PHI-SYNTAX: A THEORY OF AGREEMENT

by

Susana Bejar

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Linguistics
University of Toronto

Phi-Syntax: A Theory of Agreement

Doctor of Philosophy 2003

Susana Bejar
Department of Linguistics
University of Toronto

This thesis explores the formal mechanisms by which agreement relations are derived in syntax. The empirical focus is on some difficult cases of verb-NP agreement (Georgian, Algonquian and others) which present serious locality paradoxes when analyzed using standard mechanisms of syntactic agreement. The agreement systems in question have typically been taken to be symptomatic of a fundamental mismatch between syntax and morphology, however I treat them as evidence for a relativized formulation of conditions on Agree, under which they are in fact systematically isomorphic with well-defined derivational syntactic sequences. The claim is thus that surface complexity in these paradigms is best described in terms that are characteristically syntactic: long-distance dependencies, downwards locality/minimality, intervention effects and cyclicity (defined over syntactic operations). The analysis relies crucially on a fine-grained approach to formal phi (φ), where φ encode the grammatical properties of nominals that typically enter into agreement: person, number, gender. Three aspects of φ feature theory are developed: (A) the representation of implicational relations between φ features; (B) feature-theoretic formulations of the syntactic operations that enter into the establishment of φ agreement relations; (C) the notion of φ feature (de)activation, and its interaction with syntactic projection, and cyclicity. The theory of agreement that emerges allows for a unification of complex agreement systems — a diverse and robust group — with their simpler, more familiar counterparts.

Acknowledgements

My warmest thanks to Diane Massam, my supervisor. I have benefited greatly from her keen intuitions about language, linguistics and life in general. I 'm grateful for her generous habit of making time for long meetings, her judicious doses of encouragement, and her ability to penetrate and gleen content (and even value) in writing that had a tendency to be fast and painfully cryptic.

Likewise, I am thankful to the other core members of my committee — Elizabeth Cowper and Alana Johns — for the feedback, cooperation and support they provided during the writing process and the race to finish. It is with Elizabeth that I first poked around in the world of Georgian verbs, and her own work on inflectional feature systems has greatly enhanced the clarity of my own thinking. Alana is the person who originally turned me onto complex agreement systems when she innocently (?) suggested I read a paper by Carmack on Georgian. Her research on agreement asymmetries in Inuktitut has without question had a defining influence on the way I approach morphosyntax. I also thank Yves Roberge and Alec Marantz who came later to my committee and contributed to the defense with thoughtful and provocative feedback. It was a privilege to have Alec as my external examiner; the debt this work owes to his should be transparent. I am also thankful to him for facilitating a term that I spent at MIT in the fall of 2000.

A very special thanks to Milan Rezac for sharing my fascination with the phenomena explored in this thesis, and for an ongoing collaboration that has spanned years. Many of the ideas in this thesis took shape in the course of our discussions and joint work on related topics. Collaborating with a person like Milan, who has an uncanny ability to go further, faster, you either dig in your heels or stay on your toes. I have spent a lot of time doing both, and it has always been good exercise.

Elan Dresher and Keren Rice have also played an important role in this thesis, although it may not be obvious to outsiders. I drifted away from phonology over the years, but their research into the nature of feature systems lay the foundation for my thinking about morphosyntactic features. I'm grateful to both for being fantastic teachers, and for taking the time to discuss various feature-related or morphology-related topics with me. Elan has the additional distinction of being the maestro of the linguistics department band F-0 which has been a source of good fun for the last 5 years. And to Keren I owe additional thanks for extraordinary support she showed me when I first came to the department. Her attention was crucial to my transformation into a graduate student.

Of course, being a graduate student entails much more than 'just' writing a thesis. I first came to the University of Toronto in 1994. Three degrees and so many years later, I feel compelled to say something about the people who have shaped this experience. The beginning and end are always most salient, so I apologize to anybody from the middle years who may be unintentially overlooked.

As I was writing the thesis, a handle of people could be counted on to be putting in the same crazy schedule at Robarts that I was. Chiara Frigeni and Manami Hirayama were in the department at all hours, all winter, and their ever-presence made a huge difference, as did the small sustaining treats which would magically appear on my desk, as though left by elves ((including one not-so-small 5-cup rice cooker!).

Kenji Oda, would descend periodically from the 14th floor, usually to scold me for how messy my desk had gotten, but sometimes to take care of my chocolate cravings. Daniel Currie Hall was, as always, a constant companion. Daniel and I have been through all three degrees together. We were, in fact, the only two students in our first year of graduate school, and we've milked that relationship ever since. Milan, of course, could be counted on to pop by in the wee hours, once with hamburgers, always with something provocative to say about agreement.

Others, with saner work habits, may not have been such constant presences, but they provided a welcome counterpoint to routine. Sara Mackenzie never tired of letting it be known that it would be nice if I could make it out for some fun time, but ok that I never did. For crucial doses of kvetching, comic relief or commiseration, in and out of the department I thank Arsalan Kahnemuyipour, Alex D'Arcy, Bettina Spreng, Carolyn Smallwood, Gabriela Alboiu, Jack Chambers, Keir Moulton, Luba Butska, Naomi Cull, Nicole Rosen, Peter Avery, Ron Smyth, Suzanne Belanger and Taylor Roberts. I owe Mary Hsu and Ed Burstynsky big time, for crucial help at the very end in smoothing over some unforeseen administrative (and electrical) snafus. Likewise, I owe many thanks to Bill Forrest for his help with innumerable technical bumps in the road.

I don't want to overlook the people who first *were* the department to me. In the beginning there was David Bennett, Luba Butska, Jools Simner, and Lynn Macfie, the people with whom I learned to navigate after-hours life at Robarts Library. After them came several waves of people who made the department an especially good place to be, including David Beck, Trisha Causley, Julio Cebrian, Jun Mo Cho, Nila Friedberg, Wenckje Jongstra, Alexei Kochetov, Paivi Koskinen, Hoa Pham, and Carolyn Smallwood.

In fall 2000 I spent a wickedly fun term at MIT, and for that I thank in particular David Adger, Maria Cristina Cuervo, Daniel Harbour (I owe the catchier half of the title of this thesis to him), Stephanie Harves, Christian List and Michael Wagner. Also in the category of people from other places, I'd like to thank the following for feedback at conferences and useful discussions at various points over the years: Artemis Alexiadou, Elena Anagnostopoulou, Jonathan Bobaljik, Phil Branigan, Henry Davis, Elly van Gelderen, David Heap, Martha McGinnis, Lea Nash, Juvenal Ndayiragije Andrew Nevins, Charlotte Reinholtz, Halldor Sigurdsson, and Jochen Trommer.

My student life was made possible by funding from a SSHRC doctoral fellowship, as well as an Ontario Graduate Scholarship, U of T Open Fellowships, departmental research grants, and funds from the UTSC humanities division.

Finally, I'd like to thank people closer to home: my family; good friends along the way who I hope will now forgive me for having disappeared over these past few years; and most of all Abdel-Khalig Ali, who never disappeared at all.

Table of Contents

List of abbreviations

π	person
φ	phi-
#	number
[SMALLCAPS]	feature notation
> (as in x>y)	SU>DO
1 or [1]	1^{st} person
2 or [2]	2^{nd} person
3 or [3]	3^{rd} person
ACC	accusative
ADD	addressee
AGR	agreement/Agree
ANIM	animate
AOR	aorist
APPL	applicative
AUX	auxiliary
BPS	bare phrase structure
CSA	context-sensitive agreement
DAT	dative
def	defective
dflt	default
DIR	direct
DM	distributed morphology
DNC	dative nominative construction
DO	direct object
DU	dual
DUB	dubitative
EV	evidential
EVID	evidential
FF	formal feature
INAN	inanimate
INDIC	indicative
INDIV	individuation
INV	inverse
IO	indirect object
LI	lexical item
MASC	masculine
MIN	minimal
MP	minimalist program
NEG	negative
NEUT	neuter
NEUT	neutral
NOM	nominative
PART	participant
PASS	passive
PL	plural
PLD	primary linguistic data
POS	positive
PRES	present
PRET	preterite
PreV	preverb
Q	complex symbol
R	referring expression
SG	singular
SPKR	speaker
SU	subject
TCP	target-controller pair
TH	theme
TNS	tense
UG	universal grammar
VI	vocabulary item

Chapter 1 The Problem of Context-Sensitive Agreement

1.0 Introduction

This thesis explores the formal mechanisms by which agreement relations are derived in syntax. The empirical focus is on some difficult cases of verb-NP agreement that present serious locality paradoxes when analyzed using standard mechanisms of syntactic agreement. Canonical verb-NP agreement is typically characterized in terms of the structural position or grammatical function of the cross-referenced NP — hence the ubiquitousness of terminology like 'subject agreement' and 'object agreement'. Such characterization has prevailed despite the well-known fact that there are many languages with complex agreement systems for which it is inadequate.[1] In the languages that will be considered here, agreement is promiscuous, being sometimes with a subject, sometimes with an object (or indirect object) and sometimes with multiple arguments; the choice depends on φ-properties of co-arguments in the agreement domain, where 'φ' represents the formal properties of NPs that typically enter into agreement: person, number and gender (our concern will be primarily with the first two).[2] In general, these patterns have been considered to be special and unnatural for syntactic agreement mechanisms. Most

---

[1] Even in languages whose core agreement patterns conform to traditional characterizations, it is well known that non-canonical patterns arise systematically. Most linguists would agree that English has something like 'subject agreement' (weak though the paradigm may be). In all of the sentences in (i) the NP that *controls* agreement with the matrix verb is the syntactic subject of the matrix clause.[1]

i) a. Annabelle [seems/*seem [*t* to have misplaced her castanets]].
   b. Her castanets [seem/*seems [to have been misplaced *t*]].

But (i) is incomplete in important ways. Non-canonical controllers (i.e. controllers that are not subjects) arise systematically in a well-defined set of English sentences such as expletive constructions and locative inversion constructions, as in:.

ii) a. There are/?is thought to be three chambers in that cave.
    b. Out the window are/*is my hopes for a vacation this summer.

[2] The choice can be shown to occur independently of extenuating factors such as discourse or information structure. Such factors have been shown to be relevant in some cases (e.g. Bobaljik & Wurmbrand 1997, 2002), but I abstract away from them here.

previous approaches to this problem have either attributed the complexity to mechanisms outside of the core agreement system, or taken it as evidence for a fundamental mismatch between syntax and morphology, with the implication that agreement is not syntactic at all. The central claim of this thesis is that these ostensibly non-canonical agreement patterns are really manifestations of the same core syntactic mechanisms that give rise to their traditionally more familiar counterparts. I argue that these patterns reflect very general conditions on agreement that have hitherto not been articulated. The central claim is that so-called non-canonical agreement arises precisely in those environments where a canonical agreement relation fails to be established for independent reasons (Bejar 2000a, 2000b, 2000c, Bejar & Rezac 2003a, 2003b, Rezac 2003; cf. Bobaljik & Wurmbrand 1997, 2002). In effect, complex agreement patterns are a consequence of (anti)intervention effects. The conditions that properly characterize the environments in which these occur will be stated as restrictions on the core syntactic operations that enter into agreement: match and value. These restrictions are taken to apply uniformly across all agreement systems, being integral to the basic architecture of the computation. Given the formulation of the conditions proposed here, the complex patterns in these agreement systems can be shown to arise from the interaction of agreement operations, locality and cyclicity in syntactic computation. These interactions will be shown to be natural and predictable within the adopted syntactic framework, which is the minimalist framework of Chomsky 2000, 2001. The analysis assumes familiarity with minimalism, which is only briefly reviewed below.

The remainder of this chapter is structured as follows. Section 1.1 illustrates the core empirical phenomena explored in the thesis. Section 1.2 outlines the locality problems these data pose for syntactic approaches to agreement. Section 1.3 discusses the tension between syntactic and morphological theories of agreement, as well as relevant assumptions about morphology and the morphology-syntax interface. Section 1.4 provides a brief overview of previous syntactic approaches to agreement, and a summary of the framework adopted in the thesis. Section 1.5 outlines the proposals to be made in the thesis. Finally, section 1.6 provides a brief roadmap to the remaining chapters.

1.1 Context-sensitive agreement

To illustrate the problem of complex agreement as it is meant to be understood here,[3] compare the following descriptions of the core agreement patterns in canonical active transitive clauses for Spanish (a simple case) versus Nishnaabemwin (Algonquian), and Georgian (South Caucasian).

1) *Spanish:* a verb cross-references the person and number features of the syntactic subject.

2) *Nishnaabemwin* (based on paradigms in Valentine 2001)*:* a verb cross-references the person features of the subject if it is 2nd person, or if it is 1st person and the object is 3rd person, or if it is 3rd person and the object is a backgrounded 3rd person phrase. Otherwise, the verb cross-references the person feature of the object. If there has been person agreement with the object, then number

[3] Multiple-argument agreement systems like Nishnaabemwin and Georgian frequently manifest other forms of complexity as well, such as various incarnations of allomorphy (category neutralizations, positional variation, suppletion) patterns of syncretism and synthesis. See Noyer 1992 for discussion of these. For the most part these issues will not be dealt with in this thesis, although in chapter 4 I will show that the proposed solution to context-sensitivity has promising consequences for a theory of allomorphy. Another important source of complexity that will not be discussed here involves spurious valency alternations triggered by person hierarchy effects (Hale 2000, Nichols 2001, Bobaljik and Branigan 2003, Bejar and Rezac 2003b).

agreement must also be with the object. If, however, there has been person agreement with the subject, then number agreement is with the object only if the object is plural, otherwise it is with the subject.

3) *Georgian* (based on Aronson 1989, Hewitt 1995)*:* a verb cross-references the person feature of its object, unless the object is 3rd person, in which case the person feature of the subject is cross-referenced. Number is preferentially that of the subject, unless the subject is singular, in which case the number of the object is cross-referenced, unless the object is 3rd person.

The three systems schematized above are clearly very different from one another. The familiar one is Spanish, where agreement is canonically *controlled* by the syntactic subject.[4] In Nishnaabemwin and Georgian the controller shifts between subject and object. The choice depends crucially on ϕ-properties of other arguments in the agreement domain. For instance, in Georgian a 1st person subject will control agreement if the object is 3rd person, but not if the object is 2nd person. In Nishnaabemwin a 1st person object will control agreement if the subject is 3rd person, but not if the subject is 2nd person. I refer to this henceforth as *context sensitive agreement*. Context-sensitive agreement (CSA) systems are characterized by the appearance of intricate conditions on the identification of a controller within a paradigm, with crucial reference to the ϕ-properties of other arguments in the agreement domain. Differently put, context-sensitive agreement systems are characterized by *agreement displacement*: the same agreement 'slot' may cross-reference NPs with different grammatical functions, and a particular grammatical

[4] This usage of the term *control* is borrowed from Relational Grammar.

function may be tracked by more than one set of agreement slots. Thus, in these systems there is a many-to-many relation between arguments and agreement slots. CSA systems pose two main challenges that will be addressed in this thesis. The first is to explain how agreement relations are established within CSA systems. The second is to understand how CSA systems and non-CSA systems can both emerge from the same computational system.

1.2 The locality paradox

Agreement displacement challenges standard syntactic approaches to agreement because it creates the appearance of locality paradoxes. In the principles and parameters framework, the correlation between agreement and grammatical function is generally taken to follow from their sharing a structural locus in clause structure. So, for example, English and Spanish have subject agreement because the structural locus of the agreement (AGR)-relation in these languages is the same as the structural position that is identified with subjecthood. The choice of which argument will occupy (or at least enter into a relation with) a given AGR-position is mediated by locality; the 'closest' suitable NP in the domain of the AGR-position will enter into the relation, blocking the relation from being established with any further NPs (Rizzi 1990, McGinnis 1998, Chomsky 2000, 2001). Following Chomsky 2000, 2001, I take the domain of an AGR-position to be the syntactic object contained by its sister. The notion of 'closest' is defined in terms of highest c-command. NPs outside the domain of the AGR-position are not considered at all (see section 1.4). In the configuration shown in (4), NP2 and NP3 are in the domain of AGR^0, but NP1 is not. NP2 c-commands NP3, so by locality it will enter into AGR-relation, blocking NP3 from doing so.

4) [NP1 [AGR^0 [NP2 … [NP3 …]]]]

Systems like Nishnaabemwin and Georgian (as opposed to Spanish or English) are problematic for syntactic theories of agreement precisely because they resist straightforward characterization in terms of structural positions or syntactic relations. To illustrate further, consider just one of the asymmetries in Nishnaabemwin (Valentine 2001). In (5) the verb agrees with whichever argument is 2nd person, regardless of whether it is the subject (SU) or the direct object (DO). We can think of this as a problem of controller identification; in (a) the controller is the SU but in (b) it is the DO. (read *x>y* as SU>DO).

