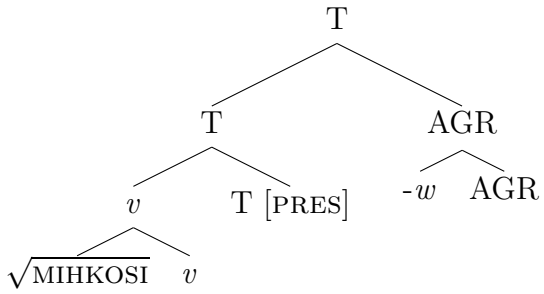
(109) Meaning target for (108)



The derivation begins with the insertion of the root VI. Next, vocabulary insertion applies to the v and T nodes with no effect; there are no features to express on v and present tense is not marked in Cree. Finally, vocabulary insertion applies to the AGR node. As discussed in the immediately preceding section, the person-encoding VI matches first. The rewrite step adds the string -w to the form, fully realizing the [P] category of the meaning target. The resulting structure is shown in (110); strikethrough of features in the ϕ-set indicates feature realization, but crucially not feature discharge or deletion (§5.4.1).

pure coreference may involve a many-to-one mapping from syntax to morphology, but not a one-to-many mapping.

(110)

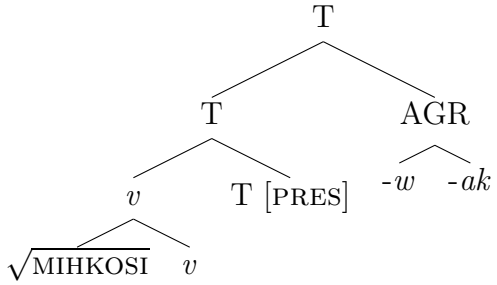


AGR:

~~P~~		
N	—	NSG
G	—	ANIM

At this point in the derivation the first cycle of vocabulary insertion at the AGR node is complete, but unrealized features (the number and gender features) of that node's meaning target remain unexpressed. These unrealized features trigger splitting of the AGR node, creating a new position of exponence and licensing a second cycle of Insert. The number- and gender-encoding VI matches next; Rewrite adds the string *-ak* and the meaning target is fully exhausted. The resulting structure is shown in (111).

(111)



AGR:

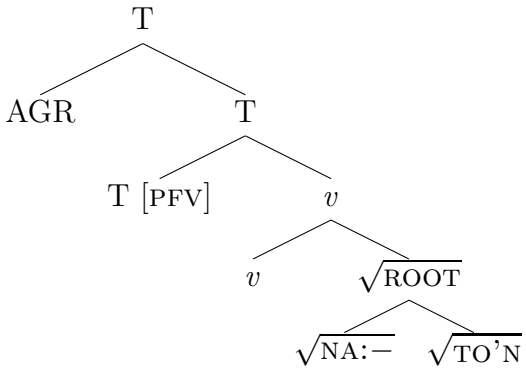
~~P~~		
~~N~~	—	~~NSG~~
~~G~~	—	~~ANIM~~

In this way, insertion applies cyclically at each morphosyntactic node until the node's meaning target is expressed as fully as possible.

Let's look next at another type of DA. Combinatorial discontinuities differ from coreferential discontinuities in the following way: whereas coreferential discontinuities involve the splitting of a set of feature categories into multiple morphemes, in combinatorial DA a single feature category is realized by more than one morph, each of which expresses a different value. For instance, recall the Hupa example from section (5.4.1), repeated here as (112), in which two distinct prefixes encode subject number. The inner prefix, *di-*, encodes nonsingular number. The outer prefix, *ya:-*, encodes strict plurality ($\geq$3). The full meaning target is given in (113).

(112) naː-yaː-s-di-l-to'n
PVB-PL-PFV-1NSGS-THM-jump
'we (PL) danced' (Sapir & Golla 2001)

(113) Meaning target for (112)

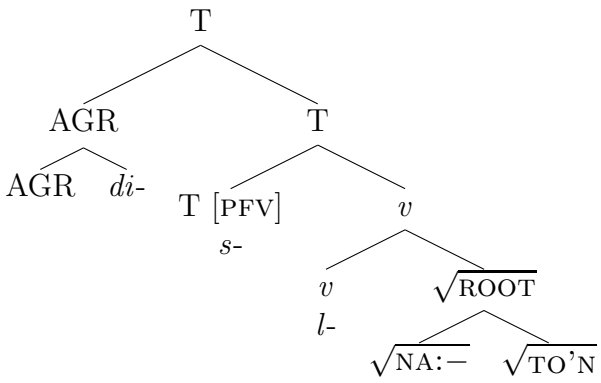


AGR:

P	—	PART	—	SPKR
|				
N	—	NSG	—	PL
|				
G	—	ANIM	—	F

The derivation of this form also begins with the insertion of the root; in this verb form the root is complex and made up of two VIs, the preverb and the main verb root. Next, vocabulary insertion applies to the *v* and T nodes, adding the strings *l-* and *s-*. The first cycle of insertion at the AGR node adds the string *di-* to the form, fully realizing the person category of the meaning target and partially realizing the number category. The resulting structure is shown in (114).

(114)

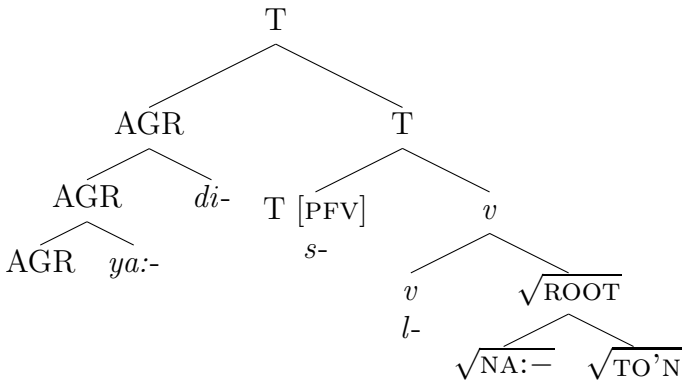


AGR:

~~P~~	—	~~PART~~	—	~~SPKR~~
|				
~~N~~	—	~~NSG~~	—	PL
|				
G	—	ANIM	—	F

Again, the unrealized features of the meaning target trigger splitting of the AGR node, creating a new position of exponence and licensing a second cycle of Insert. The discontinuous number VI matches next; Rewrite adds the string *ya:*-, leaving only the gender features unexpressed. Split applies, and a third cycle of Insert is licensed. The resulting structure is shown in (115).

(115)



AGR:

~~P~~	—	~~PART~~	—	~~SPKR~~
|				
~~N~~	—	~~NSG~~	—	~~PL~~
|				
G	—	ANIM	—	F

Although the gender category of the meaning target has not been expressed, the grammar has come as close as it can to hitting the meaning target (gender is not

marked in first person forms in Hupa). Thus, although there exists a terminal AGR node that could host an agreement marker, there is no potential match for insertion into that slot and insertion terminates at this node. The fact that gender remains unexpressed is not a problem under the assumptions of the current analysis; I have rejected the notion of feature discharge and the idea that nodes must be fully discharged in order for the derivation to be grammatical.

An important consequence of a cyclic view of vocabulary insertion is that it pushes the DM framework toward more naturally deriving coreferential DA. The creation of additional positions of exponence is not a result of some diacritic on the node or language-specific morphological rule, but rather it happens obligatorily because insertion operates relative to a meaning target. Coreferential DA arises when an earlier insertion leaves some category of the meaning target unexpressed, licensing a subsequent insertion to realize the remaining categories with additional morphs. Combinatorial DA arises when an inner affix expresses a category weakly (it only realizes part of the meaning target), licensing a subsequent insertion to realize the same category more fully.

5.4.3 Blocking insertion

Although my proposals move DM in the direction of deriving DA more naturally, it is important to retain the ability to block unnecessarily redundant or truly superfluous morphology. In early verions of DM, redundant insertion was blocked when multiple VIs competed for a single morphological slot; since only one VI could win, the other matching VIs would never be inserted. With the addition of Fission, blocking was accomplished by appealing to feature discharge. Features of the agreement node that had already been realized features were thought to be grammatically inactive, and thus unable to condition another insertion (Harley & Noyer 1999).

The present theory rejects the idea of feature discharge per se, but redundant insertion can still be blocked. Consider first an example in which a more specific VI blocks the insertion of a more general one. Tamazight Berber has both a default first person agreement suffix (116a) and a special prefix fusing person and number that is used in first person plural forms (116b). Relevant to the discussion here, the first person suffix never appears in first person plural forms, as shown by the ungrammaticality of (116c).

(116) Blocking 1: Specific blocks general

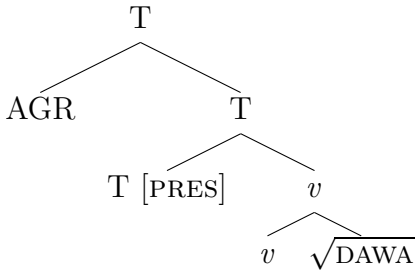
a. dawa-ɣ
cure-1
'I cure'
(Abdel-Massih 1971) Tamazight Berber

b. n-dawa
1PL-cure
'We cure'
(Abdel-Massih 1971) Tamazight Berber

c. * n-dawa-ɣ
1PL-cure-1

Taking seriously the idea that insertion is driven by expressiveness, the drive to fully express a meaning target, the intuition might be that insertion should not apply if it doesn't contribute to the meaning of the word. In keeping with this notion I proposed above that insertion must realize an as-yet-unrealized feature of the meaning target. The meaning target for a first person plural verb in Berber is given in (117).

(117) Meaning target for (116b)

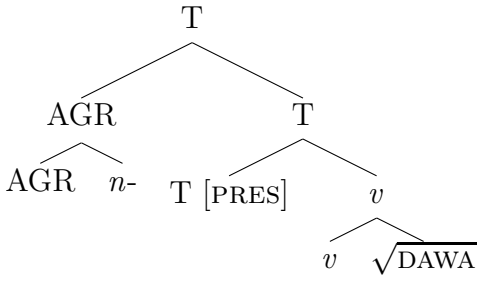


AGR:

P	—	PART	—	SPKR
N	—	NSG	—	PL
G	—	ANIM		

The derivation begins with the insertion of the verb root. Next, vocabulary insertion applies to the v and T nodes (with no effect). The first cycle of Insert at the AGR node adds the string n-[18], which realizes both person and number features of the meaning target. The unrealized gender feature triggers Split, and the resulting structure is shown in (118).

(118)



AGR:

~~P~~	—	~~PART~~	—	~~SPKR~~
~~N~~	—	~~NSG~~	—	~~PL~~
(G	—	ANIM)		

[18] Although both n- and -ɣ can fully value the person features of the meaning target and are thus eligible for first insertion, n- is additionally specified for number. Because it is more highly specified, it wins the competition for insertion by the Subset Principle.

At this point there remains one VI that matches the meaning target: -ɣ ↔ [1]. However, because the category of person has already been fully realized, this VI does not realize any feature that has not already been expressed. In other words, it does not move the meaning of the form closer to the target meaning at this node. Therefore, although there is an empty position of exponence and a matching VI, the second cycle of Insert is blocked.

Let's look at ambiguous DA for a more complicated example of blocking. Chapter 2 discussed patterns of DE that give rise to ambiguous interpretations (§2.3.4). Ambiguity sometimes arises when an agreement feature splits away from the other features in its set, as demonstrated by the various interpretations impossible for the following examples.

(119) a꞉= **ya꞉-xo-ɬ-ch'i**-de꞉-ne'
thus= **PL-3.A.O**-COM-**3.A.S**-THM-say
'he said to them' / 'they said to him' / 'they said to them'
(Sapir & Golla 2001) Hupa

(120) mo-g-k'al-i-t < mo-**v-g**-k'al-i-**t**
PVB-**1.S-2.O**-kill-THM-**PL**
'I kill you (pl.)' / 'we kill you (sg.)' / 'we kill you (pl.)'
(Hewitt 1995) Georgian

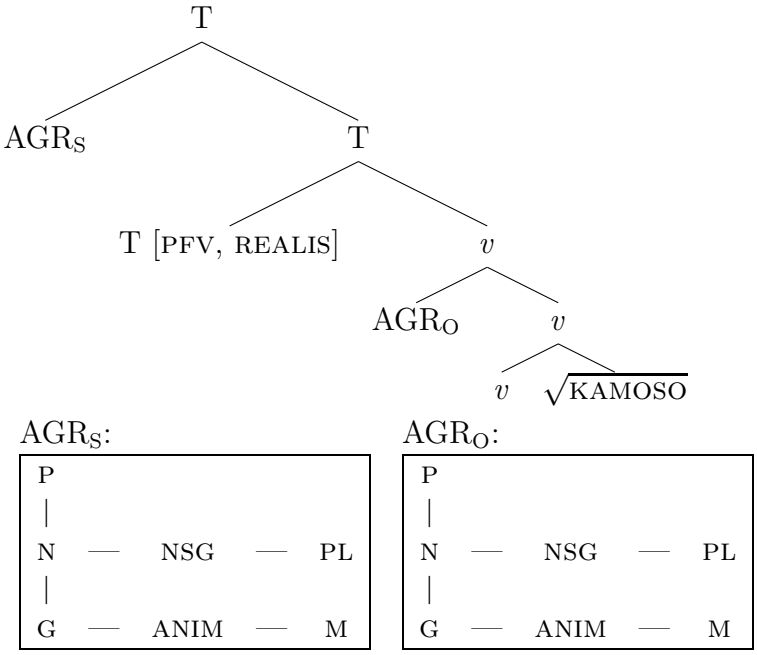
(121) **i**-kamoso-**hig**-ak-e-**ri**
3M.S-visit-**PL**-PFV-REALIS-**3M.O**
'he visited them' / 'they visited him' / 'they visited them'
(Michael 2008) Nanti

All ambiguous examples in the results of my typological study involve the expression of number discontinuously from the person features of the corresponding argument. The examples also all code more than one argument on the verb. In these patterns, the discontinuous number marker can be interpreted as cross-referencing one argument or the other, or both. Crucially, in each case there is only one plural marker even if both arguments are plural.

Ambiguity occurs when a single form can encode more than one underlying constellation of meaning targets. For instance, (121) may result from a meaning target in which the subject agreement node is plural and the object is singular, one in which the subject agreement node is singular and the object is plural, or one in which both agreement nodes are plural. The latter case must involve blocking, since there is only one plural marker at the surface (otherwise the pattern would not be ambiguous).

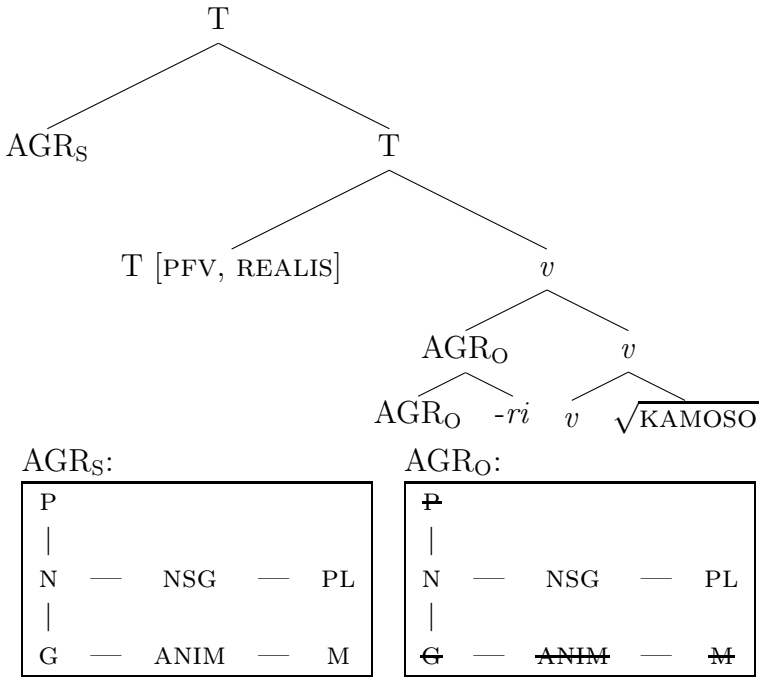
The meaning target for an ambiguous example, assuming that both subject and object are plural, is shown in (122).

(122) Meaning target for (121) with plural S and O



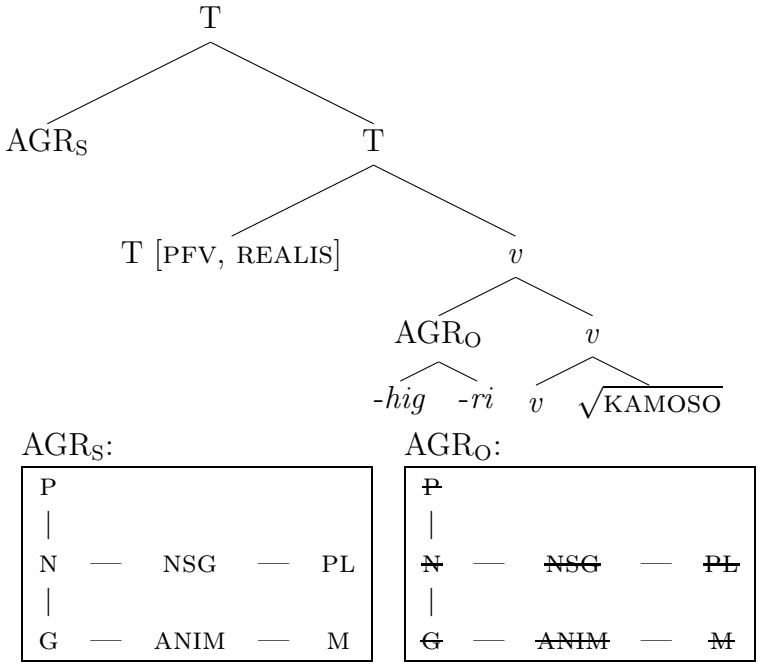
The derivation begins with the insertion of the root VI, followed by vocabularization of the v node (no effect). Next, vocabulary insertion proceeds to the object AGR node. The first cycle inserts the string *-ri*, realizing the features [3, M, ACC]. The object's number features remain unexpressed, and so Split applies. The resulting structure is shown in (123).

(123)



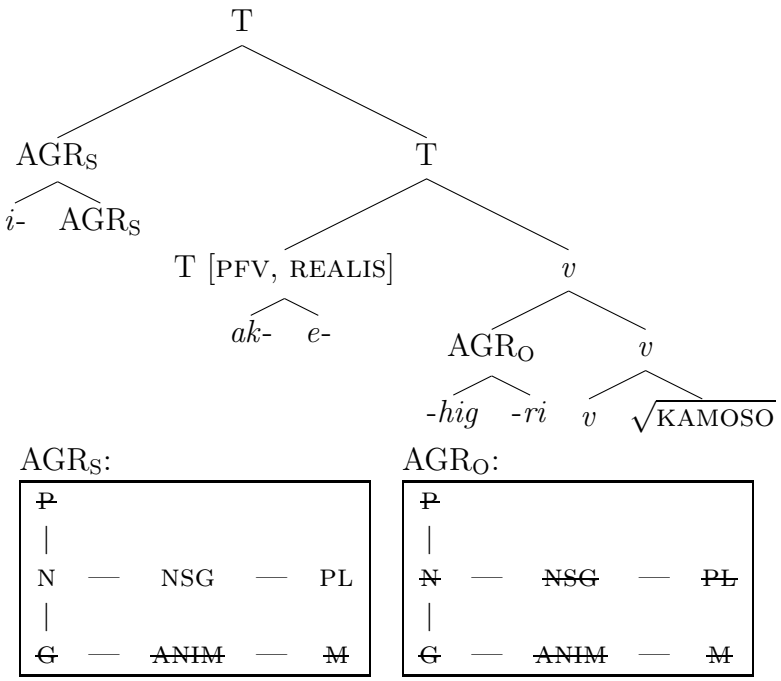
The second cycle of Insert at the object agreement node adds the string -*hig*, fully realizing the object's meaning target. The resulting structure is shown in (124).

(124)



Continuing with vocabulary insertion at the T node, *ak-* and *e-* are inserted. Finally, at the subject agreement node the string *i-* is added on the first cycle, realizing the features [3, M, NOM] and triggering Split. The resulting structure is in (125).

(125)



The AGR$_S$ meaning target would now compel us to again insert *-hig* ↔ [PL], but recall that this meaning target is just one part of the whole word meaning target in (122). I appeal here to the notion that insertion operations at outer nodes can "see" the material inserted at inner nodes (Noyer 1992). Recalling the idea that all insertions should be semantically or structurally optimizing (§5.4.1), note that both N insertions would realize the same featural content with an identical form. Since the plural-marking VI has already been inserted by the inner agreement node, re-inserting it would not move the form closer to its target meaning. In other words, the subject's number features have already been realized by a VI inserted by the object. Because the second insertion appears to contribute nothing to the output form, it fails to apply[19].

This treatment of ambiguity is a logical consequence of my view of the factors that compel and constrain morphological operations. Under this view, multiple exponence would arise just in case the second realization of N made a structural contribution to the output form, increasing its well-formedness in some way. Otherwise the second insertion is blocked, resulting in an ambiguous pattern. It is important to note that previous formulations of the Insert and Fission operations would have nothing to say about this type of ambiguity. If third person plural agreement nodes in Nanti were specified to undergo Fission, that operation would apply at both AGR nodes

[19]Note that the exponence of the subject's person/gender marker is not similarly blocked because its VI (*i-* ↔ [3, M, NOM]) differs from that of the object's person/gender marker (*-ri* ↔ [3, M, ACC]). The form of the plural marker is identical whether it cross-references the subject or the object, so I assume a single plural VI (*-hig* [PL]).

to realize all of the node's features. The idea that Insert and Split are motivated by expressiveness at both the node and the word levels is, I believe, a novel one.

To summarize, there are two conditions which must both be satisfied in order for Insert to apply:

(126) Insert only applies when
  a. It realizes an as-yet-unrealized feature of the meaning target
     AND
  b. The VI being inserted has not previously been added to the word

5.5 Conclusion

This chapter developed a theory of agreement in which richly structured ϕ-sets function to formalize meaning targets and strength of expression. Vocabulary Insertion operates relative to them and is licensed if there are features of the meaning target that remain unrealized. ϕ-sets also constrain insertion order and drive Cyclic Insertion.

Treating the fully valued ϕ-sets as meaning targets frees us from the assumption that nodes are typically restricted to a single vocabulary insertion. Instead, and without special prespecification, insertion applies at each node until its meaning target is expressed as fully as possible. I also set aside the idea that insertion discharges features (§5.4.1).

The noncanonicality and relative infrequency of discontinuous agreement are captured in terms of its increased derivational complexity (number and type of operations). The ϕ-sets thus have both descriptive power, in that they derive the patterns of discontinuous agreement, and explanatory power, in that they add complexity to the derivation.

There is a strong kinship between Béjar's Cyclic Agree and my Cyclic Insertion. Cyclic Agree rejects the idea of feature checking in syntactic agreement and relies instead on the obligatoriness of operations (Béjar 2003: 58, see also Preminger's 2011 rejection of "derivational time bombs"). Cyclic Insertion rejects the idea of feature discharge in morphological agreement and relies instead on the fact that vocabulary insertion is compelled by the drive to fully express a target meaning. Both approaches shift their focus from the outcome of operations to the application of those operations themselves.

Chapter 6

Applying and Extending the Theory

This chapter draws together the preceding material in considering both the contributions and the challenges of the theory of Cyclic Insertion. Section 6.1 illustrates the usefulness of Cyclic Insertion by showing how it can yield new insight into the proper analysis of a complex agreement system. Section 6.2 enumerates some of the predictions of the theory and shows that they are for the most part consistent with the results of the typological study presented in chapter 2. Section 6.3 examines some discontinuous patterns in pronoun formation and TAM marking, in support of the idea that the broader phenomenon of discontinuous exponence can be given a unified treatment. Finally, section 6.4 considers some challenges for future work.

6.1 Case Study: Reanalyzing Karuk

The theory of discontinuous agreement laid out in chapter 5 is concerned with the syntax of agreement, the morphology of agreement, and the relationship between the two. Thinking about agreement in this way involves a commitment to looking closely at the morphology and accounting for the way that features are bundled on the various agreement morphs. Morphological complexity is taken on its own terms, rather than being reduced to a side effect or explained away. This kind of approach can be a fruitful way to consider existing analyses of agreement data in a new light.

This section considers two previous analyses of agreement in Karuk, a Hokan language indigenous to Northern California, and proposes a reanalysis that captures certain insights of each but that, owing to Cyclic Insertion's focus on fully explaining morphology, is more descriptively adequate than both. The Karuk positive indicative

		Object					
		1sg	**2sg**	**3sg**	**1pl**	**2pl**	**3pl**
Subject	**1sg**	—	nu-́	ni-	—	ki:k-́V-ap	ni-
	2sg	ná-	—	ʔi-	kín-́	—	ʔi-
	3sg	ná-	ʔi-V-ap	ʔu-	kín-́	ki:k-́V-ap	ʔu-
	1pl	—	nu-́	nu-́	—	ki:k-́V-ap	nu-́
	2pl	ka-ná-́	—	ku-	kín-́	—	ku-
	3pl	ka-ná-́	ʔi-V-ap	kun-́	kín-́	ki:k-́V-ap	kín-

Table 6.1: Karuk positive paradigm

		Object					
		1sg	**2sg**	**3sg**	**1pl**	**2pl**	**3pl**
Subject	**1sg**	—	nú-	kán-́	—	ki:k-́V-ap	kán-́
	2sg	ná-	—	0-	kín-́	—	0-
	3sg	ná-	ʔi-V-ap	kám-	kín-́	ki:k-́V-ap	kám-
	1pl	—	nú-	nú-	—	ki:k-́V-ap	nú-
	2pl	ka-ná-́	—	ki:k-́	kín-́	—	ki:k-́
	3pl	ka-ná-́	ʔi-V-ap	kun-́	kín-́	ki:k-́V-ap	kín-

Table 6.2: Karuk optative paradigm

and optative verb paradigms[1] are presented in Tables 6.1 and 6.2[2] . Agreement is marked largely by prefixes, but some forms have both a prefix and a suffix *-ap* (namely, forms with a third person subject and second person singular object and all forms with a second person plural object). Notice that the object controls agreement in some forms, as shown by the consistent agreement marking in the first person and second person object columns in the tables below, while the subject controls agreement in other forms, as shown by the varying forms of the agreement marker in the third person object columns.

6.1.1 Previous analyses: Macaulay and Béjar

Macaulay (1992) proposes that Karuk employs a distinction between direct and inverse agreement and that the suffix *-ap* is a marker of inverse contexts. Under her

[1]Karuk also has a negative indicative paradigm, discussed by Bright (1957) and Macaulay (1992), which I do not consider here.

[2]An accent diacritic on vowels (e.g. á) is an acute accent. An accent diacritic over a hyphen is used by Bright to indicate the accent shift conditioned by the prefix.

view, a verb is marked with *-ap* just in case a subject is lower on the person hierarchy than the object; her proposed person hierarchy for Karuk is given in (127).

(127) Karuk person hierarchy (Macaulay 1992: 188)
2PL > 1 > 2SG > 3

This analysis is appealing in that it helps us understand why agreement is controlled by the subject in some contexts and the object in others. Whichever argument is higher on the hierarchy will be marked, so agreement is controlled by a first person object (e.g. *ná-* 3SG > 1SG) but by a first person subject when the object is third person (e.g., *kán-́* 1SG > 3SG). The hierarchy also makes largely valid predictions about the distribution of the *-ap* suffix, which had previously been regarded as unpredictable (Macaulay 1992: 182).