5) a. 2sg>1sg **g**-waabm-i *Nishnaabemwin* (Valentine 2001)
2-see-DIR.TH
'You see me.'

b. 1sg>2sg **g**-waabm-in
2-see-INV.TH
'I see you.'

Taking agreement to be a relation between some structural AGR-position and the closest NP in the domain of AGR, then regardless of where we situate AGR in the clause, the pattern in (5) will create problems for locality.[5] Two plausible schematizations of Nishnaabmewin clause structure are given below (precedence equals c-command):

6) … SU … AGR … DO …

7) … AGR … SU … DO …

[5] I mean AGR to designate any head that enters into an agreement relation, not necessarily the head of an AgrP.

In (6) the DO is the closest argument in the AGR domain, which predicts that it should always control agreement. In (7) it is the SU that is predicted to do so. Neither schematization predicts the attested pattern: that whichever NP is 2nd person will control agreement. (The facts are even more complicated when a wider range of argument configurations is considered).

We might suppose that both AGR positions are instantiated, and spelled out as *g-* if controlled by a 2nd person argument, otherwise being null. However, this raises new problems and ultimately begs the question. For instance, in combinations of [1]>[3] or [3]>[1] the verb manifests 1st person agreement morphology (8).[6] Given the availability of both [1] and [2] morphology in the Algonquian lexicon, the forms in (9) are plausible and their ungrammaticality requires explanation, and so we reiterate the problem of determining what conditions underlie the surface pattern.

8)	a.	1>3	**n**-waabm-aa	*Nishnaabemwin* (Valentine 2001)
			1-see-THEME[7]	
			'I see him.'	
	b.	3>1	**n**-waabm-igw	
			1-see-THEME	
			'He sees me.'	

---

[6] [1], [2], [3] denote first person, second person and third person, respectively.

[7] The THEME suffix is an obligatory component of the verb-word in Algonquian languages. The shape of it covaries with the φ-specifications of the verb's arguments. In particular, the choice of theme suffix is sensitive to whether the φ-specifications respect a 2>1>3 hierarchy, where features of the subject are more highly ranked than those of the object. See Bejar & Rezac 2003b for an account of this alternation.

9)	a.	2>1	***n-g**-waabm-i*	or	***g-n**-waabm-i*
			1-2-see-THEME		2-1-see-THEME
			'You see me.'		'You see me.'
	b.	1>2	***n-g**-waabm-in*	or	***g-n**-waabm-in*
			1-2-see-THEME		1-2-see-THEME
			'I see you.'		'I see you.

1.3 Morphology vs. Syntax

Syntactic accounts of similar paradoxes have taken two paths. One manipulates the base position of NPs, so that the controller is always base-generated in a local configuration with AGR, regardless of logical grammatical relations (e.g. for Algonquian: Bloomfield 1962, Perlmutter & Rhodes 1988, Piggott 1989, Rhodes 1994). The other posits syntactic parameterization of functional clause structure so that either the position or the φ-properties of AGR vary, depending on other structural configurations in the AGR-domain (for a variety of languages: Déchaine & Reinholtz 1998, Déchaine 1999, Brittain 2001, cf. Levin and Massam 1985, Bobaljik 1992, Bobaljik & Branigan 2003, Campana 1992, Johns 1992, 1993, 1996, Murasugi 1992, Bittner 1994, Bittner & Hale 1996, Rice and Saxon 1994, Hale 2002). But context-sensitive agreement effects have typically been taken to be symptomatic of a fundamental mismatch between syntax and morphology The leading idea in recent years has been that such cases should not be dealt with syntactically because they reflect the output of morphological operations/rules, not syntactic ones (Marantz 1991, Anderson 1992, Halle & Marantz 1993). Under this view, the syntax does not rule out agreement with either controller. All agreement relations are, potentially, made available to the derivation. The appearance of context-sensitivity arises from morphological mechanisms which suppress the expression of certain agreement relations and license the expression of others. Typically, suppression is accomplished by either of: (A) morphological slot competition (e.g. Bonet 1991, cf. Perlmutter 1971); or (B) morphological rules/operations/filters (e.g. Anderson 1992, Halle & Marantz 1993, McGinnis 1999, Stump 2001. See also Noyer 1992, Bobaljik & Wurmbrand 1997, 2002). For example, Halle & Marantz analyze a pattern analogous to (5) in Potawotami (Algonquian) as follows: the SU and the DO both enter into agreement relations with separate AGR nodes, but the two AGR nodes fuse (by a morphological operation) into a single vocabulary insertion site which forces the complementary distribution between subject and object agreement morphology.[8] The privileging of 2nd person agreement is a result of either intrinsic or extrinsic ordering of vocabulary insertion rules, such that the rule that inserts 2nd person agreement morphology is ordered before the rule inserting 1st person agreement morphology, thus blocking expression of non-2nd person agreement in (5).[9]

A central claim of this thesis is that, despite appearances, there is no morphology-syntax mismatch in these cases. I argue that the surface complexity of these paradigms is the outcome of (anti)intervention patterns and their derivational consequences. The patterns are therefore best analyzed in terms that are characteristically syntactic: long-distance dependencies, downwards locality/minimality, intervention effects and cyclicity (defined over syntactic operations). I treat the complexity as evidence for a relativized

[8] Halle & Marantz consider what I have labeled AGR to be clitic nodes rather than agreement. I abstract away from this distinction here.

[9] The Algonquian and Georgian cases are now famous testing grounds for theories of morphology. Previous works with detailed analyses of both languages include Anderson 1992, Halle & Marantz 1993, Stump 2001. See also McGinnis 1999, Trommer 2002.

formulation of the identity criteria that govern the establishment of the AGR-relations. The conclusion is that the morphological systems in question are systematically isomorphic with well-defined derivational sequences in the syntactic computation. Isomorphy between syntax and morphology is not, in and of itself, an objective of the thesis. The claim is *not* that there is perfect isomorphy between syntax and morphology. Mismatches between the two components are undeniably characteristic of the interface, since it cannot be maintained that there is a one-to-one mapping between syntactic positions and vocabulary insertion sites (see Noyer 1992 for an excellent overview of the kinds of mismatches that arise). The objective is to further understand the *limits* of isomorphy, so that questions about the mapping between these components of grammar can be formulated all the more precisely. My claim is that CSA-effects have incorrectly been grouped amongst the mismatch phenomena, and that removing them from this group provides a clearer picture of the morphology-syntax interface.

Tension between syntactic and morphological explanation is to be expected in work on agreement, a phenomenon that straddles sentence-formation and word-formation. The challenge of differentiating syntactic patterns from morphological ones is especially difficult in this case. The criteria by which this is to be done are not well understood yet, and the usual measures by which theories are compared (empirical coverage, restrictiveness, predictive power, etc.) are often difficult to evaluate. In this thesis the choice is for syntactic explanation, not because the data cannot be accounted for morphologically, but because the syntactic account is able to capture and make predictions for a wide range of data using mechanisms that are natural within the syntactic framework, without appealing to problem-specific rules/operations. Conversely, the data shed new light on the nature of the conditions governing the establishment of syntactic dependencies in the computation.

Certain assumptions about the morphology-syntax interface are crucial to my analysis. I assume a morphological component after syntax, along the lines of Distributed Morphology (Halle and Marantz 1993). In distributed morphology (DM) the formatives of syntax and morphology are intimately connected. Word-formation is a function of the manipulation of syntactic structures in the computational component as well as in a post-syntactic morphological component. Vocabulary insertion is performed on the terminal nodes of syntax after all such manipulation has taken place. These syntactic and morphological operations derive the terms that would correspond roughly to lexical items (LIs) (in the sense of Chomsky 1995, 2000, 20001), which are the bundles of formal features on terminal nodes that are replaced by vocabulary items. I adopt the core assumption that syntactic and morphological operations play a crucial role in redistributing features, prior to vocabulary insertion. Thus, in this thesis I assume a crucial distinction between lexical items (LIs), the objects that enter into the computation, and vocabulary items (VIs), which replace LIs by the vocabulary insertion operation (an operation of the morphological component).

Given these assumptions, the problem of context-sensitive agreement can be seen as a problem of feature redistribution and the proposals developed below will hopefully contribute to our understanding of the syntactic principles and operations that constrain and explain this.

1.4 Syntactic agreement

There is a long tradition in transformational grammar (and certainly in traditional grammar) of treating agreement as a syntactic phenomenon. Early rules of agreement were thought to have much the same substance as phonological rules of assimilation (Chomsky 1965: 174-175).

10)

			+N
	□ Gender		□ Gender
Article —>	□ Number	/___ …	□ Number
	□Case		□Case

A rule like (10), which results in agreement between the □-features of a determiner and a noun, was thought to be a transformational rule of the usual kind (in that it rewrites the terminal symbols of a phrase marker), except that it introduces specified features instead of only non-lexical formatives (N, V, etc.). But although (10) successfully formalizes the fact of agreement between two syntactic terms, it sheds no light on the question of how the relation between the two terms in question is established. This information is stipulated by the structural description of the rule. If we take agreement to be a relation established between a *target* (the term whose feature structure is rewritten by the operation) and a *controller*, then how is this target-controller pair (TCP) established in the first place? As obvious as the question is, it has not been given much consideration in syntactic theory, most likely because, until recently, it has been assumed that agreement is just a reflex of other syntactic relations, which were thought to establish the TCP and create a local configuration in which agreement could occur. This is not to say that agreement has been ignored. Hypotheses restricting the structural configurations in which agreement can be established emerge throughout the 1980s: Chomsky (1986) suggests heads agree with either their specifiers or their complements; Koopman (1987, 1996) proposes that all agreement relations are spec-head agreement, a restriction which is taken up in early incarnations of the minimalist framework (Chomsky 1991, 1993, 1995) and other work (Chung 1998). But such structural restrictions only exacerbate the problem of context-sensitive agreement.

Chomsky 2000 represents an important turning point in the syntactic formalization of agreement. First, the trigger of A-movement is reformulated so that the need to check φ-features is itself the impetus behind the establishment of a syntactic relation between a target and an NP. Second, the ability to establish an agreement relation is dissociated from NP-movement. φ-features of an NP may be checked in situ, making long-distance agreement the simplest case, with NP-movement arising as a reflex of the checking relation for reasons having to do with whether the target must project a specifier.[10] In this framework, agreement ceases to be a reflex of (or diagnostic for) some other relation and is instead a core syntactic operation: Agree. Context-sensitive agreement systems still pose a problem for locality within this framework, but novel approaches to the problem are made possible by the new autonomy of the agreement relation, and the flexibility of long-distance agreement. The theory of Agree is thus the starting point for the proposals to be made in this thesis.

In Chomsky 2000, 2001, an AGR-head is a core functional head (v^0, T^0, C^0) with unvalued φ-features. φ-features are assumed to be inherently valued and interpretable on N/D but not on AGR-heads. The uninterpretability of φ on AGR-heads is the basis for the Agree relation. By hypothesis, uninterpretable φ must be eliminated by the Agree relation

---

[10] Why some targets require specifiers and others do not is poorly understood. The requirement is formalized by a diacritic EPP feature on the target

in order for the derivation to converge. Agree establishes a TCP between the AGR-head and a suitable NP, with the result that the φ-values of the NP are assigned to the unvalued φ of the target. The target's φ-features, once valued, are marked for deletion, but they are not deleted immediately; deletion must be postponed at least until the derivation reaches PF, where the valued φ will serve as instructions for vocabulary insertion. In other words, the set of TCPs that is reflected in the morphological agreement system is established by the syntactic computation.

In effect, Agree consists of a sequence of three procedures: (i) probe (p): the initiation of a search for a controller (the *goal*) with valued (p) ; (ii) match: the evaluation of whether or not an object in the search space of the probe is a possible goal (i.e whether it has interpretable features that can potentially value the probe); (iii) value: the assignment of a value to the probe. Superiority effects effectively follow from conditions on match. Chomsky (2000:122) formulates these conditions as follows:

11) Matching is a relation that holds of a probe P and a goal G. Not every matching pair induces Agree. To do so, G must (at least) be in the domain D(P) of P and satisfy locality conditions. The simplest assumptions for the probe-goal system are:

(i) Matching is feature identity

(ii) D(P) is the sister of P

(iii) Locality reduces to 'closest c-command'

Thus D(P) is the c-command domain of P — what I have been referring to as the AGR-domain — and a matching feature G is *closest to* P if there is no matching G' in D(P) such that G is in the c-command domain of G'.

Agree for φ-features is just one consequence of a matching probe. Overt movement takes place if there is a matching goal and the category with the probe also has an EPP feature. Another consequence is case assignment to the controller, by the target; nominative is a reflex of Agree with T^0, and accusative with v^0. Only NPs without case are 'active' in the sense that they are able to value a probe (this is the *active goal hypothesis* of Chomsky 2000, 2001). What is generally assumed in standard implementations of the framework is that the core agreement system of a language has the underlying architecture in (12) where, in a transitive clause, the derivation includes two AGR-heads, each triggering one instance of Agree, yielding two sets of TCPs: <v^0, XP_{DO}> and <T^0, YP_{SU} >, where XP_{DO} receives accusative case, and YP_{SU} receives nominative.

12) T YP_{SU} v XP_{DO}

φ-[11] φ φ- φ

Adopting the strong thesis that the problem of controller identification reduces to the establishment of TCPs, then the extent of cross-linguistic variation in transitive agreement paradigms should be restricted to surface differences with respect to the lexical inventory available for spelling out the two Agree relations established by the computation. This cannot account for CSA effects in languages like Georgian and Nishnaabemwin. The locality paradoxes of previous frameworks still present themselves in the minimalist theory of Agree. However, the Agree hypothesis is constructed in part to cope with non-canonical agreement configurations of a certain kind (see footnote 1),

[11] The *n*-dash after 'φ-' denotes uninterpretability.

and I will exploit this in the following chapters to show that a natural account of CSA effects is made possible by this framework. Agree for φ-features is a consequence of a matching probe, subject to the restriction that the NP must not have been previously assigned Case. If the closest matching goal in the domain of P has already been assigned Case, it creates a *defective intervention effect*. The inactive NP successfully matches and halts the probe, but because it is inactive it cannot value. The result is default agreement (typically default agreement morphology is indistinguishable from 3rd person singular agreement morphology) *unless* the intervener undergoes movement (EPP) as a result of the defective AGR-relation (Chomsky 2000:131, Anagnostopoulou 2003). If it does so, then an AGR-relation may be established with a lower goal (a non-canonical controller).

13)	a.		________			NP1 matches but fails to value
			AGR	NP1	NP2	
	b.	NP1	AGR	*t*	NP2	NP1 displaces
	c.		________________			NP2 controls AGR
		NP1	AGR	*t*	NP2	

This is illustrated in the following pair of sentences from French. Unlike English, raising in French is prohibited across a dative experiencer NP (14). The dative NP is rendered inactive by inherent Case, and is thus a defective intervener. However, if the dative experiencer is a clitic (14) then for independent reasons it is able to displace. This allows a second probe; the embedded SU *Margot* thus enters into an Agree relation with matrix T^0 that results in raising-to-subject.

14)	a.	Margot	semble	(*à Francois) [*t* être intelligent]				
		M.	seems	(*to Francois)	to.be	intelligent		
		'Margot seems (*to Francois) to be intelligent.'						
	b.	Margot$_j$	lui$_i$	semble	t_i	[t_j	être	intelligent]
		M.	3.SG.DAT	SEEMS			to.be	intelligent
		'Margot seems to him to be clever.'						

Similar analyses have been given to account for non-canonical agreement in expletive constructions and locative constructions, of the sort noted in footnote 1. My claim is that these non-canonical patterns are, in essence no difference from those exhibited by languages like Georgian and Nishnaabemwin. Derivational consequences of failure to Agree with the canonical controller result in the establishment of a non-canonical target-controller pair. However current articulations of Agree cannot capture the set of environments in which agreement fails for languages like Georgian and Nishnaabemwin. In this thesis I introduce a new kind of intervention effect, as well as an anti-intervention effect, to capture the set of environments in which context-sensitive agreement arises.

1.5 Summary of main proposals

The proposals presented in the thesis address three main sets of questions: (A) what are the syntactic mechanisms by which locality paradoxes can be accounted for?; (B) what are the conditions on Agree such that agreement fails in precisely those environments where CSA-effects arise; and (C) what accounts for the striking cross-linguistic differences between agreement systems, especially between CSA and non-CSA systems?

The dependency between a target and a controller is a relation between features, not categories or positions. A prerequisite for establishment of the target-controller pair is pre-existing compatibility in the feature specification of both members of the pair, so that the controller is able to meet the criteria for match and value. By definition, the criterion for match is feature identity. Properties of the target are crucial to the determination of agreement, since the identity criterion by which locality is computed are relativized to the specification of the probe. This is the essence of relativized minimality (Rizzi 1990). Two probes (p) and (q) — both hosted by the same target — may well have different locality effects, depending on the properties of arguments in the AGR-domain. In (15) locality plays out differently for each probe. With respect to probe (p*)*, the highest NP in the AGR domain is XP, which blocks the establishment of an AGR relation with YP which is also specified for *p* and so, in principle, would be able to satisfy match. With respect to probe (q) the highest suitable NP in the AGR-domain is YP. XP is not specified for *p* and so cannot halt the probe. This is what I will refer to as an anti-intervention pattern, but it is really just the usual implementation of locality. It follows straightforwardly from the formulation of match given in (11).

15) AGR NP1 NP2
p ———— *p* *p*
q ———————— *q*

Canonical agreement patterns result when there is a one-to-one ratio of AGR-heads and arguments, *and* the highest NP in the AGR-domain successfully matches and values a probe. CSA effects (i.e. agreement displacement) arise as a consequence of failure to match or value. This is the central proposal of the thesis.

Failure to match or value occurs when conditions on these operations are not met. This failure may simply result in default agreement, but it also has the potential feed the establishment of Agree relations in 'non-canonical' configurations. This is what creates the appearance of locality paradoxes, which come in two varieties that are schematized below. The underlined YP represents the controller of the AGR-relation. (This convention will be used throughout the thesis).

16) AGR XP <u>YP</u>

17) <u>YP</u> AGR XP

(16) is paradoxical because it appears that the controller is not the closest element in the AGR-domain. (17) is paradoxical because the controller appears to be outside the AGR-domain. My claim is that, despite appearances, there is no point in the derivation is locality not implemented in the usual way. At the stage when the AGR-relation is established normal locality conditions apply.

One derivational route to the first paradox (16) is the basic intervention pattern described in the previous section, where match succeeds and value fails, but the intervener displaces with the result that an NP that is lower in the AGR-domain is then the closest NP in the AGR-domain. While the intervention pattern creates the appearance that an AGR relation is being established with a nonlocal NP, in fact this is not the case because an intermediate stage in the derivation has altered the locality configuration. And even simpler derivational route to (16) is the anti-intervention pattern schematized by probe (q) in (15). In either case, the core claim is supported: non-canonical agreement is possible only if Agree with the canonical argument fails

The second paradox (17) is also possible to resolve in a fashion that is natural within the framework and consistent with the core claim. In this case, if Agree fails with the canonical controller, then conditions are created for an expansion of the AGR-domain. By hypothesis, the canonical controller in (17) should be XP, which is the only element in the AGR-domain. If XP fails to Agree, the probe is not valued or deleted and so in principle remains active. Rezac (2003) argues that projection of the AGR-category (following merger of the specifier YP) effectively expands the AGR-domain. The proposal takes seriously the idea that the label of a phrase is just the head of the phrase. A formalization of this view results in projection being indistinguishable from merger of a new occurrence of AGR, this time with YP in its domain. Since the probe is still active, a target-controller pair can be established with YP. Here again, at the point when the Agree relation is established, normal locality conditions apply. Thus, the schematization in (17) is really (18), where YP is the closest NP in the AGR-domain.

18) AGR YP AGR XP

(Anti)intervention patterns and cyclic expansion of the AGR-domain are mechanisms which I assume are made available by the computation. These are the mechanisms that I take to underlie the appearance of locality paradoxes and context-sensitive agreement effects in the languages that will be examined in the coming chapters. However all of them rely crucially on an initial failure of agreement. The basis for this initial failure is only partly accounted for in the standard articulation of the Agree hypothesis. Therefore, the foundation of the analyses that follow consists of a reformulation of conditions on match and value. This reformulation, in turn, hinges on proposals for a new theory of formal ϕ-features.

To illustrate the inadequacy of the standard formulation of Agree, consider the anti-intervention pattern of probe (q) in (15). This pattern cannot be derived under the received view that q=φ, where φ is a bundle that contains all the φ-features of an NP. Assuming both XP and YP are NPs, then both elements in the AGR-domain have inherent φ-features that enable then to satisfy match, and no anti-intervention effect should ensure. As far as φ-agreement goes, the only case where the anti-intervention pattern can be implemented is the special case where XP is absent altogether, in which case it is trivially true that there will be no XP intervener blocking the AGR-relation with the lower element YP. This is, of course, just the standard configuration for passive and unaccusative structures, which we can think of as the most familiar of 'non-canonical' controllers. Likewise, the failure of XP to agree in the cyclic expansion pattern (18) can only occur in unergative-type structures where there is no XP internal to the AGR-domain on the initial cycle of Agree (before projection triggers a second cycle). Yet languages with context-sensitive agreement provide evidence that it must be the case that anti-intervention and cyclic expansion occur when XP is present.

Our first challenge therefore resides in specifying (q) of probe(q) when XP is neither inactive nor absent. I argue that this can be done in a straightforward manner if the operations match and value are sensitive to a finer grain of feature structure than is generally entertained. This thesis proposes a decomposition of formal features, with special attention to intrinsic entailment relations between them. A theory of φ-features is proposed, with three main components. The first exploits φ-feature representations as a means of encoding natural classes and dependencies. I show that the received view of φ-features as unstructured bundles of features is an oversimplification. There are systematic asymmetries in the ways different φ-combinations behave syntactically. In particular, certain φ-specifications successfully satisfy conditions on match and value, whereas others fail to do so, thereby triggering (anti)intervention effects and cyclic expansion. The asymmetries in question can be modeled straightforwardly using contrastive and logical underspecification, and structural hierarchies/geometries (Avery & Rice 1989, Dresher, Piggott & Rice 1994, Dresher 2000, Hall 2000 and references therein). Building on proposals made by Harley and Ritter (2002 and earlier work) I take formal features to be privative subcategories which enter crucially into intrinsic entailment relations with one another; these entailment relations will be the basis for the reformulation of conditions on match and value. The second component of the theory of φ-features proposed here is thus a feature-theoretic formulation of match and value, the syntactic operations that enter into the establishment of φ-agreement relations. The fine structure of formal features will be argued to play an integral role in the implementation of match and value. I will show that by enhancing the sensitivity of these operations to a detailed level of feature structure, it is possible to formulate new generalizations about agreement, shedding light on a range of data that have previously resisted integration into a generalized theory of agreement.

The final component of the theory of features addresses the problem of variation. The contrast between languages with and without context-sensitive agreement is a striking one, and there are also significant differences in the kinds of context-sensitive agreement patterns different languages exhibit. This raises questions that are especially interesting within the context of the minimalist framework, for which agreement is a core operation in the syntactic computation. One central minimalist thesis is that the computation is invariant across-languages, the locus of cross-linguistic variation being the lexicon. This premise is seriously challenged by the existence of these striking cross-linguistic differences between agreement systems, which have resisted a unified syntactic treatment. By demonstrating that context-sensitive agreement patterns are in fact manifestations of the core architecture of agreement, the perceived challenge posed by complex agreement systems is turned on its head; these systems turn out to provide striking evidence in favour of the existence of a uniform computational system because the attested patterns are precisely those that are natural within the overall theory of Agree (modulo the reformulations proposed here). I argue that cross-linguistic uniformity in the computation of agreement resides in the formulation of the core operations: match and value. The formulation of match provides a uniform basis for the computation of locality. The formulation of value provides a uniform basis for the determination of (anti)intervention effects. The locus of variation between languages resides in the specification and distribution of uninterpretable φ-features on AGR-heads. In other words, the locus of variation resides in the inventory of formal features and in the (perhaps) selectional choices that govern their initial distribution in a derivation. I assume that the initial selection of formal features is sensitive to the morphological contrasts encoded by the vocabulary of a language, so for example languages with a larger pronominal inventory will make a richer selection of formal φ-features. The principles that restrict the initial distribution of uninterpretable φ-features remain obscure, but whatever they are, I assume that they permit variation between languages with respect to the specification and distribution of AGR-heads. The thesis focuses on three main points of variation: (A) the number of AGR-heads in a clause; (B) the position of AGR-heads in a clause; (C) the degree of ϕ-specification of AGR-heads.

Point (A) affects the degree to which arguments appear to compete for agreement slots. If there is only one AGR-head in a clause, this will have derivational consequences in sentences with multiple arguments. Point (B) determines locality preferences, since the position of an AGR-head establishes the AGR-domain. Point (C) is crucial to the implementation of conditions on match and value, which will be formulated in terms of feature-theoretic notions of entailment, rather than in terms of identity.

The main result of the investigation is a unification of complex agreement systems — a diverse and robust group — with their simpler, more familiar counterparts. The overall contribution of the thesis is a theory of agreement. Since George and Kornfilt (1981), there have been several influential formalizations of agreement in the Principles and Parameters (P&P) tradition of generative syntax (Chomsky 1981, Kayne 1989, Pollock 1989, Schütze 1997, Chomsky 2000, 2001, among others). However it is fair to say that these have not been so much theories of agreement *per se*, but rather theories about the role of agreement in various other syntactic phenomena, such as word order, Case, and the establishment of syntactic domains. The framework developed here provides useful, and generalizable tool, for the exploration of agreement in particular and, feature-driven syntactic dependencies in general.

1.6 Outline of the thesis

The following chapters are organized as follows. Chapter 2 introduces the theory of features that is crucial to subsequent proposals. Consequences for the computation of agreement are outlined, and the implications for conditions on match and value are explored. This chapter essentially lays out all of the mechanisms necessary for the analysis of the core context-sensitive agreement patterns.

Chapter 3 applies the mechanisms developed in the previous chapter to various case studies, with greatest attention to Georgian and Nishnaabemwin. Other languages are discussed briefly in the context of exploring the typological implications of the main proposals. These languages include Abkhaz, Basque, Choctaw and Dakota. Three main points of variation between agreement systems are reviewed, with special emphasis on the contrast between languages with differing degrees of feature (under)specification, and different AGR-positions. A full set of configurations predicted by the proposals are schematized in Appendices 1 to 4.

Chapter 4 explores consequences for a theory of suppletive allomorphy and other special morphological forms. Agree affects φ-specifications on AGR heads. Projection of AGR creates new insertion sites with φ-specifications that will very depending upon the success of Agree on a previous cycle. These two facts combine to make predictions about the distribution of environments in which suppletive allomorphy occurs. Failure of Agree with cyclic expansion is predicted to create new conditioning contexts for allomorphy. It may also split features across terminal elements, creating alternations between fused and non-fused forms, where the fused forms arise in environments where cyclic expansion has not occurred, and non-fused forms occur in environments where it has. These predictions are shown to be supported by allomorphy patterns in Karok (a Hokan language) and Mordvinian (a Uralic language).

Chapter 5 examines consequences for Case theory and explores the possibility of unifying defective intervention effects with general conditions restricting occurrences of features. It ends with concluding remarks on the thesis.

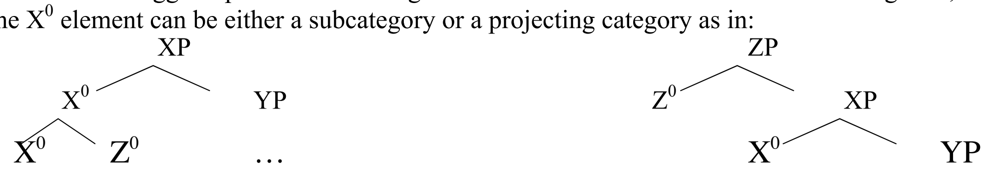
Chapter 2 Features, Underspecification and Locality

2.0 Introduction

In this chapter I explore the properties of formal φ-feature and consequences these have in syntactic computation. The status of formal features is somewhat underdefined in current theory. Minimally, they are diacritics which signal that a syntactic object has the property of being able either to trigger, or enter into, a relation R. (For agreement, the choice of φ as such a diacritic is not unreasonable, given that the AGR-relation tracks φ-feature values.) But traditionally features are also a formalization of morphosemantic categories, or subcategories (in the sense of Chomsky 1965). I will refer to both types of entities as features (also sublabels or subcategories following Chomsky 1965,1995), so the set of relations R potentially includes not only unbounded dependencies but also selectional relations (and of course, instructions for vocabulary insertion). Ontologically, it is far from obvious what the correct characterization of such diacritics is (or even what is at stake in asking this question). Features are the atomic elements from which larger syntactic objects (LIs) are assembled, but this is consistent with more than one ontological possibility. If we take LIs to be X^0 categories, one possibility is that a complex LI is simply constructed by concatenation of X^0s, in other words the sublabels of a head H are themselves X^0s (perhaps assembled by Merge).[1] Another possibility is

---

[1] This suggests possible cross-linguistic variation in the distribution of subcategories, such that the same X^0 element can be either a subcategory or a projecting category as in:



that sublabels are truly a different kind of formative (in which case the question of how they are assembled into larger objects is wide open).[2] It is not clear to me how to distinguish between these options at this point, but the differences are potentially relevant to the formulation of the mechanism by which features are shared or assigned in syntax. Before this larger questions can be addressed, a clearer understanding of the properties of features needs to be gained. What formal relationships exist between them? What are their structural properties? What conditions constrain their assembly into larger units? How is the inventory of features in a language determined?

Many of these same questions were actively engaged in early generative work when the atomization of syntactic categories was first explored (Chomsky 1951, 1955, 1956, Stockwell & Schachter 1962, Schachter 1962, Bach 1964, cited in Chomsky 1965:79). In *Aspects* (Chomsky 1965) Chomsky notes that early attempts to encode the distribution of syntactic features using "branching" re-write rules (on par with phrase-structure rules) were abandoned because subcategories were determined not to be strictly hierarchical, but rather to involve cross-classification.[3] This is not so clearly true for contextual/selectional features, but for those with morphosemantic analogues it is more apparent. For example, all combinations of the subcategories [N], [proper], [common], [human], [nonhuman] are possible, a fact that is not appropriately captured if they are introduced by branching rules. This problem motivated the importing (from phonology) of *complex symbols*, each complex symbol being a matrix of specified syntactic features.

---

[2] This suggests that no sublabel alone can ever constitute an X^0. Some formative must be added to it for it to have status as a syntactic object in its own right.

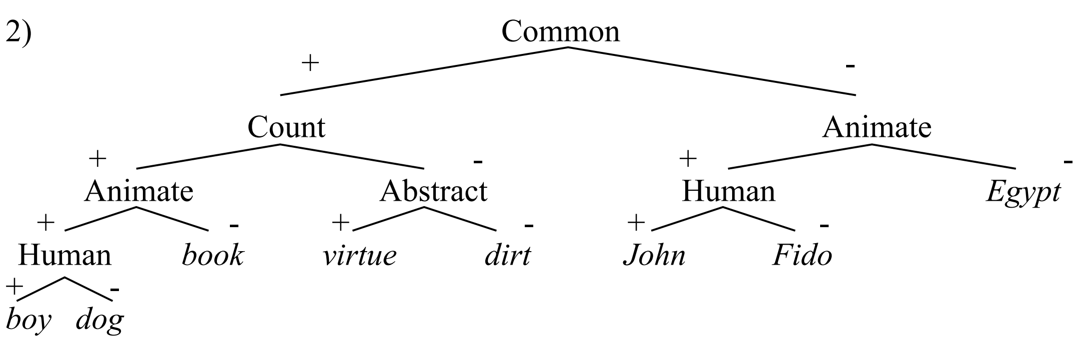
[3] Chomsky credits G.H. Matthews with being the first to work through some of these issues in the course of working on a generative grammar of German (1957-1958), but no publication is cited.

Complex symbols are mapped to terminal elements (N, V, etc.) in a phrase-structure tree by a set of replacement rules (subcategorisation rules). For example:

1) a. N —> [+N, ±Common]
   b. [+Common] —> [±Count]
   c. [+Count] —> [±Animate]
   d. [-Common] —> [±Animate]
   e. [+Animate] —> [±Human]
   f. [-Count]—> [±Abstract] (Chomsky 1965:82, (20))

Rule (a) asserts that any terminal N is to be replaced by either [+N, +Common] or [+N, -Common]. Subsequent rules follow the same conventions as phonological rules. For example, rule (e) asserts that any complex symbol Q that is already specified as [+Animate] will be replaced by a complex symbol containing all of the features of Q as well as one of the feature specifications [+Human] or [-Human]. The rules in (1) map onto a branching diagram, each path of which defines a possible category of LI for this grammar (illustrated here by the italicized VIs).[4]

2)



As written above, the subcategorisation rules encode intrinsic semantic implicational relations between subcategories. For this very reason they make predictions about possible categories in a grammar, implicitly ruling out incoherent combinations such as *[+Human, -Animate] for (1). In *Aspects* the central concern was to model

[4] Whether or not such rules *must* map onto such a branching diagram is left open in *Aspects*.

distributional and selectional restrictions that could not be adequately captured by context-free rewrite rules. φ features are not extended the same treatment, though it is acknowledged that they could be (notable related work includes Jakobson 1936, 1971, Forchheimer 1953, Greenberg 1963, Benveniste 1966).