On the other hand, Macaulay's analysis predicts the presence of *-ap* in a wider range of constructions than it actually appears in. According to the person hierarchy in (127) the inverse marker should appear in the contexts 2SG > 1SG/PL, 3SG > 1SG/PL, and 3PL > 1SG/PL, but in fact *-ap* is not found in any of them[3]. A second problem with the analysis is that it treats the prefixes *ná-* and *kaná-* as both meaning '1sg object' (Macaulay 1992: 189), thus missing the segmentability of *ka-* and *ná-*.

Chapter 4 laid out the theory of Cyclic Agree proposed by Béjar (2003) and Béjar & Rezac (2009). Béjar analyzes Karuk as a low-ϕ language, capturing the tendency for the object to control person and number agreement. In case the object is underspecified with respect to the person or number probe (when the object is third person or singular, respectively), the probe projects and may be valued by a more highly specified subject (local person or plural, respectively). The cyclic nature of agreement in Béjar's view thus serves the same purpose as the person hierarchy in Macaulay's analysis. Béjar analyzes the individual affixes as shown in Table 6.3[4].

As a low-ϕ language, Karuk locates the person and number probe on v, deriving the preference for agreement with the object. The internal structure of the probe (described in section 5.3.1) privileges agreement with local person over third ([uP, uPART]) and plural number over singular ([uNSG]).

Béjar's analysis of Karuk showcases some of the strong points of her theory of Cyclic Agree. The preferential marking of object agreement and exceptional marking of subject agreement can be understood in terms of the probe location (low in the tree so it probes the inner argument first) and the specification of its uninterpretable features (the presence of the uninterpretable second person feature means that the

[3]Macaulay's analysis does predict the distribution of *-ap* more accurately in the negative indicative paradigm.

[4]Béjar simplifies Bright's orthography in several ways. Accent marks are eliminated, collapsing some distinctions between forms, for instance nu-́ 1>2SG in the positive paradigm and nú- 1>2SG in the optative paradigm. As well, the 2PL > 3 agreement marker, *ki:k-́*, is listed by Béjar as *ki-* because it appears in that form in certain contexts.

First cycle (O) agreement morphs:

na-	↔	1
nu- ∼ʔi-	↔	2
kin-	↔	1 / PL
ki-	↔	2 / PL
-ap	↔	PL

Second cycle (S) agreement morphs:

ni-	↔	1
ʔi-	↔	2
ʔu-	↔	3
nu-	↔	1 / PL
ku-	↔	2 / PL
kun-	↔	3 / PL
ka-	↔	PL

Table 6.3: Karuk agreement morphology (Béjar 2003: 160)

probe must agree with a local person to be fully valued). This analysis thus allows for the decomposition of portmanteaux and captures the segmentability of *ka-* and *ná-*, treating the former as a plural subject marker and the latter as a first person object marker.

However, this analysis also falls short in some ways. In addition to collapsing certain formal distinctions (see footnote 2, this chapter), Béjar analyzes *-ap* as a plural marker and thus fails to account for its presence in 3SG>2SG forms and its absence in all forms involving third person plural arguments[5]. Her analysis also cannot explain why the second person object marker is sometimes *nú-* and sometimes *ʔi-*, because in this system the second person object fully values the person probe. We therefore would not expect subject person (the only featural difference between the *nú-* forms and the *ʔi-* forms) to be realized.

The analyses of Béjar and Macaulay have some overlap. They both conceive of agreement as the competition of multiple markers for a limited number of morphological positions, though they differ in the details of how this competition is implemented. Macaulay assumes that the winner is determined by the person hierarchy in (127), while Béjar assumes that local person wins over third person and (separately) that plural number wins over singular number. Both approaches have analytical advantages and weaknesses. In the next section I propose a reanalysis inspired by my own work on Cyclic Insertion that incorporates some of the advantages of each approach

[5]Béjar suggests that agreement in Karuk is never sensitive to the number of third person arguments and that this is why *-ap* doesn't pattern with these arguments. However, her presentation of the positive indicative paradigm described by Bright (1957) introduces a typographical error: Her paradigm shows the 3PL > 3PL prefix as *kun-* but it is in fact *kín-*. Macaulay's fieldwork agrees with Bright's description of this form, as does in-progress work by Andrew Garrett and Line Mikkelsen. This is an important point because the fact that 3PL > 3SG verbs are marked differently from 3PL > 3PL verbs shows that agreement is sensitive to the number of third person objects. If *-ap* is, in fact, a marker of plural objects it should appear in forms with third person plural objects, but it does not.

Positive paradigm (partial)

		Object			
		1sg	**2sg**	**1pl**	**2pl**
Subject	**1sg**	—	nu-́	—	ki:k-́V-ap
	2sg	ná-	—	kín-́	—
	1pl	—	nu-́	—	ki:k-́V-ap
	2pl	ka-ná-́	—	kín-́	—

Optative paradigm (partial)

		Object			
		1sg	**2sg**	**1pl**	**2pl**
Subject	**1sg**	—	nú-	—	ki:k-́V-ap
	2sg	ná-	—	kín-́	—
	1pl	—	nú-	—	ki:k-́V-ap
	2pl	ka-ná-́	—	kín-́	—

Table 6.4: Karuk positive and optative paradigms, first and second person

while avoiding most of the weaknesses.

6.1.2 Reanalysis: ϕ-sets and probe structure

I begin by examining more closely the distribution of the *-ap* morpheme, analyzed by Macaulay as an inverse marker and by Béjar as a plural object marker. The following tables present the agreement paradigm in a pairwise fashion; Table 6.4 shows the interaction of first and second person, Table 6.5 the interaction of first and third person, and Table 6.6 the interaction of second and third person.

The difference between the positive and optative paradigms for this subset of forms is quite subtle. The second person object forms are *nu-́* (with accent over the hyphen indicating accent shift conditioned by the prefix) in the positive and *nú-* (with acute accent on the prefix itself) in the optative[6]. Notice also in Table 6.4 that *-ap* appears in forms with first person singular or plural subjects and second person plural objects. This is problematic for both the inverse marker hypothesis and the plural marker hypothesis: if *-ap* is an inverse marker then it should appear in all 1>2 forms, not just 1>2PL, and if it is a plural object marker it should be found with 1PL objects in addition to 2PL objects.

[6]I am grateful to Line Mikkelsen for pointing out this distinction to me, and for also noting that these forms should be analyzed as distinct morphs, which Béjar's analysis does not account for (see footnote 4, this section).

Positive paradigm (partial)

		Object			
		1sg	**3sg**	**1pl**	**3pl**
Subject	**1sg**	—	ni-	—	ni-
	3sg	ná-	ʔu-	kín-́	ʔu-
	1pl	—	nu-́	—	nu-́
	3pl	ka-ná-́	kun-́	kín-́	**kín-**

Optative paradigm (partial)

		Object			
		1sg	**3sg**	**1pl**	**3pl**
Subject	**1sg**	—	kán-́	—	kán-́
	3sg	ná-	kám-	kín-́	kám-
	1pl	—	nú-	—	nú-
	3pl	ka-ná-́	kun-́	kín-́	kín-

Table 6.5: Karuk positive and optative paradigms, first and third person

Table 6.5 shows that -*ap* does not appear in contexts involving first and third person only (but see Macaulay for additional evidence from the negative indicative paradigm). This is explicable under the inverse marker hypothesis – there is no hierarchical ranking between first and third person, at least in these two paradigms – but is troubling with respect to the plural marker hypothesis, under which -*ap* is predicted in all of the plural object forms.

The forms in Table 6.6 strongly support the inverse marker hypothesis. Notice that -*ap* appears in all forms with third person subjects and second person objects and in no forms with second person subject and third person objects. Furthermore, -*ap* is not found with 3PL objects but is found with 2SG objects, regardless of subject number – two big problems for the plural marker hypothesis.

Taken together, Tables 6.4–6.6 better support an analysis of -*ap* as an inverse marker. There is evidence that second plural outranks first person (-*ap* appears in 1>2PL forms but not 2PL>1 forms) and evidence that second person outranks third person (-*ap* appears in all 3>2 forms but no 2>3 forms). If we make the simplifying assumption that second person outranks first person across all numbers and that the failure of -*ap* to appear in 1>2sg forms is due to factors external to the person hierarchy, we have established the following modified hierarchy[7].

[7]Macaulay (1992) adduces evidence from the negative imperative paradigm to further refine this hierarchy, and so it may be possible to establish a relative ranking between first and third person. I suggest, though, that the distinction between 2SG and 2PL unnecessarily complicates the issue. The

Positive paradigm (partial)

		Object			
		2sg	**3sg**	**2pl**	**3pl**
Subject	**2sg**	—	ʔi-	—	ʔi-
	3sg	ʔi-V-ap	ʔu-	ki:kˊ-V-ap	ʔu-
	2pl	—	ku-	—	ku-
	3pl	ʔi-V-ap	kunˊ-	ki:kˊ-V-ap	**kín-**

Optative paradigm (partial)

		Object			
		2sg	**3sg**	**2pl**	**3pl**
Subject	**2sg**	—	0-	—	0-
	3sg	ʔi-V-ap	kám-	ki:kˊ-V-ap	kám-
	2pl	—	ki:kˊ-	—	ki:kˊ-
	3pl	ʔi-V-ap	kunˊ-	ki:kˊ-V-ap	kín-

Table 6.6: Karuk positive and optative paradigms, second and third person

(128) Revised person hierarchy for Karuk: 2 > 1, 3

6.1.3 Cyclic insertion

Although I am proposing the revised person hierarchy shown in (128), a quick look at the paradigms in Tables 6.1 and 6.2 will convince us that Bejar's probe structure is correct. We find *ná-* marking agreement in all forms with 1SG objects, *kínˊ-* in all forms with 1PL objects, and *ki:k-* in all forms with 2PL objects[8]. In the third person object columns we can see that the forms instead show a sensitivity to subject person and number: *kánˊ-* marks agreement in 1SG > 3SG/PL, *kám-* in 3SG > 3SG/PL, and so forth. All of this is consistent with Béjar's analysis of Karuk as a language in which the probe is specified for local person and plural number and is located low in the tree on v. It first probes the internal argument, which fully values the probe if it is first or second person and plural; in this case the object controls agreement. Otherwise, the external argument is probed and may value additional features on the probe, triggering the insertion of second-cycle agreement morphology indicating that the subject controls agreement.

analysis developed here fares better because I do not make this distinction, as I will show below.

[8]There is some complexity in the 2SG object column which I take to be further evidence of a person hierarchy in Karuk: *nú-* marks 2SG object in direct contexts and *ʔi-* marks 2SG object in inverse contexts.

(129) Probe specification in Karuk

P	—	PART
|		
N	—	NSG

While the person hierarchy shows second person outranking both first and third person, the probe groups first and second person as being preferred over third person as controllers of verb agreement. In other words, the person hierarchy cuts the set of person features in a different way than probe structure does. While the person hierarchy privileges second person over first and third, the probe groups first and second person together as privileged over third. I thus carry forth Béjar's insight that many patterns of context sensitive agreement derive from the specification of the probe in combination with the way the agreement operations work, while rejecting her stronger claim that what appear to be person hierarchy effects ultimately reduce to fine-grained syntactic structure in all cases. Rather, the role of probe structure is to establish the appropriate relations between controllers and targets of agreement and the role of the person hierarchy is to structure the direct/inverse opposition.

A full reanalysis of Karuk agreement morphology in the positive indicative and optative paradigms is presented in Table 6.7; compare this with Béjar's analysis in Table 6.3. The most striking difference is the treatment of *-ap*, which for Béjar is a plural marker and for me (following Macaulay) is an inverse marker. Because I retain a notion of the person hierarchy as existing independently of probe structure I can explain the difference between *nu-́* and *ʔi-*, the forms she analyzes as being in alternation. Under my view, *ʔi-* is limited to inverse contexts (and thus is expected to co-occur with the inverse marker *-ap*) and *nu-́* is inserted elsewhere.

This reanalysis also differs from Béjar in that it introduces markers specific to the optative paradigm, which Béjar did not consider (again see footnotes 4 and 6, this chapter). These are the affixes whose insertion is conditioned on OPT in Table 6.7. The most significant analytical differences in second cycle morphology, though, lie between the two sets of double lines. As discussed in footnote 5 above, Béjar (2003) introduced a typo into the paradigm: where she has the form *kun-́* for both 3PL > 3SG and 3PL > 3PL, in fact these forms are different; the correct form for 3PL > 3PL is *kín-*. I include *kín-* here, specifying that it realizes a third person plural subject in the context of another plural argument. The *kun-́* form is thus limited to the elsewhere case and will be inserted in the context of a singular argument.

Recall that in Béjar's analysis, *-ap* was treated as a plural object marker that would be inserted on the first cycle of agreement, thus exhausting the probe's number features. She therefore did not need to rule out the appearance of the *ka-* plural subject marker with second person objects (**ka-nu-́V*). Because I am instead treating *-ap* as an inverse marker, I must constrain the insertion of *ka-* to realize a plural subject in the context of a first person argument. This prevents its overinsertion and

Inverse marker:		
-ap	↔	INV

First cycle (O) agreement morphs:		
ná-	↔	1
nu-	↔	2
nú-	↔	2.OPT
ʔi-	↔	2 / INV
kín-	↔	1PL
ki:k-	↔	2PL

Second cycle (S) agreement morphs:		
ni-	↔	1
kán-	↔	1.OPT
ʔi-	↔	2
0-	↔	2.OPT
ʔu-	↔	3
kám-	↔	3.OPT
nu-	↔	1PL
nú-	↔	1PL.OPT
ku-	↔	2PL
ki:k-	↔	2PL.OPT
kun-	↔	3PL
kín-	↔	3PL / PL
ka-	↔	PL / 1

Table 6.7: Karuk agreement morphology, reanalyzed

explains why it is not found with second person objects.