After *Aspects*, the modeling of selectional restrictions and subcategorisation moved away from this formalism and the encoding of implicational relations amongst features ceased to be a focus in generative syntax (with the notable exception of monostratal phrase structure grammars such as Generalized Phrase Structure Grammar (Gazdar et al 1985) and Head-Driven Phrase Structure Grammar (Pollard and Sag 1994)). I claim that formal features — in particular φ-features —must be decomposed in a manner reminiscent of the subcategorisation rules of *Aspects*, in the sense that the intrinsic entailment relations between them must be made specific. The past decade has already seen a resurrection of investigation into these relations, with a focus on discovering constraints on rule application in the morphological component (Bonet 1991, Noyer 1992, Harley 1994, Nevins 2002) and also constrained characterizations of morphological inventories across languages (Harley & Ritter 1997, 1998, 2002, Bejar & Hall 1999, Heap 1999, Cowper 2003a,b, Cowper & Hall forthcoming, Nevins 2002). I will argue here that these relations are not just useful to restricting inventory shapes; syntactic computation is also sensitive in part to such feature-theoretic relations (see also Bejar 1999, 2000a, 2000b, 2000c, Branigan & Mackenzie 2000, Dechaine 1999). The particular claim I make here is that these relations are in fact central to the computation of locality. The relevance to locality plays out in two ways: (A) the evaluation of (anti)identity vis a vis match must be formulated not in terms of strict match between two features F and G, but in terms of the existence of an entailment relation between F and G; (B) the determination of whether or not match will result in value will also depend on a particular entailment relation existing between the target and the controller. Whether or not value succeeds has potentially dramatic consequences for subsequent steps in the computation, feeding or bleeding subsequent locality configurations with respect to the establishment of AGR- relations. We will see that by relativizing locality in this way the context-sensitive agreement patterns introduced in the previous chapter can be captured in a precise manner, alongside standard locality patterns.

The first step in the decomposition of φ involves de-abbreviating the φ-bundle. I undertake this in the next section.

2.1 φ-features, 'complex symbols' and Target-Controller Pairs

In current terms, a complex symbol Q as defined in *Aspects* corresponds to the formal features on terminal nodes. (Halle 1992, Noyer 1992). In minimalism, a feature-driven framework, the inventory of features includes not just interpretable features (inherent or contextually assigned) but also the uninterpretable operations that trigger operations like probe. I assume for present purposes that the latter, too, are constituents of Q. Thus, the constituents of Q drive the operations of syntactic structure building, as well as the establishment of syntactic relations, and they map onto vocabulary insertion sites in the morphological component. We might represent Q of T^0 as follows:

3) (Q (PERSON
NUMBER
TENSE
__V
…))

This is is certainly consistent with evidence from vocabulary insertion, in that it is common cross-linguistically for these features to be spelled out by a single morphological formative (synthetic/fusional languages, especially). However co-constituency in Q does not necessarily mean that individual elements of Q do not also operate independently of one another. For example, φ-features operate independently of other members of Q with respect to the triggering of AGR-relations. Thus Q cannot be an unstructured bundle, but must enable differentiation between constituents for the purposes of operations that target certain features and not others:

4) $$\left(Q \left(\begin{array}{ll} \phi & \left[\begin{array}{l} \text{PERSON} \\ \text{NUMBER} \end{array} \right] \\ x & \left[\text{TENSE} \right] \\ y & \left[\text{\_\_V} \right] \\ z & \left[\ldots \right] \end{array} \right) \right)$$

Given that it must be possible to isolate individual members of Q for particular purposes, there is no a priori reason why features of the φ-bundle must be any less autonomous than other elements of Q. What I ultimately propose is that even (4) is too unstructured; the features of Q are maximally organized into subconstituents, the organizing principle being entailment relations between subcategories. We will get to this shortly, but first I revisit the nature of the asymmetries that will motivate this move.

Recall that the framework of Chomsky (2000, 2001) assumes the canonical clause structure in (5), which predicts the set of target-controller pairs (TCPs) for a canonical transitive clause given in (6)

5) T YP_{SU} v XP_{DO}

φ- φ φ- φ

6) <φ, XP_{DO}>, <φ, YP_{SU} >.

Given the assumption that the AGR-relation is established by φ-probe, the most straightforward prediction (abstracting away from possible splitting of Qs into multiple positions of exponence) is that one agreement marker will realize all φ of the DO and another will realize all φ of the SU. Now recall the description given of Georgian agreement in chapter 1:

7) Person agreement is with the DO unless the DO is 3rd person, in which case person agreement is with the SU.

8) Number agreement is with the SU unless the SU is singular, in which case a plural (non-3rd person) DO controls number agreement.

These patterns are incompatible with the set of TCPs in (6). If AGR-relations are triggered by the φ bundle in Q, as is standardly assumed, then we cannot explain how it is that person (π) and (#) enter into different locality patterns in Georgian. The contrast with the predicted outcome of agreement is made explicit when we compare the rigid set of TCPs predicted by (6) with the variable set of TCPs that arise in Georgian. Abstracting away from the structural position of the AGR-head, we find that three out of four possible TCP sets are attested in Georgian:[5]

9)	<π, SU>, <#, SU>	v-xedav-t 1-see-PL 'We see him.'	*Georgian*
10)	<π, DO>, <#, DO>	g-xedav-t 2-see-PL 'I see you(Pl).'	
11)	<π, DO>, <#, SU>	g-xedav-t 2-see-PL 'We see you(sg).'	

(9) arises when the DO is 3rd person and the SU is plural, because a 3rd person DO cannot control person agreement (therefore allowing the SU to do so) and a plural SU must control number agreement. (10) arises when the DO is a participant and the SU is singular, because a participant DO must control person agreement and a singular SU cannot control number agreement (therefore allowing a non-3rd person DO to do so). (11) arises when the DO is a participant and SU is plural, because a participant DO must control person agreement, and a plural SU must control number agreement.[6]

---

[5] The missing set of TCPs is <π, SU>, <#, DO>. See Chapter 3 for further discussion of this gap.

[6] Note the surface ambiguity of - *t* in (10) and (11). This is precisely what should be expected on the assumption that in both examples –*t* is the morphological realization of the same AGR-head and the Agree relation only cross-references φ-features and not Case or grammatical function.

It is important that these conditions on TCPs are stated on individual φ-attributes (person, number). If φ (the bundle) is the diacritic that licenses the establishment of AGR-relations, then these conditions, though accurate, are not theory-internally coherent, because match(φ) will halt the probe when it encounters the closest XP with φ in its domain. By hypothesis, conditions on match(*p*) (where *p* is any subcategory) can only be sensitive to the presence or absence of *p* in the AGR domain. A simple solution is to allow decomposition of the φ-bundle and so that individual φ-attributes can trigger Agree independently of one another.[7] This permits us to formulate the conditions on TCP sets in (9)-(11) as special cases of the anti-intervention pattern from Chapter 1, where for a probe (*p*):

12) X controls agreement unless it lacks some property *p,* in which case Y may control agreement.

The situation is analogous to the alternation in TCPs established when comparing passive and active derivations (Chapter 1). The most notable difference is that the passive/active alternation occurs across 'conjugations' whereas (12) is describing CSA effects which are in effect alternations *within* a conjugation. In the active/passive alternation *p*(X) is trivially absent because X (the SU) is suppressed. Within a conjugation the absence of *p*(X) is not trivial; *p* cannot be an attribute that is inherent to all arguments. It must be an attribute that is variably specified, and crucially, it must be demonstrably absent (on independent grounds) in precisely those configurations where the effect of (12) is manifested. Just as suppression of X is not arbitrary in the active/passive alternation, absence of *p*(X) is not arbitrary in CSA alternations.

---

[7] This has the effect of turning Agree into a cyclic operation which will apply as many times as there are active φ-features on a given AGR^0.

2.2 Absence of *p* is underspecification

By definition, no two members of a conjugation set share the same ϕ-specification, so it is conceptually plausible that *p* may be present on one member of the set, but not on another. However, plausible though it may be, it is hardly a necessary conclusion. A conjugation (and I'm using the term in its traditional sense) is simply an exhaustive cross-classification of variable attributes (usually person, number, and class) where some independent attribute(s) is held constant for all members (usually a combination of tense, mood, and aspect). In a purely taxonomic sense, every member of a conjugation set is equally specified for whatever attributes define the set; the members contrast only in their values. So what does it mean for *p* to be present for certain members but absent for others?

Recall for Georgian, π-agreement is controlled by the DO unless it is 3^{rd} person; #-agreement is controlled by the SU unless it is singular. We can capture both of these cases with the generalization that disruption of the 'usual' locality pattern (i.e.agreement displacement) is triggered when the preferred controller has the default/unmarked value of the feature that would match the probe(*p*). This asymmetry between the default and marked values appears to have real derivational consequences with respect to locality. The default values behave as though they were truly unspecified, i.e. not included amongst the formal features of Q. Thus, absence of *p* in (12) correlates systematically with X having a default specification for *p*. The implication is that the formal ϕ-feature system allows for underspecification.[8] On the underspecification of syntactic features see also Dupuis, Lemieux & Gosselin 1991, Fassi Fehri 1992, Laka 1993, Rooryck 1994, Wiklund 2001, Burzio 1991, 1992, Laka 1993, Reuland and Reinhart 1995, Cowper 2001, Cowper 2003a,b).[9] (Apparent counterexamples, such as English, where 3^{rd} person singular arguments control overt agreement morphology will be addressed in Chapter 3.)

This triggering of CSA effects in the environment of default (underspecified) φ-values is not an idiosyncratic phenomenon. The pattern is robust across languages. A few examples are given below, and we will see more in chapters 3 and 4.[10]

13) Karok (Bright 1957)

	1.DO	2.DO	3.DO
1.SU	—	nu-_____	ni-_____
		2-*base*	1-*base*
2.SU	na-_____	—	ʔi-_____
	1-*base*		2-*base*
3.SU	na-_____	nu-_____	ʔu-_____
	1-*base*	2-*base*	3-*base*

Karok (Hokan) has a core π-agreement pattern that is virtually identical to Georgian's. π-agreement is controlled by the DO unless it is 3^{rd} person, in which case the

[8] This is contrary to the usual assumption in DM that underspecification is a property of vocabulary items, not of syntax (Noyer 1992, Halle&Marantz 1993, Harley&Noyer 1999.

[9] NB: This position will be refined in section later in this chapter. Total underspecification will be distinguished from partial underspecification and I will show that both have the potential to trigger agreement displacement, but under different circumstances. In Chapter 3 the analysis for Georgian will in fact assume partial underspecification and not total underspecification for 3^{rd} persons (likewise for Dakota in Chapter 4). Algonquian, however, will be argued in Chapter 3 to have total underspecification for 3rd persons.

[10] It should not be equated with zero-morphology for default values. To diagnose the CSA effect, there must also be a blocking pattern.

controller is the SU. Also similar to Georgian, Dakota (Siouan) #-agreement displaces when the preferred controller is singular, giving rise to the same ambiguities:

14) un-kaska-pi
    1-bind-PL
    'He binds us.'
    'They bind me.' (Buechel 1939)

2.3 Feature notations

I have expressed two desiderata for a theory of features: (A) that it capture entailment relations between features; (B) that it express underspecification. The latter is justified by the fact that alternations in TCPs seem to occur in precisely those environments where the 'regular' controller is plausibly underspecified. The former has not been justified yet. Harley and Ritter (2002) ague that to the extent that any theory of features recognizes the existence of natural classes, entailment relations must be at least tacitly presupposed. Even the (seemingly) banal classification of φ-features as being either person, number or gender attributes presupposes an entailment relation between each feature and its 'type'. They go so far as to note that even the assigning of an attribute/value distinction to features tacitly encodes an entailment relation (H&R 2002:4).[11] We will see that entailment is also crucial on empirical grounds. Once a fuller range of anti-intervention patterns is considered, it will be necessary to claim that the decomposition of φ does not stop at differentiating person and number attributes. More fine-grained distinctions can

[11] I would counter that this depends upon the nature of the attribute/value distinction: in (17) it seems right to say that the value (1, 2, 3, etc.) entails the attribute 'person'; however in (18) there is no analogous relation between the Boolean value and the attribute.

also be triggers for CSA effects, and in order to capture these in a uniform way we will have to introduce complex feature specifications for more 'marked' values. This can be illustrated with data we have already seen, from Georgian. In Georgian, we just saw that 1st and 2nd person objects pattern together as a natural class, in that they control π-agreement, as opposed to 3rd persons which do not. In otherwords, we want probe (*p*) to successfully match both subcategories. What is the correct *p*? If *p*=1 then it will not match the 2nd person pronoun, and if *p*=2 it will not match the 1st person pronoun. We can resolve this if we replace the 1st and 2nd person categories with complex symbols that contain a common subcategory on the basis of which they will form a natural class, as well as distinctive subcategories, on the basis of which they will be contrastive with one another. There are numerous ways this could be done, one of which is illustrated below, using the semantic categories [+PARTICIPANT], and [±SPEAKER]. The former will unite them, and the latter will contrast them.

15) [+PARTICIPANT] —> [±SPEAKER]

(15) asserts the existence of a possible category with the feature specifications [+PARTICIPANT, +SPEAKER] (i.e. 1st person) and another with the feature specifications [+PARTICIPANT, -SPEAKER] (i.e. 2nd person. But now consider again the question of how to formulate probe (*p*). Clearly, we want *p* to be [PARTICIPANT] since the whole point was to successfully match both 1st and 2nd person pronouns. However notice that in Georgian (9)-(10) the agreement marker is not simply tracking the [PARTICIPANT] feature; there is a morphological contrast between 1st and 2nd person agreement. This means that the probe is matching on the basis of [PARTICIPANT], but it is Agreeing/valuing with more than just the one feature; it doesn't only agree with [PARTICIPANT], it also agrees with [±SPEAKER].

In principle there does not seem to be anything wrong with this, but if we are going to allow agreement with features other than *p*, then we must be able to state which of the features in Q may participate in the relation. If entailment relations between features are explicitly encoded, this problem is straightforward; we can restrict Agree/value to features that enter into an entailment relation with *p* (see section 2.7).

Crucially, not all feature notations encode underspecification and entailment. Consider the following five possibilities for the encoding of person features (not exhaustive, but they illustrate certain key differences between possible systems). In a conventional feature system, there are two components to a feature's representation: its attribute/category label, and its value. I have encoded this here using bracket embedding to represent values. In (16) and (19) the feature PARTICIPANT denotes discourse participants (1^{st} and 2^{nd} persons) and SPEAKER 1^{st} persons. The feature specifications for 1^{st}, 2^{nd} and 3^{rd} persons in each of the five systems is given in (21).

16) $\begin{bmatrix} \text{PERSON} & \begin{bmatrix} \text{PARTICIPANT} & [\text{SPEAKER}] \end{bmatrix} \end{bmatrix}$

17) $\begin{bmatrix} \text{PERSON} & \begin{bmatrix} 1 \\ 2 \\ 3 \\ 0 \end{bmatrix} \end{bmatrix}$

18) (2 [+ / - / 0] ; 1 [+ / - / 0])

2 | +
 -
 0
1 | +
 -
 0

19) (PARTICIPANT / SPEAKER)

20) (1 / 2)

21)	1st person	2nd person	3rd person
system (16)	[PERSON [PART [SPKR]]	[PERSON [PART]]	[PERSON]
system (17)	[PERSON [1]]	[PERSON [2]]	[PERSON [3]] or [PERSON [0]]
system (18)	[+1] [-2]	[-1] [+2]	[-1] [-2]
system (19)	[PARTICIPANT] [SPEAKER]	[PARTICIPANT]	
system (20)	[1]	[2]	

These five systems are in certain respects notational variants of one another (e.g. they are equal to the task of making the relevant distinctions), however they are equivalent only up to a point. They differ in the substance of their values (Boolean (18) vs. semantic (16), (17), (19), (20); they differ with respect to the amount of structure required to express a value (simple (19), (20) vs. complex (16), (17), (18) ; and they differ with respect to whether or not entailment relations between features are expressed ((16), (17) yes vs. (18), (19), (20) no).

They also differ with respect to the implementation of underspecification. In (17) underspecification would be represented by a '0' value, or alternatively 3rd person would simply have a '3' value. In (18) underspecification would also be represented by a '0' value, but 3rd person in this system is usually represented as [-2, -1]. In (16), (19) and (20) underspecification would require the actual absence of the attribute/category label itself. Given the operation match, the latter group is consistent with the locality effect that is being correlated here with underspecification (i.e. extension of the search space); the former are not. If underspecification is encoded as a '0' value, or *f*(0), then an underspecified goal has *f* and should satisfy the operation match (*f*), blocking extension of the search space. Similarly, if default forms are complex as in [-2, -1], then match (*f*) should also succeed. It is only if underspecification is encoded by the absence of *f* that we predict that match(*f*) will fail for the underspecified goal, and the search space of the probe will extend farther. Thus, by virtue of their encoding of underspecification, the notations in (16) , (19) and (20) are superior to the other two.

Now consider the encoding of entailment relations between features. Neither (19) nor (20) encodes entailment relations between features. Note that here the internal coherence of the feature system is relevant. For example, the inventory of categories in (20) is incompatible with the postulation of entailment relations(why should 1st person entail 2nd person, or vice versa?), so this would be a good notation for a theory where entailments were explicitly rejected. On the other hand, for (19) the failure to encode entailments is an oversight, since there are clearly implicational relations between features (the feature [SPEAKER] never occurs in the absence of the feature [PARTICIPANT]). Thus, in terms of internal coherence, it would seem that (20) is superior to (19), but it does not suit our purposes.

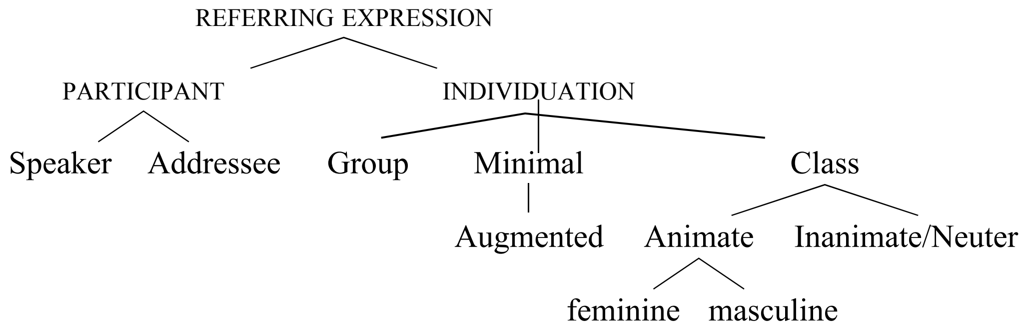
Of the five systems, (17) and (18) lack adequate encoding of underspecification; (19) is ruled out on the basis of internal incoherence. This leaves (16) and (20), the latter with no (feature-theoretic) need for entailment (22), the former with the strongest representation of entailment of all five systems (22).

22)	a.	I	[1]	b.	I	[PERSON[PART [SPEAKER]]]
		you	[2]		you	[PERSON[PART]]
		she			she	[PERSON]

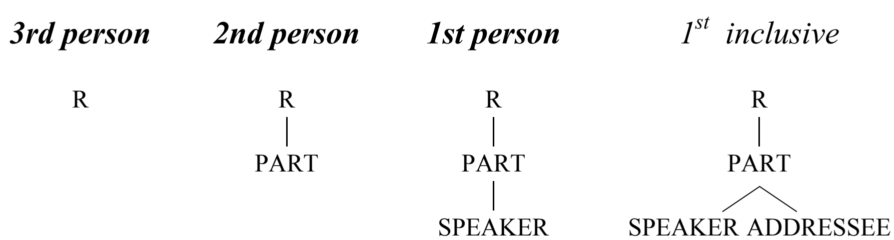
So, we are looking for a system with the properties of (16). One recent formalization of morphological φ-features with the necessary characteristics is the feature-geometry proposed in Harley and Ritter 2002. I summarize their system in the next section

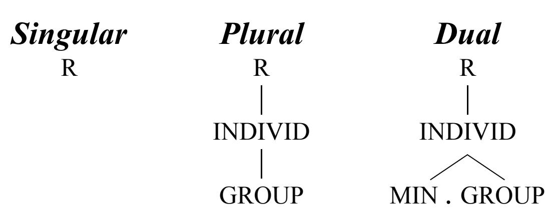
2.4 Harley and Ritter 2000, 2002

H&R propose a universal set of person and number features with which it is possible to characterize the pronominal paradigms of the world's languages. They show that the distribution of these features is highly constrained, and that their interactions are demonstrably systematic. The relations between features are modeled using a geometric representation of morphosyntactic features, like those that have been developed in phonology (Clements 1985, Sagey 1986, Avery & Rice 1989, Dresher, Piggott & Rice 1994, Dresher 2002 and many others) The features in the geometry represent the grammaticalization of fundamental cognitive categories, such as reference, plurality, and class. The organizational structure provides a principled explanation for the observed restrictions on these paradigms such as implicational dependencies and co-occurrence restrictions. They motivate the geometry through the analysis of pronoun paradigms in a broad range of genetically distinct languages, including Daga (Papua New Guinea); Kalihna (Carib); Tonkawa (Coahuiltecan); Hopi (Uto-Atezcan); Chinook (Pinutian); Yimas (Sepik); and Bouma Fijiian.

23) 

Note that H&R distinguish between features and organizing nodes (in SMALL CAPS). I will ignore this distinction. Instead, I will treat all points in the geometry as independent features, with the structural dependency represented by '|' taken as an entailment relation between the dependant feature and the feature(s) that dominate it. The root node REFERRING EXPRESSION (henceforth R), can be taken to be either a pronominal/nominal element or an agreement marker. The PARTICIPANT node encodes contrasts between person features, as in (24). The INDIVIDUATION node contains number (25) and class features (discussion of which I defer to chapter 3).

24) Person: 

25) Number: 

Crucially, H&R eschew the use of traditional categories as abbreviation. By abandoning such labels as '2nd person', 'plural' or 'masculine' they make explicit the entailment relations between feature attributes and feature values, satisfying the second desideratum discussed above. In the geometry, the entailment relation is encoded by '|'.

The geometry also satisfies the first desideratum: the possibility for total underspecification. (23) is a privative feature system; if a category is negatively specified for a particular feature, this is encoded by structural absence of the feature.

Note that underspecification is not a special property of the least-specified value of an attribute. Underspecification is systematically employed to encode contrast. There is total underspecification (as in 3rd person), but also relative underspecification, as in the distinction between 1st and 2nd person, where 2nd person is simply the default interpretation of a [PARTICIPANT] feature that has no further dependent. We have already seen that total underspecification has a clear locality effect in the computation. Later in this chapter I will show that relative underspecification also has consequences for locality. Later I will also argue that the availability of total underspecification is a point of cross-linguistic variation.

Crucially, underspecification of both kinds is sensitive to the particular inventory of a language. The richer the set of contrasts in an inventory, the more nodes/features will be 'activated'. So, for example, the activation of both [SPEAKER] and [ADDRESSEE] occurs only in languages with an exclusive/inclusive contrast.[12] Otherwise, only one is activated. For H&R the [SPEAKER] feature (denoting 1st person) is the usual dependent of [PARTICIPANT], making 2nd person the default interpretation of a bare [PARTICIPANT]. The implication is that 2nd person is universally less specified than 1st person. However, I will

---

[12] With both [SPEAKER] and [ADDRESSEE] activated one might expect that neither 1st or 2nd person is a bare [PARTICIPANT]. H&R assume this; however I see nothing in the theory to exclude the other possibilities and I will argue in Chapter 3 that for Nishnaabemwin 1st person is the default interpretation of a bare [PARTICIPANT] feature, despite the fact that [SPEAKER] is available in the inventory.

argue later that languages may vary with respect to whether 1^{st} or 2^{nd} person is the default interpretation of [PARTICIPANT]. In languages where 1^{st} person is the default, 2^{nd} person will be represented by [PARTICIPANT[ADDRESSEE]].

Note that the geometry is also meant to capture certain co-occurrence patterns between features. The success of the formalism in this regard is less straightforward. In principle it is implemented by sisterhood, but in the H&R proposal, sisterhood has 3 possible consequences:

26) 1. cross-classification: e.g. [PART] and [INDIV]

2. complementary distribution: e.g. [ANIM] vs. [INANIM]

3. combination: e.g. [GROUP]+[MIN] —> 'dual' interpretation

In certain instances, more than one of these is possible for the same pair, e.g. [SPEAKER] and [ADDRESSEE] can either be in complementary distribution, or combine if both are active in an inventory. Furthermore, combination can involve either conjunction or intersection, as pointed out by H&R themselves (citing Leslie Saxon, pc). I will not take a position on the formalization of sisterhood in the geometry and the matter of whether or not it successfully formalizes feature cooccurrence patterns. What is important for present purposes is the formalization of entailment.

2.5 A theory of formal φ- features

The representations proposed in H&R differ from standard assumptions about formal φ-features in three respects.

27) I. structured vs. 'flat' values

II. underspecification vs. full specification

III. privative features vs. n-ary feature values

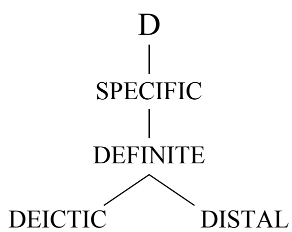
Not all of these are necessary for encoding entailment and underspecification, but they are sufficient, and so I will adopt their general approach, with some modifications. The most significant of these is that I introduce a new feature: [π] dominating (entailed by) the [PARTICIPANT] feature:

28)

```
        R
        |
        π
        |
      PART
        |
ADDRESSEE/SPEAKER
```

[π] is necessary on empirical grounds. Without it, non-participants are obligatorily underspecified aside from features of INDIVIDUATION, and perhaps the root feature R. This is too strong a claim. While there is strong evidence that certain 3rd persons *are* totally underspecified, we will see equally clear evidence that some 3rd persons *cannot* be totally underspecified for the simple reason that they can be interveners, thus, the feature specification of 3rd persons must be a point of variation. I introduced π earlier in this chapter as an abbreviation for *person,* without comment. However the actual intended substance of π is more abstract (and largely unclear at present). The intuition is that the so-called category of 'person' is just conventional, and the family of features in (28) actually encodes a progressively narrower range of restrictions on referentiality. However the nature of the cline is not at all obvious. On the one hand, as noted by H&R (cf. Jakobson 1971 and others) one clear contrast between participants and nonparticipants is that the reference of the former is inherently dependent upon shifting discourse roles, whereas the reference of the latter can be fixed. One might expect that the cline is thus somehow pertaining to deixis and the property of fixed reference (or the lack thereof). In this case the introduction of a feature category in the position of [π] is well-motivated, as there must be some means of encoding the fact that nonparticipant pronouns are also inherently deictic (though not obligatorily shifters), whereas full NPs need not be deictic, though they can be. In other words, 3rd persons subcategorize for deixis in more ways than can be represented by the binary contrast between participants and nonparticipants. (29) shows a feature geometry of nominal inflection proposed by Cowper and Hall (2002) for English:

29)



Cowper and Hall propose (29) to account for certain restrictions in the mapping between the vocabulary items that encode nominal inflection and the syntactic environments in which they occur. While they explicitly abstract away from pronominal features, a reasonable hypothesis is that H&R's [PARTICIPANT] feature would be a dependent of [DEICTIC]. It might then be the case that [π] is one of the features on the branch between D and DEICTIC, or it might be that [π] is an abbreviation for one or more points along this branch, or of course it might be that [π] is actually a distinct category that needs to be introduced either below or above DEICTIC. I take π to be a variable ranging over the highest (meaning most inclusive) element in the entailment hierarchy from [SPEAKER] to [REFERENT]:

30) SPEAKER > PARTICIPANT > DEICTIC > DEFINITE > SPECIFIC > D > R(EFERENT)

I assume that π is sensitive to language variation in the selection of formal features (FF). So for some languages π may be all-inclusive (REFERENT) whereas for others it may signal a narrower sub-class, e.g.DEFINITE (Hungarian, Vogul (Uralic)). (30) is not intended to be exhaustive. Nor do entailment hierarchies seem to be constrained by traditional taxonomies. For example, ANIMACY is typically considered to be a gender/class feature that is arbitrarily assigned to NPs. However ANIMACY is also, in many languages, entailed by participants, so we predict that π should be entailed by animacy in some languages (this seems to be the true in Algonquian, for example).

2.6 Features, entailment and conditions on match

Given the enriched feature theory presented above, the syntactic operations that range over such features must be re-examined. I begin with match. We've been assuming, following Chomsky 2000, that match is the operation that evaluates whether an element within the search space of a target has the features that are being selected for by the probe. What counts as a match is relativized to the features of the target, in effect reducing relativized minimality (Rizzi 1990) to conditions on match. (28) shows the three basic locality patterns that we expect to find under the standard interpretation of match. In (a) the matching goal is not the first object encountered by the probe, but it is the first matching object. In (b) the matching goal is the first object encountered by the probe, and the probe is halted, regardless of whether another potential match might be found farther down. In the third (c) no match is found (represented by ø as endpoint of the probe).

31) Core locality patterns

a. ________________________

X	YP	ZP
[F]	[□F]	[F]

b. __________

X	YP	ZP
[F]	[F]	[(□)F]

c. ______________________________ø

X	YP	ZP
[F]	[□F]	[□F]

Recall that for Chomsky 2000 matching is feature identity, where identity is tacitly defined over feature attribute/category labels and not values (since probe is by definition unvalued, the value of the goal cannot enter into the evaluation of match). Given the theory of features just introduced, this definition is ill-defined, as the contrast between category/attribute and value is not strictly defined. Features in the system are privative categories with no 'value' per se. In a sense, entailment dependencies introduce something like values. For instance, we could think of a [PART[SPEAKER]] as a PARTICIPANT category with the value SPEAKER. We will see later in this chapter that in certain respects entailment relations do function as though dependants were 'values' for their dominating features. However it is nevertheless unmotivated to claim that there is a substantive distinction between categories and values; we could just as easily say of PARTICIPANT that it is a value of the category π. Thus, PARTICIPANT looks to be a value from one perspective, and a category from another. This raises a question about the criterion for match. If a probe [F] encounters in its search space an NP with features [F[G]], will this be a match? In other words, if match is feature identity, does [F]=[F[G]]?

The answer has consequences for locality, as can be seen by comparing the locality patterns in (32) and (33):

32) Locality pattern if [F]≠[F[G]]

a. ________________________

X	YP	ZP
[F]	[F[G]]	[F]

b. __________

X	YP	ZP
[F]	[F]	[F[G]]

c. ________________________________ø

X	YP	ZP
[F]	[F[G]]	[F[G]]

d. __________

X	YP	ZP
[F]	[F]	[F]

33) Locality pattern if [F]=[F[G]]

a. ___________

X	YP	ZP
[F]	[F[G]]	[F]

b. __________

X	YP	ZP
[F]	[F]	[F[G]]

c. ___________

X	YP	ZP
[F]	[F[G]]	[F[G]]

d. __________

X	YP	ZP
[F]	[F]	[F]

In principle, either answer has some intuitive merit. There is a natural sense in which [F]≠[F[G]]. To say otherwise would, for example, be equivalent to saying 1^{st} person = 2^{nd} person, or singular = plural:

34)

[F]		[F[G]]	
[PART]	1st person	[PART[ADDRESSEE]	2nd person
[INDIVID]	singular	[INDIVID[GROUP]	plural

On the other hand, there is an equally natural sense in which we might expect [F] to match [F[G]]. To the extent that the criterion for match can be stated simply in terms of presence or absence of [F] it would some that that [F] *should* find a match in [F[G]] because [F] is a subset of [F[G]]. The prediction is then that a probe for [PART] will match an NP that has [PART[ADDRESSEE]] and a probe for [INDIVID] will match an NP that has [INDIVID[GROUP]], just as a probe for [π] will match an NP with [π[PART]]. This seems right. Participants are the canonical controllers of person agreement, just as plural arguments are the canonical controllers of number agreement.

Thus, intersection of the features of the probe and the goal seems to be sufficient for match; identity is not required. Given the theory of entailment, intersection can be stated by the following principle:

35) Match is defined by entailment.

36) Probe (F) and Goal (F') match if Goal (F') entails Probe (F).

Differently put, if probe (F) is a subset of goal (F') — including identity — then F and F' match. It follows from the entailment principle that if F is a proper subset of F' then the root of F is entailed by the root of F'. Intersection necessarily holds from the 'top' of the feature tree. This reduces to the following:

37) Match is evaluated at the root.

Where:

38) Probe (F) and Goal (F') match if Goal (F') entails root Probe (F).

The postulation of entailment relations makes the precision of these definitions essential. Note that a consequence of encoding entailment is that if a category is specifed for a feature [F] then it is automatically also specified for every feature entailed by [F]; that is to say every feature (in a branching diagram of feature relations) on the path between [F] and the root node. Crucially, however, I will assume that this only holds of interpretable features. This forces us to posit an important stipulation, not yet stated:

39) Q is not automatically specified for all features entailed by a uninterpretable feature [F-] in Q.

This prevents overapplication of match, which would otherwise occur. This is illustrated below, where features automatically specified by entailment are in grey text:

40) Two entailment scenarios:

Scenario A (assume to be correct):
match correctly predicted to fail

Probe	Goal
H	NP
[π-]	[R(eferent)]
	[Definite]

Scenario B (assume to be incorrect)
match incorrectly predicted to succeed

Probe	Goal
H	NP
[R(eferent)]	[R(eferent)]
[Definite]	
[π-]	

If uninterpretable features could trigger specification by entailment, as in Scenario B, then we would incorrectly predict that there should be no anti-intervention effect in, say, the environment of 3^{rd} person, because although, lacking a π feature, the goal would still have at least an R specification, which would successfully match the root of the probe.