The commitment of Cyclic Insertion to look beyond syntactic relations and say more about the morphology of agreement has helped us to mediate between the analyses of Macaulay and Béjar and ultimately to improve on both. Retaining the notion of an overt inverse marker whose distribution is determined by a person herarchy gives my analysis some advantages over Béjar's. I can predict the distribution of *ap-* more accurately and I can explain certain alternations that Béjar was unable to account for. Retaining certain insights of Béjar also gives this analysis some advantages over Macaulay's. The notion of a probe that is sensitive to person and number features of both object and subject allows me to capture the segmentability of the *ka-ná´-* agreement markers; *ná´-* marking first person object and *ka-* marking subject plurality. Finally, by simplifying the person hierarchy I reduce the over-insertion of the inverse marker *-ap*.

This reanalysis also led to the important conclusion that it is necessary to retain a direct/inverse opposition conditioned by an independent person hierarchy. I have thus falsified Béjar's strong claim that the person hierarchy can in all cases be reduced to probe structure and distribution. I showed that the person hierarchy and the characteristic probe structure divide up the same set of values in different ways: in the person hierarchy, second person is privileged over first and third as determined by

the distribution of the inverse marker. In probe structure, local person is privileged over third person in that only local person can fully value a probe. This is important evidence that probe structure and person hierarchy are not equivalent.

6.2 Consequences and Predictions of Cyclic Insertion

Several predictions follow from the theory of Cyclic Insertion. Before turning to these predictions, I briefly review some of the key ideas that will be relevant to the discussion below.

A foundational idea that has been a thread throughout the development of this thesis is that there is a direct relationship between increased derivational complexity and decreased typological frequency (for instance, see section 1.2.1 and the discussion of canonicality in chapter 3). Taking this idea to its logical conclusion has several consequences for the theory, as will be discussed in sections 6.2.1 and 6.2.2 below.

In chapter 5 I laid out a theory of discontinuous agreement that treats discontinuous exponence as a morphological consequence of a particular syntactic state of affairs in a language with a particular kind of lexical inventory. It arises when the morphology cannot fully express the meaning target at a terminal syntactic node with a single vocabulary insertion because the language does not include a vocabulary item that expresses all of the relevant features. A Cyclic Insertion analysis thus relies on meaning targets that are established in the syntax and that the morphology aims to express as fully as possible; consequences of this assumption will also be discussed in sections 6.2.1 and 6.2.2.

The evidence for and the widespread acceptance of the Fusion Assumption (§3.1) led us to believe that the relationship between a piece of morphological form and an element of meaning is one-to-one in the default case but that morphological operations may disrupt that relationship. Specifically, I proposed the morphological operation Split that is obligatorily invoked when a vocabulary insertion fails to fully realize its node's meaning target (§5.4.2). Split creates an additional position of exponence into which another vocabulary item may be inserted. In this way, greater morphological complexity follows directly from greater derivational complexity, driven by the overarching goal to fully express meaning. The fact that a failure to fully express a meaning target incurs an application of Split is relevant to sections 6.2.1 and 6.2.2.

(130) Key ideas for sections 6.2.1 and 6.2.2

a. Increased derivational complexity corresponds with decreased typological frequency.

b. The syntax sets up meaning targets that the morphology aims to express as fully as possible.

c. Split applies and creates an additional position of exponence just in case a morphological insertion does not fully realize its node's meaning target.

The analysis advanced in chapter 5 also rests on a particular view of the internal structure of sets of agreement features, which I have termed "rich ϕ-set structure" (§5.2), that captures relations among feature categories and also decomposes the categories themselves into primitive features with relations among them. The structure of ϕ-sets has consequences for affix ordering, as discussed in section (6.2.3).

Finally, one way in which the Split operation proposed here differs from previous formulations of Fission is that it is structure-building: each application of Split adds another layer to the internal structure of the morphological agreement node. This structure reflects the rich structure of ϕ-sets and has implications for affix order, as discussed in section 6.2.3.

(131) Key ideas for section 6.2.3

a. The rich internal structure of ϕ-sets encodes relations both among and within feature categories.

b. Split is structure-building; the earliest insertions are hierarchically superior to later insertions.

6.2.1 Frequency of fused agreement morphs

Cyclic Insertion predicts that fused markers are preferred over bipartite markers expressing the same features. This is predicted both within a particular language and crosslinguistically. Within a language the prediction follows directly from (130c) above, the idea that the Split operation applies only when a vocabulary insertion does not fully realize a meaning target. If a node's meaning target contains values P and N and the language has a fused marker that realizes P.N and two discontinuous markers realizing P and N, respectively, then the P.N marker will always be inserted first (recall the discussion of Noyer's Spell-Out Ordering Hypothesis and specifically his application of Panini's Principle in section 5.4.2). Once the P.N marker is inserted the meaning target is exhausted, and no features remain unexpressed to condition an application of Split which would create a new position of exponence.

Under Cyclic Insertion there is also the expectation that fused markers should be crosslinguistically more common than bipartite markers. The synchronic prediction discussed immediately above means that in a language with discontinuous markers P and N, if P through frequent co-occurrence with N semantically shifts to mean P.N, it will in effect "crowd out" the discontinuous marker that encodes N. The way that the Split operation works synchronically thus sets up a diachronic pressure toward fused markers: split markers should fall into disuse when they are no longer necessary. This is consistent with the idea, expressed in (130a) and discussed in more

detail in section 1.2.1 above, that increased derivational complexity corresponds with decreased typological frequency.

6.2.2 Frequency of subtypes of discontinuous and multiple exponence

The typological study presented in chapter 2 revealed several crosslinguistic tendencies of discontinuous exponence. This section shows that the predictions of the theory presented in chapter 5 are aligned with these tendencies.

Cyclic Insertion is predicated on the idea that DE is noncanonical and is intended to capture its status as a variation from the default one-to-one relation between syntactic feature sets and morphological feature expressions. The foundational idea of derivational complexity (130a), which captures the noncanonicality of DE itself, also predicts that double discontinuities should be rarer than single discontinuities (§6.2.2). Given the implementation details of the Split operation, it also follows that impure discontinuities and combinatorial exponence should be relatively rare (§6.2.2) and that multiple exponence should be rarer than single exponence (§6.2.2).

Double discontinuities

Chapter 3 established the noncanonicality of discontinuous exponence, and I took this noncanonicality as support for the Fusion Assumption (3.1). The theory of Cyclic Insertion relies on an extra operation, Split, that applies in derivations of discontinuous agreement but not in derivations of fused agreement. The increased derivational cost incurred by this extra operation captures the fact that DE is typologically rarer than fused agreement. By the same logic, I predict that double discontinuities such as (132b), in which a ϕ-set is realized by three distinct morphs, should be rarer than single discontinuities like (132a), in which the ϕ-set is realized by just two distinct morphs.

(132) a. *P-V-N.G*
b. *P-V-N-G*

The typological study in Chapter 2 produced 14 patterns that express all three agreement features affixally. Of these, only 3 patterns encode all three features with a dedicated morph and thus show a double discontinuity; the remaining 11 patterns show a single discontinuity. The survey data thus suggest that two-way split patterns are more common than three-way split patterns, which is consistent with the prediction.

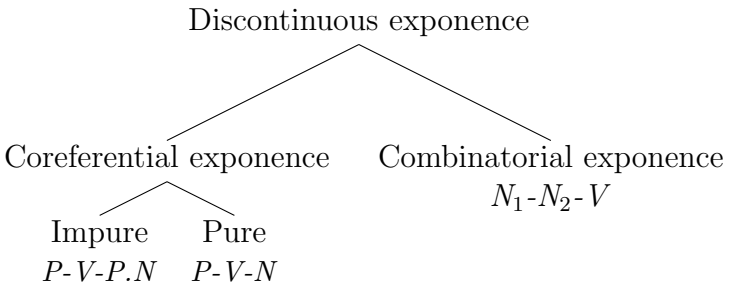


Figure 6.1: Parameters of discontinuous exponence (partial)

Morphosyntactic Type and Purity

Recall from chapter 2 that the phenomenon of discontinuous exponence can be decomposed into a number of subtypes. One subtype distinction was made on the basis of morphosyntactic type (2.3.1), which distinguishes coreferential exponence from combinatorial exponence. Another distinction was based on morphological purity (2.3.2), distinguishing patterns in which the ϕ-set splits cleanly from those in which one feature is encoded by multiple morphs. The relationships among the subtypes of discontinuous exponence were illustrated in Figure 2.2, repeated here in part as Figure 6.1.

Consider now the lexical insertions required to give rise to these various types of discontinuous exponence. After Split applies, in order for a subsequent insertion to be optimizing (that is, to bring the node closer to its meaning target) it must express features that have not already been expressed. It is easier for this to happen in coreferential versus combinatorial exponence and in pure versus impure exponence because there are simply more ways to derive the former types than the latter types.

To be more explicit, if the first insertion at a node realizes P, then a subsequent insertion of N or G (or P.N, P.G, N.G) would create a coreferential discontinuity while only the insertion of a morph realizing a different value of P would create a combinatorial discontinuity. The implementation of the Split operation thus predicts that combinatorial DE should be much less common than coreferential DE (assuming that languages don't typically have a disproportionately high number of P-expressing morphs). Similarly, if the first insertion at a node realizes P, then a subsequent insertion of N, G, or N.G would create a pure discontinuity while the insertion of a morph realizing P.N or P.G would create an impure discontinuity. There are more ways to get a pure discontinuity, so I expect that impure discontinuities should be rarer than pure discontinuities.

The typological study produced 69 patterns involve some type of coreferential discontinuity[9]. Of these 69 patterns, the discontinuity is pure in 56 patterns and

[9]The careful reader will note that the number of examples reported in rows (a) and (b) of Table 6.8

	Morphosyntactic Type	Subtype	Pattern Count
a.	Coreferential	Pure	56
b.	Coreferential	Impure	17
c.	Combinatorial	N/A	6

Table 6.8: Number of survey patterns showing various types of discontinuity

impure in 17 patterns. Only six patterns show a combinatorial discontinuity of an agreement or TAM feature.

Again, the study results are consistent with the predictions. Coreferential discontinuities are more common than combinatorial discontinuities. Within the coreferential type, pure discontinuities are more common than impure discontinuities.

Multiple exponence

A final prediction about exponence type is that multiple exponence should be less common than single exponence, because multiple exponence must involve Split (or some similar operation) and its derivations are thus more complex[10]. My typological study of discontinuous exponence did not systematically search for multiple exponence discontinuities, so I leave open the question of whether this prediction is borne out.

6.2.3 Ordering tendencies

As I noted in section 5.4.2, a complete account of all morpheme ordering in all the languages surveyed in chapter 2 is beyond the scope of this thesis. However, I proposed that the structure of ϕ-sets conditions an insertion order in which person is inserted first, then number, and then gender. This insertion order, in combination with the details of the Split operation, creates a complex, layered agreement node with person on top, number in the middle, and gender at the bottom. Given standard assumptions about linearization, namely that that linearization starts at the root and proceeds outward (Harbour 2008 and references therein), a layered agreement node makes certain predictions about the ordering of coreferring agreement affixes.

do not total 69. This is because a single pattern may show more than one type of discontinuity. For instance, the pattern *P-V-P.N-G* (Tamazight Berber) shows an impure person-number discontinuity and a pure person-gender discontinuity. There are 69 total patterns showing 73 distinct coreferential discontinuities.

[10]Note that Cyclic Insertion is not intended to derive ME, but does suggest that it should be more costly to derive because some additional operation is required to create the extra positions of exponence.

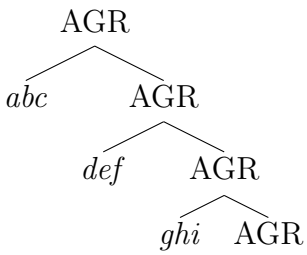
Before I turn to those predictions it is important to note that these agreement affixes are not always contiguous, because morphs that realize other, non-agreement features can intervene between morphs inserted under the complex agreement node (for instance, see example (112) in which the perfective marker intervenes between the first person subject marker and the plural number marker). There are various devices that could explain this kind of morpheme intervention: morphophonological well-formedness conditions, morphological templates, and operations that raise or lower morphosyntactic nodes have all been appealed to in analyses of morpheme order.

In short, the general principles discussed here make predictions about the relative ordering of agreement markers while leaving open the possibility that other kinds of morphs will intervene, disrupting the contiguity, but not the relative ordering, of the agreement markers. This represents progress toward a complete understanding of morpheme order, but there is still much work to be done.

Order of PNG markers

The internal structure of agreement nodes is such that P-expressing morphs are hierarchically superior to N morphs, which are in turn superior to G morphs (131b); this structure is schematized in (133).

(133) Complex agreement node generated by Split



abc ↔ P
def ↔ N
ghi ↔ G

Because linearization is by assumption root-outward, I expect the root to be linearized before any other morpheme. Agreement prefixes will of course precede the root, and suffixes will follow it in linear order. Because the vocabulary items that are inserted first are linearized first, the highest layer of morphological agreement node structure is linearized first and the lowest layer is linearized last. Earlier linearization means closer proximity to the root, and so I expect person to be marked inside of number, which in turn is marked inside of gender. This is shown in (134a) for a prefixing agreement language and in (134b) for a suffixing agreement language.

(134) Order of P, N, and G morphs in discontinuous exponence
  a. G-N-P-VERB
  b. VERB-P-N-G

Of course other morphological operations may disrupt this situation, giving variant orderings at increased derivational cost, but the order schematized in (134) is what I expect to find most frequently across languages.

Trommer's Generalization (65), discussed in sections 4.2.2 and 5.2.1, was called out as an empirical generalization that a complete theory should capture and so it is not itself a prediction of the theory. However, in explaining affix ordering tendencies using hierarchical structure inside the agreement node I have made a slightly different prediction. While Trommer's generalization is that person marking *precedes* number marking, the prediction of Cyclic Insertion is that person is marked *inside* of number[11].