The technicalities of how match is implemented could proceed in various ways, but as far as I can tell the details of this implementation do not have consequences for the derivation. One possibility would be that match evaluates every feature of the goal and holds of any that satisfies (38). This would, for example, have [π] and [PART] each being evaluated independently. Another possibility would be that [π[PART]] are evaluated as a single constituent, in which case entailment might be evaluated against the whole. In either case the outcome is the same.

A sample of outcomes predicted by the entailment principle is shown below:

41)

[F] Probe	[F'] Goal	Match
[π]	[π]	yes
[π]	[π[PART]	yes
[π]	[INDIVID]	no

Understanding match in this way allows the introduction of a novel question that tests the limits of the system. What if probe (F) is more highly specified than goal (F')? For example, what if probe is [π [PART]] and goal is [π]? This is not an implausible scenario. If the category/value distinction is just an artefact of nonprivative systems, then the uninterpretable feature(s) of the probe could in principle be any feature structure that is available in the inventory. Starke (2001), who also assumes that syntactic features are hierarchically organized, assumes an anti-identity principle such that if features of the goal are not of the same feature class as those of the target, then NP cannot be an intervener (in our terms this means the NP cannot satisfy match). If the probe is more highly specified than an NP in its search domain, then it effectively belongs to a sub-class to which the NP does not, and so match is impossible. The prediction is thus:

42) Predictions of Starke's (2001) anti-identity principle

	[F] Probe	[F']	Match
a.	[π[PART]]	[π]	no
b.	[π[PART]]	[π[PART]]	yes
c.	[π[PART]]	[π[PART[ADDRESSEE]]	yes

However, Starke's anti-identity principle is incompatible with the entailment principle in (38). If match is evaluated at the root (the condition being that [F'] of the potential goal must entail the root of [F] of the probe), then (42) is predicted to satisfy match, contrary to Starke's proposal. The difference boils down to whether match is evaluated against the root of the probe or against the entire probe. There are empirical grounds for choosing the former (38). We will see that there are languages for which it can be shown that probe must be highly specified, as in (42) but where a less specified NP is nevertheless an intervener, contra (42). The argument relies on new proposals regarding the operation value/Agree, to which I turn next.

2.7 Match versus Value

The operation match identifies a potential controller, however by hypothesis Agree *qua* agreement — meaning the transmission of features of the controller to the target —is implemented by the operation value. In this section I examine the conditions on this operation in light of the developing theory of features. In the framework of Chomsky

2000 it is clear that conditions on value must be more restrictive than conditions on match, because there exist configurations where match succeeds but value fails. Match is necessary, but not sufficient, for Agree.[13] Thus, whatever the correct formulation of conditions on value may be, the set of configurations that satisfy these must be a subset of those that satisfy match. In effect, match implements a visibility condition identifying a goal as being eligible for Agree/value, but not necessarily establishing the goal as a controller. Henceforth I use the term *controller* in a technical sense to mean a goal that matches and values some probe. In the terminology of Chomsky 2000, a goal that matches but fails to value is an *intervener* (this is not quite accurate; see discussion in section 2.9). An element that is accessible to the probe but fails to match will be referred to henceforth as a *bystander*.[14]

The purpose of this section is to explore the operation value in light of the empirical challenges posed by context-sensitive agreement effects and the hierarchical theory of features adopted above. The central question is: how do we formulate conditions on value so that the operation is systematically licensed or prevented in the appropriate contexts? However, before entering into this topic directly, I consider the following question: what does it mean to value given a feature theory in which the category/value distinction is not substantive?

2.8 What can it mean to value if there are no feature values?

One might argue that the value operation ought to be eliminated; that without 'values' there can be no valuing. All else remaining equal, this would involve a return to feature

[13] The opposite scenario is impossible, since match is a prerequisite for value.

[14] I thank Daniel Currie Hall for suggesting the term *bystander*.

checking (Chomsky 1995). Whereas in a framework with both match and value, the configurations that satisfy conditions on value are a subset of those that satisfy match, in a feature-checking framework, the narrower distribution of successful valuing configurations must be derived from match directly. In other words, conditions on the establishment of TCPs must be stated entirely as conditions on match. The only way to accomplish this (without reintroducing some notational variant of value) is to define match strictly such that the target and the controller must have identical feature specification (relative to the selectional feature/probe of the Target). Of the three checking configurations presented below — all of which would satisfy the entailment condition in (38) — only (43) would succeed under strict match.

43) Strict Match

	[F] Probe	[F']	Match
a.	[π[PART]]	[π]	no
b.	[π[PART]]	[π[PART]]	yes
c.	[π[PART]]	[π[PART[ADDRESSEE]]	no

In a framework with strict match it must be the case that AGR-heads are free to enter the numeration with whatever combination of φ-features necessary to satisfy strict match. Thus the [π[PART] probe is only successful in a derivation where the goal is 2nd person; in a derivation where the goal was either 1st or 3rd person, the feature specification of the probe would have to change accordingly. Presumably, conditions on convergence would eliminate all configurations that fail to satisfy strict match from the output of the computation, leaving just properly matching outputs.

There can be no syntactic account for CSA effects in this view. Consider, again the special status of 2nd person controllers in Nishnaabemwin. Whichever argument is 2nd person will control agreement. In (44) the 2nd person is SU and in (a) but a DO in (b); the form of agreement is the same in either case, with 2nd person controlling the prefix.

44) a. g-waabm-i 2>1

2-see-DIR.THEME

'You see me.'

a. g-waabm-in 1>2

2-see-INV.THEME

'I see you.'

However, if there is no 2nd person argument, then it *is* possible for a 1st person argument to control π-agreement.

45) n-waabm-aa

1-see- DIR.THEME

'I see him.'

Clearly, 1st person is able to satisfy conditions on match and value, or 1st person agreement would never be possible.

If match is strict, and the numeration freely selects AGR features, nothing in the derivation predicts the ungrammaticality of the following Algonquian form:

46) a. *n-waabm-i 2>1

1-see-DIR.THEME

'You see me.'

It should in principle be possible to create a numeration where match will be satisfied; all that would be required is merger of a probe with 1st person φ-features: [PARTICIPANT[SPEAKER]]). An explanation of the Algonquian agreement pattern would require a retreat to the Suppression Hypothesis (see chapter 1). Thus, the thesis that CSA effects are syntactic requires that a distinction between match and value be maintained.

We are back to the starting point, then: what does it mean to value given a feature theory in which the category-value distinction does not exist? The answer I propose is that value is structure building, where the features (identified by match) of the goal are 'copied' to the probe. Following Chomsky 2000, I also assume that valuing in some sense 'deletes' the probe (though perhaps not literally, it at least renders it inert).

47) Probe (F) Goal (F')

~~π~~ π

| |

~~PARTICIPANT~~ PARTICIPANT

|

SPEAKER

Where the probe and the goal have equally specified features, structure-building is vacuous (but we will assume, successful) and deletion occurs as usual. In the next section I will argue that where the goal is *less* specified than the probe, match may succeed but value fails.

2.9 Conditions on value

Recall from 2.7, conditions on value are more restrictive than conditions on match and it is possible for match to succeed but value to fail. In the framework of Chomsky 2000 there are two environments in which this occurs:

48) (A) where match identifies an inactive goal (i.e. a goal that has already been assigned Case);

(B) where match identifies a goal with 'defective' □-features.

(A) accounts for data such as (49) where a nominative NP can control subject agreement but a dative NP (b) cannot because it has inherent/□-related Case and so cannot be assigned Case by the target (T^0).

49) a. Strákarnir leiddust/*leiddist *Icelandic*

the.boys.NOM.PL walked.hand.in.hand3.PL/*3.SG)

b. Strákunum leiddist/*leiddust

the.boys.DAT.PL bored.3SG/*3.PL

'The boys were bored.' Sigurdsson 1996:1

A cross-linguistically robust characteristic of such constructions is that long-distance agreement with a lower (previously inaccessible) controller is possible.

50) Henni leiddust strákarnir *Icelandic*

her.DAT bored.3.PL boys.NOM.PL

'She found the boys boring.' Sigurdsson 1996:1

As discussed in Chapter 1, this is because the intervener has displaced and the probe, not having been deleted, is able to initiate a second search for a controller.

The type (B) case arises with 'true' expletives like English *there*. By hypothesis, these are φ-defective, meaning they have only a partial set of φ-features (Chomsky 2000:40). *There* is thought to have only a person feature. Incomplete φ ostensibly disables the expletive's ability to value. [15] The expletive is able to match (and is thus able to displace to subject position [Spec, TP]) but value of the uninterpretable φ-probe on T^0 is not possible. Here too, displacement of the intervener enables a long-distance agreement relation, also characteristic of such expletive constructions.

51) *There* were thought *t* to be three spies at the meeting.

Type (B) scenarios are not actually referred to as intervention effects in Chomsky (2000), though it is not clear to me why not.[16] Under the definition here of an intervener as a goal that matches but fails to value, *there* is as much an intervener as the dative NPs in (49) and (50).[17]

Conditions on value are not made explicit in MI, which makes intervention effects of type A rather mysterious. One possibility is that prior Case assignment to the

[15] The framework treats probes and goals symmetrically in this respect. A probe with a defective φ-bundle can seek a matching goal, but it cannot assign that goal Case (where Case assignment is valuation of the goal's uninterpretable Case feature). The consequence for the goal is that it remains active and is free to enter into another target-goal relation with a higher AGR-head. This system derives NP-licensing patterns and freezing effects: defective T is nonfinite T, defective *v* is passive/unaccusative *v*.

[16] There are competing analyses of such structures, not all of which assume that the expletive raises, but Chomsky's analysis is not one of these.

[17] Note that in both scenarios, the claim that the intervener cannot value is not quite straightforward. Along with the possibility for long-distance agreement, a second characteristic of intervention scenarios is that only partial agreement is possible with the long-distance goal, which can control number, but not person. As far as I know there have been no cases reported where the reverse is true. Thus, it would seem that interveners *do* agree partially, to the extent that they somehow absorb the ability for person agreement on the second cycle of the φ-probe. In the framework of Chomsky 2000 this π-absorption is mysterious. However given the theory of Agree proposed here there is a straightforward explanation, which we will see in chapter 4.

intervener somehow cloaks its φ-features rendering them inert. However one might then expect that such NPs might also fail to match. Whatever the solution, it must capture the fact that Case assignment renders the capacity of a goal to value inert without affecting its ability to match. I return to this problem in chapter 5. Ideally, there will be a unified proposal that accounts for both type A and type B scenarios.

To recap, there are two standard configurations where intervention effects arise: inactive goals and defective goals. I now introduce a third intervention scenario, evidence for which comes from context-sensitive agreement patterns. In these cases the intervener is neither inactive nor 'defective' in Chomsky's sense. Nevertheless the goal satisfies conditions on match but fails to satisfy conditions on value (to be made explicit below). A second cycle of Agree is then enabled.[18] I illustrate this here with data from Georgian dative nominative constructions (DNC). [19]

In Georgian DNCs the AGR-head that is the locus of π-agreement has both the dative and the DO in its search domain (52). Datives in Georgian are fully able to control agreement, as shown in (53)

52) [π-Probe … [DAT … [DO …]]]

53) m-i-naxav-s

1-APPL-EVID.see-TNS

I(DAT) have evidently seen him.'

However in (54) the dative *fails* to control agreement. Instead, the person prefix cross-references features of the DO, and a different 1st person marker is realized.

---

[18] Note that the evidence for this third kind of intervention scenario is similar in spirit to the first two: the first cycle triggers displacement and somehow affects the quality of the second (cf. footnote 9).

[19] An indepth analysis of Georgian agreement is given in chapter 3.

54) v-u-naxav-var

1-APPL-EVID.see-TNS

'He(DAT) has evidently seen me.'

We know from (53) that the failure to value in (54) cannot be attributed to the dative Case of the intervener. By hypothesis, the failure to value correlates with the difference between having 1^{st} and 3^{rd} person φ-specification. The simplest hypothesis would be that Georgian 3^{rd} person is totally underspecified and therefore cannot match, so the DO in (54) is the first accessible goal in the derivation. However there are two reasons for thinking that in fact the dative is an intervener and that agreement with the DO results from a second cycle of π-probe. First, the dative displaces (ultimately to subject position). Second, the AGR-relation with the actual controller is somehow affected by there having been a previous cycle. There is what I will refer to as a *second-cycle effect* in the morphology. This is manifested in the alternation between the *m-*/*v-* morphology that realizes 1^{st} person agreement. A first-cycle 1^{st} person controller is always spelled out by *m-* as in (53) and also (55) where there is no intervening dative and the DO values π-AGR on the first cycle.

55) m-xedav-s

1-see-TNS

'He sees me.'

Second-cycle controllers (which only ever occur when the first matching NP is 3^{rd} person) are spelled out by *v-* as in (54) and also (56).

56) v-xedav

1-see

'I see him.'

We will see in chapter 3 that the *m/v* alternation is generalized across the Georgian agreement system, and can be seen in nominative/accusative and ergative/absolutive constructions as well, under parallel circumstances.

Georgian 3rd persons evidently cannot be totally underspecified, since they are able to be interveners. This suggests that they are minimally specified. Suppose they have just a [π] feature; this will enable them to satisfy match for any probe with [π] at its root. Yet, they are clearly unable to value. 1st and 2nd person goals always control agreement, so the inability of a 3rd person goal to value must relate to the absence of a [PARTICIPANT] feature.[20] Thus, failure to value is a consequence of relative underspecification (whereas failure to match is a consequence of total underspecification).

57)

3rd person	2nd person	1st person
π	π	π
	|	|
	PARTICIPANT	PARTICIPANT
	|	|
	(ADDRESSEE)	(ADDRESSEE)

I propose that the correct generalization for conditions on value is that the matching feature(s) of the goal (i.e. those that entail the root of the probe) must be a superset of the probe:

58) Condition on value: G(oal) values P(robe) iff f'(G) entails f(P)

---

[20] I abstract away from the question of whether 1st or 2nd person is more highly specified in Georgian. As far as I can tell nothing in my analysis of the language hinges on this choice, the relevant split being between participants (1/2) and nonparticipants (3).

The outcome of value predicted by (58) for a sample of configurations is shown below:

59)

	[F] Probe	[F'] Goal	Value
a.	[π]	[π]	yes
b.	[π]	[π[PART]]	yes
c.	[π]	[π[PART[ADDRESSEE]]	yes

60)

	[F] Probe	[F'] Goal	Value
a.	[π[PART]]	[π]	no
b.	[π[PART]]	[π[PART]]	yes
c.	[π[PART]]	[π[PART[ADDRESSEE]]	yes

61)

	[F] Probe	[F'] Goal	Value
a.	[π[PART[ADD.]]]	[π]	no
b.	[π[PART[ADD.]]]	[π[PART]]	no
c.	[π[PART[ADD.]]]	[π[PART[ADDRESSEE]]	yes

In Georgian any argument with π is an intervener, but only participants can be controllers. This is precisely the pattern seen in (60), so I assume that the π-AGR in Georgian has uninterpretable [π[PART]]. By match, any argument with π will be an intervener, but only participants meet the entailment condition on value.

Recall that whatever the articulation of conditions on value, the set of configurations that satisfy these conditions must be a subset of those that satisfy conditions on match because it is possible for value to fail even though match succeeds.

Whereas entailment for match is evaluated just with respect to the root of the probe, for value it is with respect to the entire probe; this captures the added restrictiveness of value, and correctly predicts the outcomes seen above for Georgian.

Summary: value fails just in case the intervener has less feature structure than the probe. Failure to value triggers a second-cycle of Agree. The quality of the AGR-relation on the second cycle is somehow affected by there having been a previous cycle (for reasons to be explored further in section 2.12). In Georgian the evidence for this was an alternation in the morphology. In section 2.13 we will see an analogous second-cycle effect in Algonquian where the existence of a second cycle is signaled not by an alternation in the morphology but by an alternation in the environments where value succeeds or fails. But first I will explore derivational consequences of failure to match and value.

2.10 The consequences of failure: a locality paradox and a morphological change

Having defined the conditions on the two core operations that enter into Agree, we are now in a position to ask in concrete terms what the consequences are for the derivation when either of these operations fails to apply. Match fails either because there simply is not a suitable element Merged in the search space of the probe, or because a potential goal is totally underspecified with respect to the probe.[21] Value fails either because the goal is inactive or because it fails to satisfy the entailment condition on value (a result of relative underspecification with respect to the probe). There are two types of

[21] More precisely, total underspecification is what would be said of a *bystander* for which [F] of the probe was an interpretable category. If the *bystander* is an object for which no value of [F] is ever specified this would not be underspecification (but the effect is the same with respect to match).

consequences when either of these operations fail: consequences for locality, and morphological consequences. I begin by exploring the former.

Where match fails — and here I am speaking not of the probe failing to find a match, but of an element in the search space failing to *be* a match — the immediate consequence for locality is that an anti-superiority effect becomes possible. The highest accessible NP can be neither an intervener or a controller (it is a *bystander*), and so the next highest NP in the domain of the target becomes accessible to the probe (62). This was illustrated in chapter 1 by the number agreement patterns in Georgian and Dakota. The examples are repeated in (63) and (64).

62) a. _____________

#-AGR   SU$_{PL}$ DO

b. __________________________

#-AGR   SU   DO$_{PL}$

63) g-xedav-t   Georgian

2-see-PL

'I see you all.'

'We see you.'

64) un-kaska-pi   Dakota

1-bind-PL

'He binds us.'

'They bind me.'   (Buechel 1939)

An anti-superiority effect is also possible if value fails. Failure to value triggers a second-cycle of Agree. Here, the anti-superiority effect is contingent on displacement of the intervener, which enables previously inaccessible to NPs to match and perhaps value the probe.

There is nothing unexpected about such anti-superiority effects given conventional assumptions about locality and the cumulative assumptions made so far in this thesis. But recall, now, that this is not the only pattern we have seen. I repeat, yet again, the descriptions of π-agreement in Georgian, from chapter 1:

65) Georgian π-agreement:

a verb cross-references the person feature of its DO, unless the DO is 3rd person, in which case the person feature of the SU are cross-referenced.[22]

66) Georgian: DO as controller

a. m-xedav-s

1-see-TNS

'He sees me.'

b. g-xedav-s

2-see-TNS

'I sees you.'

cf. *m-xedav, *v-xedav with this meaning

67) Georgian: SU as controller

v-xedav

1-see

'I see him.'

[22] Technically this should say that a verb cross-references the person feature of its highest internal argument; we have seen that a higher dative blocks agreement with the DO unless it is 3rd person.

Here is a highly remarkable state of affairs. We have a paradox: a lower argument is behaving as an intervener with respect to a superior argument. By virtue of having each satisfied match, both the intervener and the controller must be in the domain of the probe (which we take to be its complement, following Chomsky 2000). By definition, an intervener must be the highest matching element in the search domain of the probe If there is a higher NP that is accessible to the probe (the SU in (66)) then *it* should have been the first match, thus obviating the intervention effect altogether. The very existence of an intervention effect in these cases would seem to contradict our basic assumptions about locality: there should be no intervention effect in these constructions because by definition an intervener cannot be structurally lower than the second-cycle controller. Yet, by all independent measures (binding, θ-theory, wh-movement), the controller is superior to the intervener. For example, in Georgian, the SU controller binds the DO intervener, but not vice versa

68) Georgian binding (McGinnis 1997:5-6)

a.	vano	tavis tav-s	xaṯav-s
	Vano(NOM)	self-DAT	draw-PRES
	'Vano draws himself.'		
b.	*vano-s	tavis-i tav-i	xaṯav-s
	Vano-DAT	self-nom	draw-PRES
	'Himself draws Vano.'		

Similarly, in multiple wh-movement constructions the SU must be higher than the DO (reflecting superiority, plus Richard's (1997) tucking in):

69) Georgian wh-movement (Harris 1981:15 (13))

a.	nino	purs	qidulobs	
	Nino	bread	buys	
	'Nino is buying bread.'			
b.	vin	ras	qidulobs?	
	who	what	buys	
	'Who is buying what?'			cf. *Who what buys?
c.	*ras	qidulobs	vin?	
d.	*vin	qidulobs	ras?	cf. Who buys what?
e.	*ras	vin	qidulobs?	

Bejar (2000) proposes that in fact at the point in the derivation where the intervention effect is created, the controller is not in the search space of the probe. The idea is that in (66) there are two cycles of Agree. In the first, because there is no participant in the domain of π-AGR, the uninterpretable π-feature is not deleted (70). The second cycle is triggered by head movement to a higher AGR-head (*v*—>T). By hypothesis, remerger of π-AGR (which still has an undeleted uninterpretable π-feature) is what triggers the second probe, this time with an expanded domain in which the SU (now the closest accessible goal) successfully matches and values (70):[23]

[23] The assumption that a head-adjoined X^0 can probe into the domain of a dominating segment in the adjunction structure is not so innocent. There are obvious locality concerns: if the search space of probe is its sister, then the domain of adjoined X^0 will not necessarily extend into the clause, but rather into the adjunction structure. Roberts 1991 argues for two kinds of head-adjunction, one where each part remains separate, but another where they make a new head. In the latter case, the domain of the probe is unproblematic.

70) a. DO cannot delete uninterpretable π

v	DO
π	π
|	
PART	

b. SU values and deletes uninterpretable π

T+v	SU	t_v	DO
~~π~~	π		π
|	|		
~~PART~~	PART		

Unfortunately, while this accounts nicely for the agreement pattern, the proposal is undermined by the absence of independent evidence for *v*—>T movement in Georgian. Nash 1995 discusses this in depth.

Despite Nash's conclusions, I do not rule out the possibility that *v*—>T occurs without pied-piping the verb (in fact there is no clear evidence that the lexical V even moves to *v* in Georgian, so it could be that V remains low because it is never in *v*). However, I will not seek to pursue a head-movement analysis for Georgian because there is an elegant and more general alternative that has been proposed by Rezac (to appear, 2002).

2.11 Dynamic search domains (Rezac 2003)

At the heart of the paradox posed by (65) is the assumption that the search domain of a probe is static and always restricted to the complement of the target (see discussion in Rezac 2002, to appear). Rezac argues that this assumption is a questionable stipulation because complements are not privileged in Bare Phrase Structure (BPS). Given that under BPS the projection of a head □ is the □ itself, then projection following Merge of a specifier effectively renders the specifier a constituent in the domain of the head (Rezac to appear: 3, 2002). On this basis, Rezac proposes that the most natural implementation of cyclicity would have the search space for feature checking increase throughout the derivation with the growth of a phrase marker. [24]

71) {□, Complement} □ {□, {Specifier, {□, Complement}}}

Cyclicity is, by hypothesis, at the core of the structure of syntactic computation; it determines the order in which structures are built, as well as the order in which operations are peformed on those structures. Chomsky 2000 formulates cyclicity by (72):[25]

72) Properties of the probe/selector □ must be exhausted before new elements of the [Numeration]… are accessed to drive further operations.

By hypothesis, if the properties of □ are not exhausted, the derivation will crash because □ cannot be accessed by the derivation once it ceases to be the root of the tree. A similar

[24] Compare to Chomsky 1995 where feature checking between a head and its specifier is allowed as a last resort.

[25] Compare to Chomsky 1993 where cyclicity is implemented by the principle that all operations must necessarily expand the tree; and Chomsky 1995 where the principle is that a strong feature must be checked *as soon as possible* after being introduced into the derivation, before any new structure is merged. See Richards 1999 for a comparison of these two formulations.

formulation of this is the Earliness Principle of Pesetsky and Torrego (2001:400).[26]

73) Earliness Principle: An uninterpretable feature must be marked for deletion as early in the derivation as possible.

I take feature 'exhaustion' (72) and 'mark[ing] for deletion' (73) to be equivalent. Given present assumptions, exhaustion/deletion is a consequence of value.

Rezac's point is that, without further stipulation, the earliness principle and BPS in fact predict that if no match is found in the complement of □ and □ has a specifier, then the search space of □ naturally expands to include this specifier. This is precisely the scenario in (65) and (66), except that the condition is not failure to match, but failure to value.[27] The effect is the same as that proposed in Bejar 2000: it is not the case that a lower argument is behaving as an intervener with respect to a higher argument. There are two cycles of Agree; on the first it is the intervener that is the closest accessible goal, and on the second it is the controller that is the closest accessible goal.[28]

74) a. π-probes, match succeeds, value fails, Uninterpretable π not deleted

[*v*	[V	DO]]
π	□	π
|		
PART		

[26] This is similar in spirit to the formulation in Chomsky 1995, the differences being a consequence of their respective formulations in terms of 'strong' and 'uninterpretable' features. All strong features in Chomsky 1995 are uninterpretable, but not all uninterpretable features are strong.

[27] In either case, the uninterpretable feature of □ is not deleted and expansion of the search domain is predicted. Rezac is building on the proposals for Georgian in Bejar (2000) where failure to value is not considered and it is assumed that 3rd persons in Georgian are totally underspecified, thus failing match.

[28] In fact, at the level of abstraction at which Rezac 2002 formalizes Project head movement and projection are indistinguishable. Both reduce to second merge of a head □.

b. Merge SU, Project Label *v* with uninterpretable π

[v_{label}	[SU	[*v*	[V	DO]]]]
π	π	π	□	π
|	|	|		
PART	PART	PART		

c. π-probes, match and value succeed, Uninterpretable π deleted

[v_{label}	[SU	[*v*	[V	DO]]]]
~~π~~ □	π	π	□	π
|	|	|		
~~PART~~	PART	PART		

In (74) I label the projection v_{label} only as a matter of expository convenience. There is in fact no difference between the two occurrences of *v*, they are distinguishable only by their sisters. However, had value had been successful on the first cycle of (74), the projection of *v* would not include the deleted uninterpretable □. The generalization would appear to be:

75) Project preserves undeleted uninterpretable properties of □

76) a. π-probes, match and value succeed, Uninterpretable π deleted

[*v*	[V	DO]]
~~π~~	□	π
|		|
~~PART~~		PART

b. Merge SU, Project Label *v*

[v_{lab}	[SU	[*v*	[V	DO]]]]
	π	π	□	π
	|	|		
	PART	PART		

In (76) the uninterpretable properties of *v* are exhausted and no further operations are triggered. Introduction of the next projecting element in the numeration is hence possible, satisfying both Chomsky's (72) and the Earliness Principle (73).

That said, this approach raises various questions concerning the formulation of target deactivation (vis a vis exhaustion/deletion of probes) and its interaction with cyclicity. I turn to these next.

2.12 Target deactivation, cyclicity, and default agreement

We have seen that a target is coded as active by the presence of an unvalued uninterpretable φ feature (i.e. by an unvalued φ-probe).[29] It remains active until its φ-features are valued/deleted by a controller, rendering it inert. By (72) and (73), exhaustion/deletion of every probe is a prerequisite for convergence. However, it actually cannot be the case that failure to match or value will crash a derivation; we know this because there are clear cases where default agreement occurs in grammatical sentences:

[29] This is symmetric to the Active Goal Hypothesis (AGH) for which a goal is coded active by an unvalued Case feature. Assignment of Case to the active goal *deactivates* DP rendering it inert for the purpose of future A-movement.

77) Default number agreement in Georgian in the absence of match

v-xedav-(ø)

1-see-(NUMBER)

'I see him.'

78) Default person agreement in Icelandic in the absence of value

Strákunum	leiddist/*leiddust
the.boys.DAT.PL	bored.3SG/*3.PL

'The boys were bored.' Sigurdsson 1996:1

If the success of match and value are (at least in some instances) irrelevant to convergence, then the conceptual motivation for the Earliness Principle would appear to be weakened. A more accurate generalization is that a probe must be valued and deleted *if it can be* before the tree is extended. Failure to do so when a controller *could have been* found in the domain of the probe is what crashes a derivation. It would seem that a better formulaton would be:

79) Revised Earliness Principle:

An uninterpretable feature must initiate a probe as early in the derivation as possible.

However, it may be possible to restore the EP to its full strength, given a theory of default agreement. I will sketch the beginnings of a proposal here. Let us take as given that a default agreement mechanism *is* in fact made available by the computation to rescue/delete any unvalued uninterpretable features. A straightforward assumption is that default agreement is deletion without valuing. The immediate question is, when can this happen? It cannot be the case that a probe is deleted as soon as match fails. This would render anti-superiority effects (e.g. the number agreement patterns in Georgian and Dakota) impossible. Nor can it be the case that a probe is deleted as soon as value fails, as there would then be no second cycle of Agree. It must therefore be the case that default agreement is a last resort operation, the next questions being how late it can wait, and what consequence waiting has for our ideas about target deactivation. I address each of these in turn.

There is reason to believe that there are in fact two types of default agreement: partial default agreement (the feature is impoverished, but remains active) and total default agreement (the feature is deactivated). Partial default agreement is triggered upon the failure of value. The idea is that this occurs either when a probe is halted by an intervener, or when (on a given cycle) a probe exhausts its search domain without finding a match. (80) illustrates partial default agreement in an intervention configuration. All but the root of the probe is deleted. This signals failure of value, but still allows a second cycle of Agree if π-AGR projects (81).

80) π-probes, match succeeds, value fails, Partial Default Agreement

[π-AGR [… XP]]

π π

|

~~PART~~

81) Merge Specifier, Project Label *v* with uninterpretable π, π probes

[π-AGR [YP [π-AGR [… XP]]]]

~~π~~ □ π π □ π

| |

PART ~~PART~~

If YP satisfies match and value as in (81), then default agreement is obviated. Otherwise, assuming no further specifiers are introduced, total default agreement will be triggered, meaning the root of uninterpretable π deletes. How late this can occur is an open question. Chomsky's implementation of cyclicity (72) would have it occur before any further element from the numeration is introduced. However, it seems that nothing really hinges on this. An uninterpretable [F] that has failed to find a controller on all cycles is invariably marked for partial default agreement. In effect, partial default agreement serves as a diacritic 'telling' the derivation that an attempt to find a controller was made (satisfying the revised earliness principle (79)). I assume that default agreement is impossible in the absence of this diacritic.[30] If [F] possesses this diacritic, then it makes no difference to the derivation when deletion of the root occurs. There is no risk that [F] will crash the derivation, because it is already marked for deletion. There is no risk that it will trigger counter-cyclic operations because, by hypothesis, once the tree is extended by the addition of new material from the numeration, [F] is no longer active.[31]

[30] Alternatively, one could argue that the nature of a principle is that it is inviolable and so if there are any undeleted uninterpretable features remaining at the end of the derivation, (given the revised earliness principle) it must be the case that all probes were attempted as required.

[31] There may be an empirical basis for deciding this, that follows from our notion of what exhausts a probe. If it could be established that, as proposed in Bejar 2000, remerger of a head □ is in some languages able to trigger a subsequent cycle of Agree, then this would show that total default agreement

Let us now be explicit about our notion of what exhausts a probe, and the active/inactive distinction. Given the proposals just made regarding default agreement, it must be possible to extend the tree despite not having exhausted/deleted all uninterpretable □. Thus, exhaustion of a target must be taken to mean that all probes have been attempted once. In keeping with the proposals made by Rezac (2002, to appear) and Bejar (2000), I assume that evaluation of the aforementioned attempts is with respect to a given occurrence of the AGR-head. If the AGR-head remerges (by projection, and perhaps by head movement) then all remaining undeleted □ must probe again.

These proposals concerning default agreement are still sketchy, and the details are in certain respects irrelevant to the main concerns of this thesis. What is crucial, however, is the assumption that default agreement is licensed by a failed *attempt* to value. Not all undeleted uninterpretable features will crash the derivation. Those that are marked with partial default agreement will obviate a crash. We thus have a somewhat constrained picture of default agreement.

2.13 Second-cycle effects

In this section I revisit second-cycle effects in light of the theoretical apparatus accumulated in the previous sections. Recall from the discussion of Georgian in section 2.9 that it is possible for there to be a morphological correlate of cyclic Agree. In Georgian this was seen in the *m*/*v* alternation that distinguishes 1st cycle controllers from 2nd cycle controllers. I now suggest that this differentiated spell-out of AGR-features is a

---

occurs at the end of the derivation, and not immediately after closure of the functional complex headed by □ and prior to introduction of the next object from the numeration. Unfortunately I do not have a clear case of this with which to settle the question, although in chapter 3 I entertain the possibility that perhaps the situation arises in a narrow range of constructions in Georgian.

reflection of the derivational difference between the structures in (82) and (83) where in the former a controller is found on the first cycle of Agree, and in the latter the controller is found on the second cycle. The underlying structural contrast between *m*- and *v* configurations resides in the following detail. Either the π-probe is [π[PART]], valued with [SPEAKER] on the first cycle of Agree, as in (82), or π-probe is [π] vaued with [PARTICIPANT[SPEAKER]] on the second cycle of Agree, as in (83):

82) *First Cycle*

Probe (F) Goal (F')

π □ π Match, value succeed

| | Spell out *m*-

PARTICIPANT PARTICIPANT

|

SPEAKER

83) i. *First Cycle*

Probe (F) Goal (F')

π □ π Match succeeds,

| Value fails triggering partial default

~~PART~~

ii. *Second Cycle*

Probe (F) Goal (F')

π □ π Match, value succeed

| Spell out *v*-

PART

|

SPEAKER

It is perhaps surprising that spellout should be sensitive to the contrast between these two outcomes. One might expect the two to be treated identically since the actual feature structures are indistinguishable. However we will see in Chapter 3 that *m/v*-type alternations are a robust phenomenon across languages, from which I conclude that it must be the case that the derivation is able to distinguish between the original structure of a probe and consequences of value.[32]

A second prediction of second-cycle effects is that the criterion for value should be weaker on the 2nd cycle than on the 1st, because partial default agreement strips the probe of structure below the root, making the entailment condition on value trivially satisfiable for any goal that satisfies match.[33] We will see in chapter 3 that this mechanism is the cause of person-hierarchy effects in Algonquian languages (and others). To illustrate, let us take a quick look at Algonquian (see chapter 3 for a full discussion). Recall that Algonquian 2nd persons appear to be privileged in the agreement system. Whichever argument is 2nd person will control agreement, regardless of its grammatical function. As we saw before, in (84) the 1st person argument simply cannot control the person prefix. Crucially, however, if the 1st person has a 3rd person as co-argument rather than a 2nd person, then it succeeds in controlling the π-probe (85).