I would now like to return to the results of the typological study with respect to the order of person and number marking. Table 6.9 shows all patterns involving coreferential discontinuities of person and number. The pattern is schematized in the first column, the second column shows how many times the abstract pattern is instantiated in my study results, the third column indicates whether the pattern supports the P-before-N hypothesis, and the fourth column indicates whether the pattern supports the P-inside-N hypothesis.

These results do not strongly support the P-before-N hypothesis over the P-inside N hypothesis, and may in fact slightly favor the P-inside-N hypothesis. Among the patterns where agreement affixes are prefixed, 3/9 patterns support P-before-N versus 6/10 that support P-inside-N. Among the patterns where agreement affixes are suffixed, 21/29 patterns support P-before-N versus 21/30 that support P-inside-N. All of this suggests that Trommer's generalization may be reformulated as a hierarchical rather than a linear generalization, and that his alignment-based analysis may be better captured in terms of linearization.

Weak inside strong

A final consequence of Cyclic Insertion is that when two values of a single feature category are marked, the weaker expression should in the default case occur closer to the root than the stronger expression. This prediction follows the same basic logic as that for the order of PNG markers, immediately above.

Recall from section 5.4.1 that the combinatorial double expression of a single feature category arises when the first vocabulary insertion only partially expresses the meaning target. The unrealized features of the meaning target trigger an application

[11] Thanks to Sharon Inkelas for raising the possibility of a hierarchical interpretation of Trommer's Generalization.

Pattern	Count	P-before-N	P-inside-N
Prefixing:			
P-N-V	3	Y	N
N-P-V	6	N	Y
P-V.N	1	–	N
		3Y; 6N	6Y; 4N
Suffixing:			
V-P-N	20	Y	Y
V-N-P	8	N	N
V.N-P	1	–	N
P-V-P-N	1	Y	Y
		21Y; 8N	21Y; 9N
Mixed:			
P-V-N	25	Y	–
N-V-P	1	N	–
		25Y; 1N	

Table 6.9: Linear order of person and number in discontinuous agreement (my survey)

of Split, creating another position of exponence into which the stronger expression can be inserted. Because the partial expression is inserted first, it is structurally higher in the agreement node and will be linearized first and will thus be marked closer to the root.

As discussed above, combinatorial exponence is relatively rare. Additionally, it's not always possible to determine the relative order of the exponents; for instance, in Kiowa the two underlying expressions of number fuse into a single surface form (Watkins & McKenzie 1984: 122) and in Kashmiri the two expressions of tense flank the verb root (Wali & Koul 1997: 225). For these reasons it is difficult to test this prediction against the results of my language study. The clearest example of combinatorial exponence comes from Hupa and is consistent with the prediction, as seen in (135) in which the weaker expression of number (*di-* 1NSGS) is marked closer to the root than the stronger expression (*ya:-* PL).

(135) na:-ya:-s-di-l-to'n
PVB-PL-PFV-1NSGS-THM-jump
'we (PL) danced' (Sapir & Golla 2001)

The weak-inside-strong prediction is also consistent with an important premise of Inkelas & Caballero's work. They cite diachronic literature on "hypercharacterization, a change in stem or word form brought upon paradigmatic analogy when an

inner marker is not marking a category transparently enough, and an outer, more productive marker is available" (Caballero & Inkelas to appear: 22). The diachronically older, less transparent marker, having been grammaticalized earlier in the language's history, appears closer to the verb root and is inserted first in synchronic derivations. The newer, more productive marker appears farther away from the root and is inserted later.

6.3 Related Phenomena

Chapter 2 argued that discontinuous exponence is a unified phenomenon that can and should be studied on its own terms. Chapters 3 through 5 focused on the patterns of discontinuous verb agreement as the clearest and most numerous examples of discontinuous exponence, and thus the easiest to generalize over. In this section I look again at discontinuous exponence in nonverbal domains (6.3.1) and of other types of features (6.3.2). I will show that the theory developed in Chapter 5 is general enough to be extended to these related patterns.

6.3.1 Discontinuous exponence in pronouns

Within the results of the typological study in chapter 2, the pronominal patterns are particularly interesting because within a language the order of the features as expressed by pronoun formatives tends to parallel the order of features as expressed by agreement affixes on the verb. In some cases, features are realized with the same phonological form in both the verbal and the pronominal domain. Furthermore, the featural distinctions made tend to be subject to the same constraints in both systems of a given language. These observations suggest the possibility for a unified account of verb agreement and pronoun formation.

One example of this overlap comes from Classical Nahuatl (Andrews 1975: 14, 17, 175–176), in which there are a number of formal and featural parallels between pronouns, which are built on the pronominal stem *eh*, and agreement markers on verbs (see Table 6.10). Perhaps most striking is that the nominal and verbal morphology both show the same prefix syncretism between first person plural and second person singular forms.

The similarity between pronouns and agreement markers is discussed explicitly by Arregi (1999) for Basque. Table 6.11 shows the Basque pronominal and agreement forms, and (136) gives an example of both kinds of inflection in the same clause. Arregi argues for the same morphosyntactic analysis of both pronouns and verb agreement affixes.

(136) **su**-**e**-k Boston-ea **s**-ixus-**e**-n
2-**PL**-PL.ABS Boston-ALL **2.ABS**-go-**PL**-ABS-PST

	Independent Pronoun	Agreement Affix(es)
1SG	n-eh	ni-
1PL	t-eh-meh	ti-. . . -h
2SG	t-eh	ti-
2PL	am-eh-meh	am-. . . -h

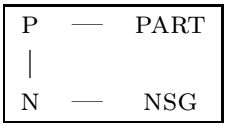
Table 6.10: Nahuatl inflection

	Independent Pronoun	Agreement Affix(es)
1SG	n-i	n-. . . -e
1PL	gu	g-
2SG	su	s-
2PL	su-e	s-. . . -e

Table 6.11: Basque inflection (Arregi 1999: 240)

'you (pl.) went to Boston' (Arregi 1999: 249) Basque

(137) ϕ-set structure for second person plural in Basque



The derivation of the second person plural pronoun will proceed in much the same way as the derivation of a verb marked for second person plural agreement. When vocabulary insertion begins, the pronominal stem (here, phonologically null) will be inserted first. Next, the person-encoding VI matches and the string *su-* to the form, fully realizing the person category of the meaning target. The resulting structure is shown in (138); recall that strikethrough of features in the ϕ-set indicates feature realization, but not feature discharge or deletion (§5.4.1).

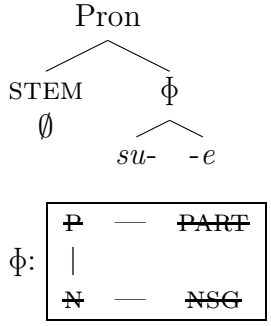
(138)



At this point in the derivation, the number features of the meaning target remain unexpressed. These unrealized features trigger the Split operation, creating a new

position of exponence and licensing a second cycle of Insert. The number-encoding VI matches next; the string *-e* is inserted and the meaning target is fully exhausted. The resulting structure is shown in (139).

(139)



This simple derivation illustrates several points about the extensibility of the Cyclic Insertion theory laid out in chapter 5. First, the concept of meaning targets that the morphology aims to express can be applied to domains other than verb agreement. Second, because they are not tied to syntactic agreement or even to the verbal domain, the operations of Split and Insert are general enough in their implementation to be useful in pronoun formation. At this point in the derivation, it doesn't matter whether the features on a node came to be there through syntactic agreement (as with agreement nodes) or inherent valuing (as with pronouns). Third, not only can these operations be used, but they work in the same way in the two different situations, capturing the crosslinguistic tendency of pronoun formation to parallel verb agreement in languages with discontinuous exponence. This is a major advantage offered by my theory over the others considered in chapter 4, which I discussed in particular for Cyclic Agree (§4.2.1) and Distributed Optimality (§4.2.2).

6.3.2 Discontinuous exponence of TAM

Discontinuous exponence of TAM features is significantly less pervasive than that of agreement features, but it is found in 10 of the 40 languages surveyed in chapter 2. When TAM features are expressed discontinuously, the discontinuity is generally combinatorial in that a single feature category (e.g., tense) is realized by more than one morph, each of which expresses a different value.

In Kashmiri, the remote past tense form (RPST) of an intransitive verb is built on the indefinite past tense form (IPST), which in turn is built on the proximate past tense form (PPST). Example (32) shows the stacking of three tense suffixes, all of which must be present to yield a remote past interpretation.

(140) su a**ː-yeː-yoː-v** vakht-as
3SG.M come-**PPST-IPST-RPST** time-DAT
'He came on time.' (Wali & Koul 1997: 225) Kashmiri

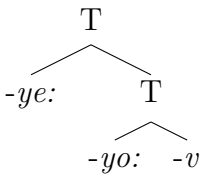
Leaving aside other verb morphology, I assume that Kashmiri remote past entails indefinite past, which in turn entails proximate past. If this is correct, then the meaning target for the tense node is as follows.

(141) Tense meaning target for (140) T:

PPST — IPST — RPST

I further assume that these tense vocabulary items have a contextual limitation so that remote past can only be inserted in the context of indefinite past, which can only be inserted in the context of proximate past. There is thus only one vocabulary item eligible for the first insertion at the T node: *-ye:* (PPST). Next *-yo:* (IPST) will be inserted, followed by *-v* (RPST). Each of the first two insertions only partially realize the meaning target, which triggers the Split operation and creates a new position of exponence. Once the final tense marker has been inserted, the meaning target is fully realized.

(142)



T:

~~PPST~~ — ~~IPST~~ — ~~RPST~~

It is thus possible to extend the treatment of combinatorial agreement – for instance, the derivation of Hupa example (112) in chapter 5 – to the combinatorial discontinuous exponence of TAM features.

6.4 Challenges for Future Work

In this section I discuss two problems that arise when applying Cyclic Insertion to languages from my survey with specific patterns of agreement. The first problem is found in Cree, which has both fused and discontinuous markers of person and number; the markers are found in different syntactic environments. The challenge, according to the prediction in section (6.2.1), is that the more highly specified fused marker should block the insertion, and perhaps even the presence in the language, of the discontinuous markers.

The second problem I will look at begins with Béjar's analysis of Georgian as a "split-ϕ" language, meaning in this case that the person probe is located low in the tree on *v* while the number probe is located higher in the tree on T (Béjar 2003: 117-127). When I carry this analysis through to the morphology, it becomes clear that vocabulary insertion at the lower node is sometimes sensitive to number

features, and vocabulary insertion at the higher node is sometimes sensitive to person features, challenging Béjar's notion that the person and number probes are separate and independent of one another.

6.4.1 Blocking in Cree

The Cree verb has one prefix position and eight suffix positions. There are two distinct ways in which agreement is marked and affixes are arranged on the verb, termed the "independent" and the "conjunct" order. In the independent order, which occurs in main clauses only, person is marked by a prefix and number by a suffix, shown in (143a). In the conjunct order, which occurs both in main clauses and in embedded clauses, person and number are fused on a different suffix, shown in (143b).

(143) Person and number marking in Transitive Inanimate Clauses[12]

a. **ni**-waːpahtːe-**n**
1-see.INAN-THM-**SG**
'I see it' (Dahlstrom 1986: 36) Independent

b. eː-waːpahtːam-**aːn**
PVB-see.INAN-THM-**1SG**
'I see it' (Dahlstrom 1986: 36) Conjunct

The person-number discontinuity seen in the independent order collapses into fused agreement in the conjunct order. The challenge, then, is how to block the insertion of the more highly-specified conjunct order morph *-aːn* (1SG) in the independent order.

One obvious line of attack is to make the fused marker unavailable in independent order by limiting its distribution to the conjunct order. The idea that the two orders have fundamentally distinct morphosyntax is common in the Algonquianist literature, so it is not a huge leap to suggest that there are two distinct sets of agreement markers.

(144) Agreement markers and morphological order in Cree

a. Independent order markers
niː ↔ [1]
-n ↔ [SG]

b. Conjunct order markers
-aːn ↔ [1SG]

[12]The theme sign alternation here is regular: *-eː* is used in the independent order with a first or second person subject, otherwise *-am* is used (Dahlstrom 1986: 36).

Next, I would need to assume that there are two different types of morphological agreement nodes. This is a more serious problem than the distinction between the postive, optative, and negative paradigms in Karuk (§6.1) because here not only are the two kinds of nodes are vocabularized with distinct forms (as in Karuk), but the forms themselves appear in different morphological positions.

(145) Agreement nodes

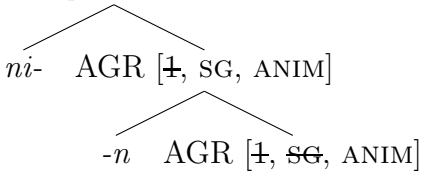
a. Independent order node
AGR$_{\text{indep}}$ [1, SG, ANIM]

b. Conjunct order node
AGR$_{\text{conj}}$ [1, SG, ANIM]

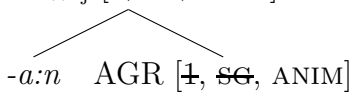
Finally, if I limit the insertion of the fused conjunct order marker in (144b) to the agreement node type in (145b) then it is no longer available for insertion in the independent order. Since there is no longer a competition for insertion between *-n* (SG) and *-a:n* (1SG) in the independent order, the proper discontinuous markers are inserted (146a).

(146) Vocabularized agreement nodes

a. Independent order
AGR$_{\text{indep}}$ [1, SG, ANIM]



b. Conjunct order
AGR$_{\text{conj}}$ [1, SG, ANIM]



6.4.2 Problems with split probe in Georgian

The verb morphology of Georgian can be described with reference to the template in (147). Note that agreement is marked in both prefix position 2 and suffix position 11.

(147) Georgian verb template (Hewitt 1995: 526)
PVB$_1$-**AGR**$_2$-VERSION$_3$-ROOT$_4$-CAUS$_5$-INC/PASS$_6$-THM$_7$-PFV/STV$_8$-IMPF$_9$-MOOD$_{10}$-**AGR**$_{11}$

	S	O	P-contr	N-contr	Example
a.	PL	1/2SG, 2PL	O	S	m-xedav-en 'they see me' g-xedav-en 'they see you (sg/pl)'
b.	PL	1PL	O	S&O	gv-xedav-en 'they see us'
c.	SG	1/2PL	O	O	g-xedav-s 'he sees us' g-xedav-t 'he sees you (pl)'
d.	1/2PL	3	S	S	v-xedav-t 'we see him/them' ∅-xedav-t 'you(pl) see him/them'
e.	1/2SG	3PL	S	*	v-xedav 'I see him/them'

Table 6.12: Georgian transitive agreement patterns

The patterns of person marking in Georgian are similar to those in Karuk (§6.1). A first or second person object controls person agreement, but if the object is third person then the subject controls person agreement. Recall that under Béjar's (2003) theory this is characteristic of a probe that is located low in the tree and is specified for local person agreement. It first probes the lower argument and, if its features are not fully valued (in case of a third person object), a second cycle of Agree is licensed and it probes the higher argument for additional person features.

Number marking in Georgian also shows this kind of cyclic agreement, but privileging the outer argument instead. A plural subject controls number agreement, but if the subject is singular then the object controls number agreement. This is characteristic of a number probe located high in the tree, on the T node, and specified for plural number agreement. If it is not fully valued by the higher argument, it probes the lower argument for additional number features. Béjar thus analyzes Georgian as a "split-ϕ" language (Béjar 2003: 117-127).

Table 6.12 illustrate these patterns of argument control and realization in agreement in Georgian transitive verbs. The first two columns indicate the person and/or number of the subject and object, respectively. The third column shows which argument controls person agreement and the fourth shows which controls number agreement. Rows (a)–(c) show that first or second person objects control person agreement (Set 2 markers); rows (d)–(e) show that when the object is third person, person agreement is with the subject (Set 1 markers). Rows (a), (b), and (d) show that a plural subject controls number agreement; (b) further shows that just in case the object is first person plural it also contributes to number agreement. Rows (c) and (e) demonstrate that in the context of a singular subject, an object can control number agreement if it also controls person agreement.

Georgian agreement markers are traditionally organized into two sets. In light of the preceding discussion, it should now be clear that set 1 (148a) marks subject-

controlled person agreement (second cycle) and number agreement (first cycle); set 2 (149a) marks object-controlled person agreement (first cycle) and number agreement (second cycle).

(148) a. Agreement markers: Set 1 (Hewitt 1995: 526)

	SG	**PL**
1	v-	v-...-t
2	∅/x	∅/x-...-t
3[13]	-s/a/o	-en/nen/es

b. Vocabulary Items: Set 1

v-	↔	1
∅-	↔	2
-s	↔	3 / PRES|FUT
-en	↔	3PL / PRES|FUT
-t	↔	PL

(149) a. Agreement markers: Set 2 (Hewitt 1995: 526)

	SG	**PL**
1	m-	gv-
2	g-	g-...-t
3	∅/s/h-	∅/s/h-...-t

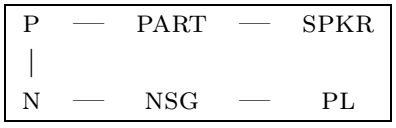
b. Vocabulary Items: Set 2

m-	↔	1
gv-	↔	1PL
g-	↔	2
∅-	↔	3
-t	↔	PL

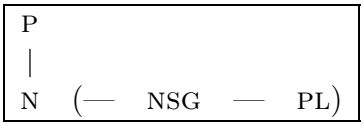
Turning now to an application of Cyclic Insertion to the Georgian data I consider the derivation of (150), in which the subject controls both person and number agreement. The subject and object ϕ-sets are shown immediately below the example.

(150) **v**-xedav-**t**
1-see.PRES/FUT-**PL**
'We see him/them'

S:



O:



In the syntax, the following steps take place.

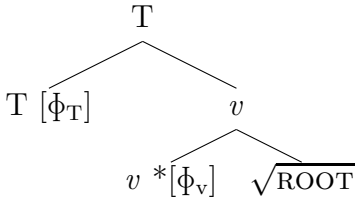
1. v probes the object (first cycle) and gets a third person value: [***u*3–3**, u2]
2. v probes the subject (second cycle) and gets a first person value: *[(u3–3), ***u*2–2–1**]

   (The * diacritic indicates that the probe is valued by second cycle of Agree.)
3. T probes the subject and is fully valued for plural number: T[***u*pl–pl**]

[13]The third person agreement markers in Set 1 are synthetic with tense. The first alternant occurs in the present and future tenses.

The meaning target constructed by the syntax is schematized in (151). The number probe on T has been fully valued by first-cycle agreement with the plural subject and the person probe on v has been fully valued by second-cycle agreement with the first person subject.

(151)



Internal structure of ϕ_T:

N	—	NSG	—	PL

Internal structure of ϕ_v:

P	—	PART	—	SPKR

The morphology in this example is fairly trivial because (following Béjar) the probes are split in the syntax.

1. At $*\phi_v$, the diacritic indicates second-cycle person morphology (set 1). Insertion of v- fully realizes the meaning target, as shown by the list of vocabulary items in (148b).

2. At ϕ_T, the lack of a diacritic indicates first-cycle number morphology (also set 1). Insertion of -t fully realizes number features (148b).

Next I look at the derivation of (152), in which the object controls both person and number agreement. Again, the subject and object ϕ-sets are shown below the example.

(152) **gv**-xedav-**en**
1PL-see.PRES/FUT-**3PL**
'They see us'

S:

P				
|				
N	—	NSG	—	PL

O:

P	—	PART	—	SPKR
|				
N	—	NSG	—	PL

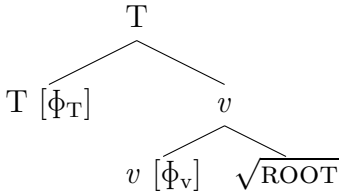
In the syntax, the following steps take place.

1. v probes the object (first cycle) and is fully valued for first person: [***u*3–3**, ***u*2–2–1**]

2. T probes the subject (first cycle) and is fully valued for plural number: [***u*pl–pl**]

The meaning target is schematized in (153). The number probe on T has been fully valued by first-cycle agreement with the plural subject and the person probe on v has been fully valued by first-cycle agreement with the first person object.

(153) Meaning target for (152), first pass



Internal structure of ϕ_T:

N	—	NSG	—	PL

Internal structure of ϕ_v:

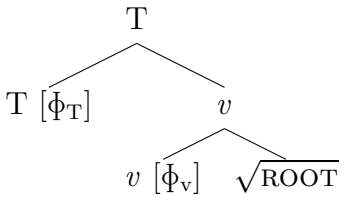
P	—	PART	—	SPKR

Although the features of the meaning targets in (151) and (153) are the same, they differ in that ϕ_v in (151) has a diacritic * while ϕ_v in (153) does not. This reflects the fact that different arrangements of features gave rise to these morphological meaning targets: compare the syntactic nodes in (150) and (152). The consequence is that set 1 will realize person agreement here.

At ϕ_v, insertion of *m-* (set 2) would fully realize the meaning target. However, this isn't the right form! Example (150) shows that the correct form to insert is *gv-* 1PL. It must therefore be the case that the lower agreement node is sensitive to number.

There is a similar problem at the higher node, ϕ_T. Insertion of *-t* would fully realize number features, but again this is the wrong form; it should be *-en* 3PL. It appears that the higher agreement node is also sensitive to the person features of the subject. The meaning target needs to have the distribution of features shown in (154).

(154) Meaning target for (152), second pass



Internal structure of ϕ_T:

P				
|				
N	—	NSG	—	PL

Internal structure of ϕ_v:

P	—	PART	—	SPKR
|				
N	—	NSG	—	PL

Just as section 6.1 showed how a commitment to fully explaining the morphological realization of agreement could lead to a better analysis of an agreement system, I have demonstrated here that the same commitment can make us aware of deeper shortcomings of existing analyses that would otherwise be masked. It is beyond the scope of this chapter to propose a refinement to Béjar's theory, but I have presented clear evidence that a complete split between person and number probes is untenable.

Chapter 7

Conclusions

The three broad aims of this dissertation were to establish discontinuous exponence as a robust phenomenon, to develop generalizations about the behavior of agreement features in discontinuous agreement, and to provide a satisfying analysis of discontinuous agreement.

Chapter 2 provided the empirical base for the discussion. I examined the patterns of discontinuous exponence based on a typological survey of 40 genetically and geographically diverse languages, which allowed me to arrive at several empirical generalizations and to distinguish major subtypes of discontinuous exponence. In Chapter 3 I argued that there is good reason to accept the assumption that agreement features are most commonly fused on a single agreement marker, and worked within the framework of Canonicality Theory to explain why this should be so.

The findings of this work thus support a view of discontinuous exponence as a robust phenomenon that is interesting to study in its own right and that can lead to a better understanding of the mechanisms of agreement. However, it is also the case that the phenomenon is noncanonical and represents a deviation from normal expectations about the relationship between syntax and morphology.

Chapter 4 turned to the theoretical implications of discontinuous exponence. I demonstrated that currently popular theories of agreement fail to fully explain the morphological patterns of discontinuous agreement and cannot offer a unified account of discontinuous exponence more generally. In chapter 5 I proposed Cyclic Insertion, a theoretical framework that quite naturally derives discontinuous exponence while also capturing its noncanonicality. Chapter 6 illustrated the power of the theory by applying it to a reanalysis of one language's agreement system and also suggested how the theory may be extended to account for discontinuous TAM morphology and pronominal formatives.

This thesis also supported the idea that there is a connection between derivational complexity on one hand and noncanonicality and infrequency of use on the other (§1.2.1, 4.1.2, and 6.2.1). The noncanonicality of discontinuous exponence is thus

captured in Cyclic Insertion by its derivational complexity in terms of number of operations. Derivations of forms with fused agreement markers involve only a single vocabulary insertion operation, while derivations of discontinuous agreement involve at least a first insertion, a split, and a second insertion. The same logic argues that single split agreement should be more frequent than double split agreement (§6.2.2), a prediction borne out by the typological results.

Cyclic Insertion rests on the idea that vocabulary insertion is relative to a target meaning. For my purposes, meaning targets take the form of morphological agreement nodes that are valued for all relevant agreement features. They receive their values from the agreement probe(s), which in turn receive their values from one or more of the verb's arguments (§5.3.1). In the morphology, Cyclic Insertion applies as many times as it can until the meaning target is fully expressed or no vocabulary items remain that could contribute to the meaning of the node. This represents a significant departure from standard Distributed Morphology, which posits that features are "discharged" when they condition an insertion and that discharged features are no longer available to condition insertion (§5.4.1). My principle of full expression and Distributed Morphology's principle of feature discharge both serve to block redundant exponence, but my approach allows for multiple insertions at a node without relying on external stipulations.

Finally, chapter 5 drew together prior work on structure among and within feature categories to propose a two-dimensional structure within ϕ-sets (§5.2). These ϕ-sets originate as specifications on the DP arguments of the verb and their values are copied in whole or in part to the agreement probe(s). Importantly, probes located on the same syntactic node may be valued by different ϕ-sets; for instance, a number probe may be valued by the subject and the corresponding person probe by the object. I showed, following Béjar (2003) and Béjar & Rezac (2009), how this structure can derive patterns of context sensitive agreement, for instance in the agreement patterns of Karuk in chapter 6.

A commitment to fully explaining the agreement morphology of Karuk led to the important observation that, contra Béjar and Rezac, an independent person hierarchy must be retained. In other words, person hierarchy effects in agreement cannot be fully reduced to characteristic probe structure. I showed that the facts of agreement in Karuk argue strongly for a direct/inverse opposition in which second person outranks first and third person, while the patterns of context sensitive agreement in the language show a characteristic probe structure in which first and second person are privileged over third person (6.1.3). Because the language's inverse marking system divides up the space of person values in a different way than do the context-sensitive patterns of agreement, there must be something external to probe structure that derives the shape of inverse marking.

A careful study of the phenomenon of discontinuous exponence has illuminated much about the typology and theory of agreement. I have aimed to show that a

commitment to accounting for the syntax and morphology of an agreement system – and the interface between the two modules – can lead to some very interesting insights about the necessary features of a good theory. I hope that future work will extend this approach to other domains of inflectional morphology.

Bibliography

Abdel-Massih, Ernest T. 1971. A Reference Grammar of Tamazight. Ann Arbor: Center for Near Eastern and North African Studies, The University of Michigan.

Abondolo, Daniel. 1982. Verb Paradigm in Erzya Mordvinian. Folia Slavica 5.11–24.

Adger, David. 2003. Core Syntax: A Minimalist Approach. Oxford: Oxford University Press.

Aissen, Judith. 1987. Tzotzil Clause Structure. Dordrecht: Reidel.

Anceaux, J. C. 1965. The Nimboran language; phonology and morphology. 's Gravenhage: M. Nijhoff.

Anderson, Stephen R. 1992. A-Morphous Morphology. Cambridge: Cambridge University Press.

Andrews, J. Richard. 1975. Introduction to Classical Nahuatl. Austin: University of Texas Press.

Aronoff, Mark. 1976. Word Formation in Generative Grammar. Cambridge, MA: MIT Press.

Aronson, Howard. 1990. Georgian: A Learner's Grammar. Columbus, OH: Slavica Publishers.

Arregi, Karlos. 1999. Person and Number Inflection in Basque. In Papers on Morphology and Syntax, Cycle Two, ed. by Vivian Lin, Cornelia Krause, Benjamin Bruening, & Karlos Arregi, volume 34 of MIT Working Papers in Linguistics, 229–264. Cambridge, MA: MITWPL.

——, & Andrew Nevins. 2012. Morphotactics: Basque Auxiliaries and the Structure of Spellout, volume 85 of Studies in Natural Language and Linguistic Theory. Springer.

Baerman, Matthew. 2004. Directionality and (Un)natural Classes in Syncretism. Language 80.807–827.

Baker, Mark C. 2008. The Syntax of Agreement and Concord. Cambridge: Cambridge University Press.

Bale, Alan, Michael Gagnon, & Hrayr Khanjian. 2011. On the Relationship between Morphological and Semantic Markedness. Morphology 21.197–221.

Béjar, Susana, 2003. Phi-Syntax: A Theory of Agreement. University of Toronto dissertation.

Béjar, Susana. 2008. Conditions on Phi-Agree. In Phi-Theory: Phi-features across Modules and Interfaces, ed. by Daniel Harbour, David Adger, & Susana Béjar, 130–154. Oxford University Press.

——, & Milan Rezac. 2009. Cyclic Agree. Linguistic Inquiry 40.35–73.

Bergsland, Knut. 1997. Aleut Grammar. Fairbanks: Alaska Native Language Center.

Bickel, Balthasar, & Johanna Nichols. 2007. Inflectional Morphology. In Language Typology and Syntactic Description: Grammatical Categories and the Lexicon, ed. by Timothy Shopen, 169–240. Cambridge University Press.

Blake, Barry J. 1969. The Kalkatungu Language, volume 20 of Australian Aboriginal Studies. Canberra: Australian Institute of Aboriginal Studies.

Bloomfield, Leonard. 1984. Language. Chicago: University of Chicago Press. Reprint. Originally published 1933 by Holt, Rinehart, and Winston, New York.

Bobaljik, Jonathan David. 2008. Where's Phi? Agreement as a Post-syntactic Operation. In Phi-Theory: Phi-features across Modules and Interfaces, ed. by Daniel Harbour, David Adger, & Susana Béjar, 295–328. Oxford University Press.

Bonet, María Luisa, 1991. The Lengths of Propositional Proofs and the Deduction Rule. University of California, Berkeley dissertation.

Bright, William. 1957. The Karok Language, volume 13 of University of California Publications in Linguistics. Berkeley: University of California Press.

Bybee, Joan. 1985. Morphology: A Study of the Relation between Meaning and Form. Amsterdam, Philadelphia: John Benjamins.

Caballero, Gabriela, & Sharon Inkelas. to appear. Word Construction: Tracing an Optimal Path through the Lexicon. Morphology .

Carstairs, Andrew. 1987. Allomorphy in Inflexion. London: Croom Helm.

Chomsky, Noam. 1991. Some Notes on Economy of Derivation and Representation. In Principles and Parameters in Comparative Grammar, ed. by Robert Freidin, 417–454. Cambridge, MA: MIT Press.

——. 1993. A Minimalist Program for Linguistic Theory. In The View from Building 20 : Essays in Linguistics in Honor of Sylvain Bromberger, ed. by Kenneth Hale & Samuel Jay Keyser, 1–52. Cambridge, MA: MIT Press.

——. 1995a. Bare Phrase Structure. In Evolution and Revolution in Linguistic Theory: Essays in honor of Carlos Otero, ed. by Hector Campos & Paula Kempchinsky, 51–109. Cambridge, MA: MIT Press.

——. 1995b. The Minimalist Program. Cambridge, MA: MIT Press.

——. 2000. Minimalist Inquiries: The Framework. In Step by step: Essays on minimalist syntax in honor of Howard Lasnik, ed. by Roger Martin, David Michaels, & Juan Uriagereka, 89–156. Cambridge, MA: MIT Press.

Corbett, Greville. 2007. Canonical Typology, Suppletion and Possible Words. Language 83.8–42.

——. 2010. Canonical Derivational Morphology. Word Structure 3.141–155.

Corbett, Greville G. 2006. Agreement. Cambridge: Cambridge University Press.

Dahlstrom, Amy, 1986. Plains Cree Morphosyntax. University of California, Berkeley dissertation.

Daniel, Michael. 2011. Plurality in Independent Personal Pronouns. In The World Atlas of Language Structures Online, ed. by Matthew S. Dryer & Martin Haspelmath. Munich: Max Planck Digital Library. Available online at http://wals.info/chapter/35.

Dasgupta, Dipankar. 1978. Linguistic studies in Juang, Kharia Thar, Lodha, Mal-Pahariya, Ghatoali, Pahariya. Calcutta: Anthropological Survey of India.

Derbyshire, Desmond C. 1979. Hixkaryana, volume 1 of Lingua Descriptive Studies. Amsterdam: North-Holland Pub. Co.

Dimmendaal, Gerritt J. 1983. The Turkana Language, volume 2 of Publications in African Languages and Linguistics. Dordrecht: Foris.

Donaldson, Tamsin. 1980. Ngiyambaa: The language of the Wangaaybuwan. Cambridge: Cambridge University Press.

Embick, David, & Rolf Noyer. 2007. Distributed Morphology and the Syntax-Morphology Interface. In The Oxford Handbook of Linguistic Interfaces, ed. by Gillian Ramchand & Charles Reiss. Oxford: Oxford University Press.

Emonds, Joseph E. 1980. Word Order in Generative Grammar. Journal of Linguistic Research 1.33–54.

Evans, Nicholas. 2003. Bininj Gun-wok : A Pan-dialectal Grammar of Mayali, Kunwinjku and Kune, volume 541 of Pacific Linguistics. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Foley, William A. 1991. The Yimas Language of New Guinea. Stanford, CA: Stanford University Press.

Galloway, Brent Douglas. 1993. A Grammar of Upriver Halkomelem, volume 96 of University of California Publications in Linguistics. University of California Press.

Georg, Stefan. 2007. A descriptive grammar of Ket (Yenisei-Ostyak), volume 1 of Languages of Asia. Global Oriental.

Glinert, Lewis. 1989. The Grammar of Modern Hebrew. Cambridge: Cambridge University Press.

Goddard, Pliny Earle. 1905. Morphology of the Hupa Language, volume 3 of University of California Publications in American Archaeology and Ethnology.

Golla, Victor, 1970. Hupa Grammar. University of California, Berkeley dissertation.

Gordon, Lynn. 1986. Maricopa Morphology and Syntax, volume 108 of University of California Publications in Linguistics. Berkeley: University of California Press.

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Universals of Human Language, ed. by Joseph H. Greenberg, 73–113. Cambridge, MA: MIT Press.

Halle, Morris. 1997. Distributed Morphology: Impoverishment and fission. MIT Working Papers in Linguistics 30.425–449.

——, & Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In The View from Building 20 : Essays in Linguistics in Honor of Sylvain Bromberger, ed. by Kenneth Hale & Samuel Jay Keyser, 111–176. Cambridge, MA: MIT Press.

Harbour, Daniel. 2008. Discontinuous Agreement and the Syntax-Morphology Interface. In Phi Theory: Phi-Features across Modules and Interfaces, ed. by Daniel Harbour, David Adger, & Susana Béjar, 185–230. New York: Oxford University Press.

Harley, Heidi. 1994. Hug a tree: Deriving the Morphosyntactic Feature Hierarchy. In MITWPL, volume 21, 289–320.

——, & Rolf Noyer. 1999. Distributed Morphology. In GLOT, volume 4.4, 3–9.

——, & Elizabeth Ritter. 2000. Structuring the Bundle: A Universal Morphosyntactic Feature Geometry. In Pronouns: Features and Representation, ed. by Horst Simon & Heike Weise, 23–39.

——, & ——. 2002. Person and number in pronouns: A feature-geometric analysis. Language 78.482–526.

Hewitt, B. G. 1995. Georgian: A Structural Reference Grammar. Amsterdam, Philadelphia: John Benjamins.

Hock, Hans Henrich. 1991. Principles of Historical Linguistics, Second Edition. Berlin; New York: Mouton de Gruyter.

Hopper, Paul J., & Elizabeth Traugott. 2003. Grammaticalization. Cambridge: Cambridge University Press.

Hualde, Ignacio, & Jon Ortiz de Urbina. 2003. A Grammar of Basque. Berlin; New York: Mouton de Gruyter.

Jakobi, Angelika. 1990. A Fur grammar: Phonology, morphophonology, and morphology. Hamburg: Helmut Buske Verlag.

Johnson, Keith. 2003. Acoustic and Auditory Phonetics. Oxford: Blackwell.

Julien, Marit. 2002. Syntactic Heads and Word Formation. New York: Oxford University Press.

Kayne, Richard S. 1994. The Antisymmetry of Syntax. Cambridge: MIT University Press.

Kim, Yuni, 2008. Topics in the Phonology and Morphology of San Francisco del Mar Huave. University of California, Berkeley dissertation.

Kiparsky, Paul. 2005. Blocking and periphrasis in inflectional paradigms. In Yearbook of Morphology, ed. by Geert Booij & Jaap van Marle, volume 2004, 113–135. Springer Netherlands.

Leer, Jeffrey A. 1991. Evidence for a Northern Northwest Coast language area: Promiscuous number marking and periphrastic possessive constructions. International Journal of American Linguistics 52.158–193.

Lewis, Geoffrey. 2000. Turkish Grammar. Oxford: Oxford University Press.

Macaulay, Monica. 1992. Inverse Marking in Karuk: The Function of the Suffix -ap. International Journal of American Linguistics 58.182–201.

Matthews, Peter H. 1974. Morphology: An Introduction to the Theory of Word-Structure. Cambridge: Cambridge University Press.

—— 1991. Morphology. Cambridge: Cambridge University Press.

McCarthy, John, & Alan Prince. 1994. The emergence of the unmarked: Optimality in prosodic morphology. In Proceedings of the North East Linguistic Society 24, 333–379.

Michael, Lev D., 2008. Nanti evidential practice: Language, knowledge, and social action in an Amazonian society. University of Texas, Austin dissertation.

Mithun, Marianne. 1999. The Languages of Native North America. Cambridge: Cambridge University Press.

Moravcsik, Edith A. 1994. Inflectional Morphology in the Hungarian Noun Phrase: A Typological Assessment. In EUROTYP Working Papers VII/22, 127–185.

Mous, Marten. 1993. A Grammar of Iraqw. Hamburg: Helmut Buske Verlag.

Nash, David G., 1980. Topics in Warlpiri Grammar. MIT dissertation.

Nichols, Johanna. 1986. Head-Marking and Dependent-Marking Grammar. Language 62.56–119.

——, & Balthasar Bickel, 2009. The autotyp Genealogy and Geography Database: 2009 Release. Electronic database, accessible online at http://www.uni-leipzig.de/ autotyp.

NORDLINGER, RACHEL. 1995. Split tense and mood inflection in Wambaya. In Proceedings of the Berkeley Linguistics Society, ed. by Jocelyn Ahlers, Leela Bilmes, Joshua S. Guenter, Barbara A. Kaiser, & Ju Namkung, volume 21, 226–236. Berkeley: Berkeley Linguistics Society.

——. 1998. A Grammar of Wambaya, Northern Australia. Canberra: Pacific Linguistics.

NOYER, ROLF, 1992. Features, positions and affixes in autonomous morphological structure. MIT dissertation.

——, 1993. Optimal Words: Towards a declarative theory of word formation. Unpublished ms., Princeton University.

——. 1997. Features, positions and affixes in autonomous morphological structure. New York: Garland.

PREMINGER, OMER, 2011. Agreement as a Fallible Operation. MIT dissertation.

PRINCE, ALAN, & PAUL SMOLENSKY. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. RUCCS TR-2: Technical Reports of the Rutgers University Center of Cognitive Science.

RAY, SIDNEY HERBERT, & E. BAXTER RILEY. 1933. A Grammar of the Kiwai Language, Fly Delta, Papua. Port Moresby, Papua: E.G. Baker.

ROSCH, ELEANOR. 1973. Natural Categories. Cognitive Psychology 4.328–350.

——. 1978. Principles of Categorization. In Cognition and Categorization, ed. by Eleanor Rosch & Barbara B. Lloyd, 27–48. Hillsdale, NJ: Lawrence Erlbaum.

SALMINEN, TAPANI. 1997. Tundra Nenets Inflection, volume 2 of Publications of the Databank for Endangered Finno-Ugrian Languages. Helsinki: Suomalais-Ugrilainen Seura.

SAPIR, EDWARD, & VICTOR GOLLA. 2001. Hupa Texts, with Notes and Lexicon. In The Collected Works of Edward Sapir, ed. by Victor Golla & Sean O'Neill, volume XIV, 19–1011. Berlin; New York: Mouton de Gruyter.

SCHIFFMAN, HAROLD F. 1999. A Reference Grammar of Spoken Tamil. New York: Cambridge University Press.

SMEETS, INEKE. 2008. A Grammar of Mapuche. Berlin; New York: Mouton de Gruyter.

Stonham, John T. 2004. Linguistic Theory and Complex Words: Nuuchahnulth Word Formation. Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.

Sullivan, Thelma D. 1988. Compendium of Nahuatl Grammar. Salt Lake City: University of Utah Press.

Trommer, Jochen, 2001. Distributed Optimality. Universität Potsdam dissertation.

——. 2002. The interaction of morphology and syntax in affix order. In Yearbook of Morphology 2002, ed. by Geert Booij & Jaap van Marle, 283–324. Kluwer.

Vajda, Edward J. 2004. Ket, volume 204 of Languages of the World. Munich: Lincom Europa.

van den Berg, René. 1989. A Grammar of the Muna Language. Dordrecht, Providence: Foris.

van Driem, George. 1993. A Grammar of Dumi. Berlin; New York: Mouton de Gruyter.

Wali, Kashi, & Omkar N. Koul. 1997. Kashmiri. London; New York: Routledge.

Watkins, Laurel J, & Parker McKenzie. 1984. A grammar of Kiowa. Lincoln: University of Nebraska Press.

Weber, David John. 1989. A Grammar of Huallaga (Huánuco) Quechua, volume 112 of University of California Publications in Linguistics. Berkeley: University of California Press.

Wichmann, Søren. 2007. Popoluca de Texistepec, volume 27 of Archivo de Lenguas Indgenas de México. Mexico City: Colegio de México.

Williamson, Janis, 1984. Studies in Lakhota Grammar. University of California, San Diego dissertation.

Wurm, S.A. 1975. The Trans-Fly Stock. In Papuan Languages and the New Guinea Linguistic Scene, ed. by S.A. Wurm, number 38 in Pacific Linguistics Series C, 323–344. Canberra: Linguistic Circle of Canberra.

Appendix A

Language Survey Results

Language	**Pattern**	**Source**
Aleut	V-P-N	(Bergsland 1997: 84)
Basque	P-V-N	(Hualde & de Urbina 2003: 207)
Berber	P-V-(P.)N-G	(Abdel-Massih 1971: 190)
Cree	P(.T)-V.G-N	(Dahlstrom 1986: 28-29)
Cree	V.G-P-N.G	(Dahlstrom 1986: 16, 29)
Cree	P1-V.G-P1.N1-P2.N2	(Dahlstrom 1986: 40)
Dumi	V-P-N	(van Driem 1993: 143)
Dumi	V-P1.P2-N	(van Driem 1993: 130)
Dumi	V-P1.N1-P1.N1.P2-N2	(van Driem 1993: 133)
Fur	P(.G)-V-N	(Jakobi 1990: 92)
Georgian	P1-P2-V-N	(Hewitt 1995: 128-9, 133)
Georgian	P-V-P.N	(Hewitt 1995:133)
Georgian	P-V-N	(Aronson 1990: 41-45)
Halkomelem	V-P-N-P	(Galloway 1993: 192)
Hebrew	P-V-N.G	(Glinert 1989: 470)
Hebrew	P.G-V-N	(Halle 1997: 435)
Hixkaryana	P1.P2-V-N	(Derbyshire 1979: 148)
Hixkaryana	P-V-N	(Derbyshire 1979: 145)
Huave	P-V-N	(Kim 2008: 249)
Huave	V-P-N	(Kim 2008: 252)
Huave	P-V-P-N	(Kim 2008:252)
Hupa	N-P1.G-P2.G-V	(Sapir & Golla 2001)
Hupa	N-P.G-V	(Sapir & Golla 2001)
Hupa	N-P.N-V	(Sapir & Golla 2001)

Language	**Pattern**	**Source**
Iraqw	V-P-(P.)N	(Mous 1993: 159, 162)
Kalkatungu	V-P-N	(Blake 1969: 50)
Kalkatungu	Aux-P-N	(Blake 1969: 50)
Karuk	N-P-V	(Bright 1957: 64)
Karuk	(P1.)P2-V-N2	(Bright 1957: 64)
Karuk	N-P-V	(Bright 1957: 64)
Karuk	N-P-V-N	(Bright 1957: 64)
Kashmiri	V-G.N-P	(Wali & Koul 1997: 248)
Ket	P1(.G1)-P2.N2(.G2)-V-N1	(Georg 2007: 191), (Vajda 2004: 53ff)
Ket	P.N-V.N	(Georg 2007: 218-219)
Kiowa	P1-N1-N2-N2-V	(Watkins & McKenzie 1984: 117ff)
Kiwai	P-N-V	(Ray & Riley 1933: 46)
Kiwai	P-V-N	(Ray & Riley 1933: 47)
Lakhota	P.N-V-N	(Williamson 1984: 78)
Mapudungun	V-P-N	(Smeets 2008: 152ff)
Maricopa	N2-P1.P2-V.N1	(Gordon 1986: 21, 22, 148)
Mayali	P.(N.)-N-V	(Evans 2003: 261)
Mordvin	V-P2-N-P1.N1	(Abondolo 1982: 14)
Muna	P(.N)-V-N	(van den Berg 1989: 51ff)
Nahuatl	P1.N1-(P2.N2-)V-N1	(Sullivan 1988: 31)
Nanti	P1.G1-V-N-P2.G2	(Michael 2008: 260)
Nenets	V.N2.M-P.n-p.N	(Salminen 1997: 103-105)
Ngiyambaa	X-P-N	(Donaldson 1980: 124, 131)

Language	Pattern	Source
Nimboran	V-N1-N2-G2-P1	(Anceaux 1965)
Nimboran	V-N2-G2-N1-P1	(Anceaux 1965)
Nimboran	V.N-P(.G)	(Anceaux 1965: 93ff)
Nimboran	V-N-P.N	(Anceaux 1965: 85-86)
Nootka	V-P-N	(Stonham 2004: 67)
Nootka	N-V-P.M	(Stonham 2004: 83)
Nootka	V-N-P.N	(Stonham 2004: 97)
Quechua	V-N-P	(Weber 1989: 144)
Quechua	V-P-N	(Weber 1989: 96)
Quechua	V-N-P-N	(Weber 1989: 96, 182)
Tamil	V-P-G-N (?)	(Schiffman 1999: 116)
Tamil	V-P.G-N	(Schiffman 1999: 116)
Turkana	P.N-V-N	(Dimmendaal 1983: 120ff)
Turkana	P-V-N	(Dimmendaal 1983: 120ff)
Turkish	V-P-N	(Lewis 2000: 105, 107)
Turkish	V-P.T-N	(Lewis 2000: 105, 98, 123, 132)
Turkish	V-P.M-N	(Lewis 2000: 105, 98, 138)
Tzotzil	P-V-N	(Aissen 1987: 46-49)
Tzotzil	P-V-P.N	(Aissen 1987: 46-49)
Warlpiri	Aux-P-N	(Nash 1980: 172)
Yimas	P.N-P.N-V-N	(Foley 1991: 216)

Table A.1: Language survey results: Verb agreement

Language	Pattern	Source
Hupa	A-V.A	(Sapir & Golla 2001)
Kashmiri	T-V-T	(Wali & Koul 1997: 225)
Ket	T-V.T	(Georg 2007: 221)
Kiwai	P.T-V-T	(Ray & Riley 1933: 48)
Maricopa	P-V-P.A	(Gordon 1986: 27)
Mayali	P.(N.)T-V-T.A	(Evans 2003: 401)
Nahuatl	A-V-A.N	(Sullivan 1988: 50)
Nanti	M-V-M.Voice	(Michael 2008: 276)
Nootka	V-T-T	(Stonham 2004: 85ff)
Warlpiri	Aux-T V-T	(Nash 1980: 44)

Table A.2: Language survey results: TAM

Language	Pattern	Source
Berber	Neg-V.Neg	(Abdel-Massih 1971: 172-173)
Dumi	Neg.T-V-P.N.Neg.T	(van Driem 1993: 119)
Dumi	V-Neg.A-Neg	(van Driem 1993: 124)
Karuk	Neg-V-Neg.N	(Bright 1957: 64, 67, 137-138)
Maricopa	Neg-V-Neg	(Gordon 1986: 71-73)
Turkish	V-Neg-T.Neg	(Lewis 2000: 114)

Table A.3: Language survey results: Negation

Language	Pattern	Source
Basque	P.Pr-N	(Hualde & de Urbina 2003: 206-207)
Berber	P.Pr-G	(Abdel-Massih 1971: 35)
Dumi	P.Pr-N	(van Driem 1993: 82-84)
Fur	P.Pr-N	(Jakobi 1990: 92)
Georgian	P.Pr-N	(Hewitt 1995: 77)
Halkomelem	G-P.Pr-(P.)N	(Galloway 1993: 192)
Hebrew	P.Pr-N.G	(Glinert 1989: 51)
Hixkaryana	P.Pr-N	(Derbyshire 1979: 127-128)
Huave	P-(P.)Pr-N	(Kim 2008: 222)
Hupa	N-P-Pr	(Goddard 1905:)
Iraqw	P.Pr-P.N	(Mous 1993: 112)
Juang	P.Pr-N	(Dasgupta 1978:)
Kalkatungu	P.Pr-N-C	(Blake 1969: 39)
Kashmiri	P.Pr-N-G	(Wali & Koul 1997: 196)
Ket	P.Pr-N	(Georg 2007: 163)
Ket	P.Pr-G.C	(Georg 2007: 164)
Kiwai	N-P.Pr(-N)	(Ray & Riley 1933: 19)
Mapudungun	P.Pr-N	(Smeets 2008: 97)
Mayali	P.(N.)Pr-N	(Evans 2003: 261)
Muna	P.Pr-N	(van den Berg 1989: 68, 81)
Nahuatl	P(.N)-Pr-N	(Sullivan 1988: 36)
Nenets	P.C.Pr-p.N	(Salminen 1997: 131)
Ngiyambaa	P.Pr-N	(Donaldson 1980: 122-123)
Ngiyambaa	P.Pr-N-N.C / P.N.Pr-N.C	(Donaldson 1980: 122-123)
Ngiyambaa	P.Pr-N-P / P.N.Pr-P	(Donaldson 1980: 123)

Language	**Pattern**	**Source**
Quechua	P.Pr-N	(Weber 1989: 37)
Tamil	P.(G.)Pr-N	(Schiffman 1999: 59)
Tamil	P.Pr-G	(Schiffman 1999: 59)
Turkish	P.Pr-N	(Lewis 2000: 64)
Warlpiri	P.Pr-N	(Nash 1980: 169, 172)
Yimas	N.Pr1-P-N.Pr2	(Foley 1991: 111)

Table A.4: Language survey results: Pronouns

Appendix B

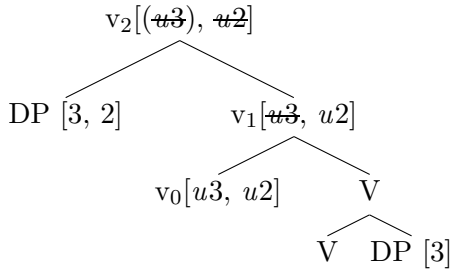
Testing the syntactic bias of Cyclic Agree

Chapter 4 introduced the theory of Cyclic Agree ((Béjar 2003), (Béjar & Rezac 2009)), which has been extremely influential for recent work on agreement. Developed to account for Person Hierarchy sensitivity, it also provides a unified treatment of intervention effects and has consequences for Case theory. Although the system focuses on the syntax, it also makes specific predictions about agreement morphology. Specifically, Béjar (2003) proposes that particular syntactic configurations favor the insertion of markers that fuse person and number features.

Under Béjar's system, crosslinguistic variation in Person Hierarchy sensitivity arises from the distribution of interpretable and uninterpretable agreement features. For instance, in Karuk the DO controls person agreement if it is a participant, otherwise S controls person agreement. Likewise the DO controls number agreement if it is plural, otherwise S controls number agreement. Béjar locates the P(erson)- and N(umber)-probes on v, deriving the preference for control by the internal argument (155). The P-probe is specified as [$u3$, $u2$]; when it encounters a third person goal the $u3$ feature matches and is valued, but the $u2$ feature creates an 'active residue,' licensing a second cycle of Agree to probe the external argument (Béjar 2003: 159–160).

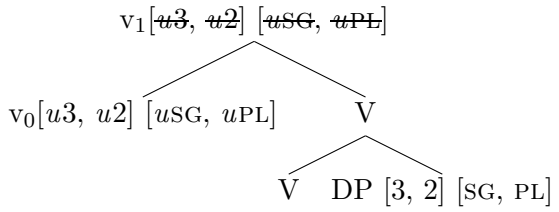
(155) Unchecked $u2$ on v_1 projects to probe the external argument[1] (N-probe not shown)

[1]The parentheses around $u3$ on v_2 in (155) indicate that the feature was checked and valued on a lower head.



On the other hand, if the P-probe encounters a first or second person DO then its uninterpretable features are fully valued, just as the N-probe is fully valued when it encounters a plural DO. This situation results in a single insertion site (v_1 in (156)) that is valued for both person and number: "a natural candidate for a vocabulary insertion rule that makes reference to both sets of features" (Béjar 2003: 160–161).

(156) P-probe and N-probe are fully valued on the same head



Generalizing Béjar's claim, we would expect the agreement systems of the world to show the following tendencies:

(157) Morphological predictions of characteristic probe type

a. low-ϕ (P-probe and N-probe on v)
Fused P/N markers when the internal argument fully values both probes (e.g., 1/2PL DO in Karuk) or when the external argument values both by second cycle Agree (e.g., 3SG DO in Karuk).

b. high-ϕ (P-probe and N-probe on T)
Fused P/N markers when the external argument fully values both probes or when the internal argument values both by second cycle Agree.

c. split-ϕ (P-probe on T and N-probe on v, or vice versa)
No fused P/N markers because P and N features are never valued on the same head.

d. double-ϕ (P-probe and N-probe on T, and P-probe and N-probe on v)
All fused P/N markers because second cycle Agree is never licensed.

To test the reality of this syntactic bias, I checked agreement patterns in 10 representative languages chosen from the study reported in chapter 2. The results of this check are summarized in Table B.1, in which the rightmost column indicates how

Language	**Macro-Area**	**Probe**	**Predicted?**
Dumi	South and Southeast Asia	double-ϕ	SG & PL only
Fur	Africa	high-ϕ	1 & 2 only
Georgian	Western and Southwestern Eurasia	split-ϕ	2 & 3 only
Huave	Central America	double-ϕ	3 only
Hupa	Western North America	double-ϕ	2 only
Ket	Northern and Central Asia	double-ϕ	no
Lakhota	Eastern North America	double-ϕ	no
Nanti	South America	double-ϕ	no
Nimboran	New Guinea and Oceania	double-ϕ	no
Warlpiri	Australia	double-ϕ	no

Table B.1: Survey languages checked for Béjar's syntactic bias

much of the agreement paradigm is consistent with the expectations laid out above. For instance, in Hupa (a double-ϕ language, as determined by its target-controller patterns) we expect by (157) to find only fused markers, but in fact we find person and number marked discontinuously in the first and third persons. In other words, Hupa satisfies the prediction in the second person only.

These results suggest that Béjar's predictions do not adequately explain the distribution of discontinuous agreement, looking across languages and grammatical persons.