84) a. g-waabm-i 2>1

2-see-DIR.THEME

'You see me.'

[32] Note that while second-cycle effects in some cases seem to differentiate the grammatical relations of 1st cycle and 2nd cycle controllers (e.g. Georgian *m*- for DOs and *v* for SUs), this is not an inherent characteristic of the phenomenon. The discussion of Algonquian in the remainder of this section provides an example where second-cycle effects do not correlate at all with grammatical relations. The difference lies in the presence or absence of total underspecification for default categories.

[33] I thank M. Rezac for suggesting that the diminishing restrictions on value could be derived by partial default agreement. This pattern in Algonquian was the starting point for the proposals on partial default agreement in the previous section.

a. g-waabm-in 1>2

2-see-INV.THEME

'I see you.'

85) a. n-waabm-aa 1>3

1-see- DIR.THEME

'I see him.'

b. n-waabm-ig(w) 3>1

1-see- INV.THEME

'He sees me.'

Clearly, 1st person is able to satisfy conditions on match and value, or 1st person agreement would never be possible and the forms in (85) should be ungrammatical. But 1st person's ability to satisfy conditions on value depends on whether its clause-mate is 2nd person or 3rd person. Suppose that π-probe in Algonquian has the structure shown below:

86) Probe

π

|

PARTICIPANT

|

ADDRESSEE

By match, any argument with π will be an intervener, but by value only a 2^{nd} person object can be a controller, as the feature structures for 3^{rd} and 1^{st} person are both proper subsets of the probe and thus fail the entailment condition.[34]

87) 3   1   2

3	1	2
	π	π
	|	|
	PART	PART
		|
		ADDRESSEE

88) Outcomes of match and value: First cycle of Agree

	Probe	Goal	Match	Value
a.	[π[PART[ADDRESSEE]]		no	no
b.	[π[PART[ADDRESSEE]]	[π[PART]]	yes	no
c.	[π[PART[ADDRESSEE]]	[π[PART[ADDRESSEE]]	yes	yes

If we assume that π-AGR in Algonquian is low, so that only a DO is in its domain on the first cycle of Agree, then we predict that only 2^{nd} person objects can be 1^{st} cycle controllers. If the DO is 1^{st} or 3^{rd} person, π-AGR undergoes partial default agreement, and so on a second cycle, π-AGR will be minimally specified, and the outcomes of match and value will be predicted to be the following:

[34] Note that I assume total underspecification of 3^{rd} persons in Algonquian. This is discussed further in chapter 3.

89) Outcomes of match and value: Second cycle of Agree

	Probe	Goal	Match	Value
a.	[π]		no	no
b.	[π]	[π[PART]]	yes	yes
c.	[π]	[π[PART[ADDRESSEE]]	yes	yes

The contrast between the first cycle conditions (88) and second cycle conditions (89) precisely predicts the agreement patterns in (84) and (85).

90)	Argument structure	outcome of 1st cycle	outcome of 2nd cycle	Example
a.	2>1	Match Succeeds Value Fails Partial Default	Highest accessible goal is 2nd person SU. SU matches and values	(84)
b.	1>2	2nd person DO matches and values	N/A	(84)
c.	1>3	Match Fails Partial Default	1st person SU matches and values	(85)
d.	3>1	Match Succeeds Value Fails Partial Default	3rd person is totally underspecified so 1st person DO is highest accessible goal. It matches and values	(85)

2.14 Summary: Features, Underspecification and Locality

The outcomes of match and value combine to create three possible scenarios:

91) A. Failure to match (resulting from total underspecification)

B. Match without value: (resulting from partial underspecification)

C. Match with value: (resulting from full specification relative to P)

The fundamental factor that determines which of A, B, C will arise at any particular point in the derivation is underspecification. Failure to match arises from total underspecification; match without value results from partial underspecification with respect to the probe, and match with value arises when the -specification of the goal is a superset of the probe.

The consequences of C are fixed: it exhausts a probe. The consequences of A and B depend on independent variables in the derivation. There is no unique locality effect correlated with (A) or (B). Failure to match (A) a *bystander* may result in match with a lower NP, or last resort default agreement, or cyclic expansion of the search domain and match/value with a higher argument. These three outcomes are also potential consequences of match without value (B), the only difference being that agreement with a lower NP is contingent upon the intervener displacing.

By hypothesis, match and value constitute the core infrastructure of the system; they provide the uniformity that anchors and unites agreement systems across languages. The locus of variation lies in the properties of the features over which match and value are defined. In the next chapter I show that there are three main parametric dimensions underlying crosslinguistic variation: the specification of interpretable , the specification of uninterpretable , and the initial distribution of uninterpretable .

Chapter 3 Agreement: Uniformity and Variability

3.0 Introduction

In the previous chapter I explored the core operations that enter into Agree: match and value. It was proposed that what would otherwise be syntactically intractable agreement patterns can be accounted for if conditions on match and value are sensitive to a fine grain of feature structure. By hypothesis, the conditions on these operations are uniform across languages, as are the logically possible consequences when either or both operations fail to apply:

1) Consequences for locality:

> A. downward *extension* of the search space past a non-intervener (*bystander*) or past a displaced intervener
>
> B. upward *expansion* of the search space by projection or head movement of □ with undeleted AGR feature(s)

2) Morphological consequences:

> A. Agreement displacement
>
> B. Default agreement

Context-sensitive agreement effects were argued to be amongst the results that are naturally allowed in this system. Crucially, not all languages manifest CSA effects, and amongst those that do, not all manifest them in the same way. In this chapter I show that the computational core of chapter 2 is the unifying foundation for a rich typology of agreement systems. Variation arises as a consequence of a restricted set of cross-linguistically variable properties of feature systems. We saw in chapter 2 that minimal differences in the specification of φ-features had potentially striking consequences for TCP patterns within a language. We will see here that they are likewise responsible for striking typological differences in the TCP patterns across languages. A core typology of agreement systems will be shown to follow from three points of variation:

3) A. variation in the specification of interpretable φ-feature structures

B. variation in the specification of uninterpretable φ-feature structures

C. variation in the distribution of uninterpretable φ

There appears to be a fixed range of variation within each of these areas, but the combinations allow for impressive typological variety in agreement systems that includes both complex and 'simple' systems. For A the range of variation that will be considered here is just the degree of underspecification for default 'values' of interpretable attributes; as seen in chapter 2, in some languages all categories with default values are interveners (e.g. English (Germanic), Spanish (Romance), Abkhaz, Choctaw), whereas in other languages they are not (e.g. Algonquian, Georgian, Dakota). I will thus distinguish between totally underspecified defaults (*bystanders*) and minimally specified defaults (interveners). For B the range of variation involves the degree of specification for AGR probes; in other words whether probes are minimally specified (e.g. [π]-probe), or whether an AGR head selects a cluster of uninterpretable features related by entailment (e.g. [π[PART[ADD]]]-probe). We will see that there are consistent differences between languages that have minimally specified probes (e.g. English, Spanish, Abkhaz, Choctaw), versus those with more highly specified probes; and furthermore, there is a difference between languages with an intermediate degree of probe specification (Georgian, Dakota, Mordvinian, Karok) and those with a high degree of probe specification (Algonquian). Finally, the range of variation for C involves the actual distribution of AGR probes across functional heads in the clausal domain; variation in C affects the establishment of initial AGR-domains (abstracting away from 2nd-cycle effects), with crucial consequences for locality. Contrary to standard assumptions, will see that there is considerable freedom (up to a point) in the initial distributions for uninterpretable φ attested across languages, though there are gaps which I address in section 3.11. I will focus on what appears to be a major split between languages with just one set of φ-features (Georgian, Algonquian, Mordvinian, Karok, Dakota) and those with more than one set of φ-features (Choctaw, Abkhaz, Germanic, Romance). Various permutations on these two classes of φ-distribution patterns are attested; I will focus on 4 salient sub-types. These four types correlate not only with varieties of CSA effects, but also (at least partly) with traditional typological distinctions between ergative, accusative and active agreement systems.

The parametric space defined by the full cross-classification of variants of (3) is very large. It is beyond the scope of this thesis to provide examples of each possible type. First of all, it is not yet clear what the true limit on possible types is, and the range of variants in each of the three categories is deliberately restricted here. Considerations with respect to (A) are being restricted just to the contrast between languages with total underspecification and those without, but a broader range of contrasts could easily have been included (e.g. languages with an inclusive/exclusive distinction versus those without; languages with dual/trial/paucal number contrasts, versus those without). I have restricted (B) to three degrees of specification, and (C) to four patterns of φ-distribution, but this restriction is somewhat artificial in that it reflects just the range of variation attested in a relatively small sample of languages. Even so, the parametric space for the range of variants that *are* being considered here is approximately 70 language types. Instead of showing examples of each potential type of language, I will illustrate each potential type of 'parametric' choice.

To begin, consider just the interaction between points A and B. These two, in and of themselves, are responsible for the outcome (meaning the success or failure) of any given operation match and value, since the evaluation of entailment relations ranges over the specifications of interpretable and uninterpretable φ.[1]

In (4) I give a cross-classification of interpretable (A) and uninterpretable (B) person specifications, with consequences for the derivation listed for each intersection of properties. All but one of the predicted patterns are attested in the sample of languages I have examined. The missing '??' slot I take to be an accidental gap in the sample. Note that two slots at the top left of the grid have identical consequences for the derivation, and it is not clear to me that they would be distinguishable to a learner.

---

[1] They are not, in and of themselves, responsible for the outcome in a larger sense (consequences for locality and morphological consequences) as independent variables such as argument structure, φ-specifications of other arguments, and distribution of uninterpretable φ must also be taken into the consideration.

4) Cross-classification of A and B for person features

	A (Underspecification of interpretable features)	
B (Specification of probes)	**total underspecification**	**relative underspecification**
π	•1 and 2 are interveners •1 and 2 value 1^{st} cycle •3 triggers 2^{nd} cycle •1 and 2 value 2^{nd} cycle e.g. Dakota	•1, 2 and 3 are interveners •1, 2 and 3 value 1^{st} cycle •there are no 2^{nd} cycle effects e.g. English, Spanish, Choctaw, Abkhaz
π[PART]		•1, 2 and 3 are interveners •1 and 2 value 1^{st} cycle •3 triggers 2^{nd} cycle •1, 2 and 3 value 2^{nd} cycle e.g. Georgian
π[PART[ADDRESSEE]][2]	•1 and 2 are interveners •2 values 1^{st} cycle •1 and 3 trigger 2^{nd} cycle •1 and 2 value 2^{nd} cycle e.g. Algonquian	•1, 2 and 3 are interveners •2 values 1^{st} cycle •1 and 3 trigger 2^{nd} cycle •1, 2 and 3 value 2^{nd} cycle e.g. ??

The typological space increases considerably when a second feature is brought in. For example if we cross-classify these with interpretable and uninterpretable specifications for number (with parallel 3-way contrast: sg, pl, du) then there are 36 points in the parametric typology. However I have not found any language which seems to have a number probe based on dual feature specifications, so a more accurate approximation of the typological space would be 24 points. A cross-classification of interpretable and

[2] I leave open the possibility that in some languages the uninterpretable π might select for a speaker rather than an addressee. This would account for languages with person-hierarchy effects where 1^{st} persons are privileged (e.g. Kashmiri, cf. Nichols 2001) in being the only NPs able to value a probe on the 1^{st} cycle..

uninterpretable number features is shown below. Note that here too there is an ambiguity between two of the cells.

5) Cross-classification of A and B for number features

<table>
<tr><td></td><td colspan="2">A (Underspecification of interpretable features)</td></tr>
<tr><td>B (Specification of probes)</td><td>total underspecification</td><td>relative underspecification only</td></tr>
<tr><td>#</td><td rowspan="2">•pl (and du) interveners
•pl (and du) value 1st cycle
•sg triggers 2nd cycle
•pl (and du) value 2nd cycle

e.g. Georgian, Dakota</td><td>•sg, pl (and du) are interveners
•sg, pl (and du) value 1st cycle
•there are no 2nd cycle effects

e.g. English, Spanish, Choctaw, Abkhaz</td></tr>
<tr><td>#[GROUP]</td><td>•sg, pl (and du) are interveners
•pl (and du) value 1st cycle
•sg triggers 2nd cycle
•sg and pl (and du) value 2nd cycle

e.g. Algonquian</td></tr>
</table>

The most striking differences between agreement systems (in the context of this inquiry) involve the presence vs absence of ϕ-sensitive second-cycle effects. In the next section I go through the reasons why CSA effects are present in certain languages and not others.

3.1 Simplicity versus Complexity

Obviation of second-cycle effects (hence of context-sensitive agreement effects) occurs in just one configuration: where probe is minimally specified (π or #) and there is minimal specification for default values of interpretable features. [3] The crucial characteristic of languages without second-cycle effects is that all categories can be controllers on the 1st cycle of Agree. Thus, there is no sensitivity to permutations on φ-specifications. This is considered to be the 'normal' case, familiar from Romance and Germanic languages (including English)). Thus, in English, a 3.sg subject halts and values agreement. It is not possible for either π-probe or #-probe to see past the SU and enter into an AGR-relation with a more highly specified DO.[4]

6) a. They are eating the cake.

b.*He are eating the cakes.

7) a. She is tutoring me.

b.*She am tutoring me.

The standard assumption for English is that *v* and T are both AGR heads. For now I just assume this to be true (in chapter 5 this claim will be examined more closely). If *v* is an AGR head then it is only ever spelled out as -*ø*. The derivation of a transitive English clause is shown below, with the morphologically vacuous lower AGR relation in grey-type:

---

[3] In chapter 5 I will show that in fact even in these languages second-cycle effects can arise in certain configurations, not as a result of A/B interactions but as a result of 'traditional' intervention effects.

[4] Given the assumption that all φ-values in English successfully match and value, the absence of overtly contrastive agreement morphology with anything other than 3.sg NPs (leaving aside the suppletive *be* paradigm) must be taken to be a superficial morphological fact. In effect, the analysis must be that default AGR in English is realized as –*s* , while valued AGR is –ø.

8)

key:	≠ *no Match*		□ *Match, no Value*		□ *Match, Value*
…	$AGR_{T/NOM}$…	SU	…	$AGR_{V/ACC}$…	DO
	π- □	$\pi_{1/2/3}$		π- □	$\pi_{1/2/3}$
	#- □	$\#_{SG/PL}$		#- □	$\#_{SG/PL}$

Questions arise in intransitive clauses with respect to the lower AGR head. One might simply suppose that (by Burzio's generalization) there is no lower AGR in intransitives because *v* is in some way defective. But in fact this assumption is not necessary. By the analysis of unergative verbs of (Hale and Keyser 1993, 1997a, 1997b, 2002) there is no NP in the domain of the lower AGR and one might expect, as a result, a second cycle of Agree with expanded search domain in which SU would be the highest accessible NP. Nothing rules this possibility out. Assuming that the intransitive verb has no Case to assign (Burzio's generalization), the SU would remain active and able to enter into Agree with the higher AGR at a subsequent stage of the derivation.

9)

key:	≠ *no Match*		□ *Match, no Value*			□ *Match, Value*
$AGR_{T/NOM}$…	SU	…	$AGR_{V/ACC}$…		Ø	AGR_V probes, no match is found π- and #- not deleted
π-	$\pi_{1/2/3}$		π-	≠	Ø	
#-	$\#_{SG/PL}$		#-	≠	Ø	

$AGR_{T/NOM}$…	$AGR_{V/ACC}$		SU …	$AGR_{V/ACC}$…		Ø	AGR_V Projects, with uninterpretable π and #, which probe again. SU matches and values, but remains active because AGR_V cannot assign Case
π-	π- □		$π_{1/2/3}$	π-	≠	ø	
#-	#- □		$\#_{SG/PL}$	#-	≠	ø	
$AGR_{T/NOM}$…	$AGR_{V/ACC}$		SU…	$AGR_{V/ACC}$…		Ø	AGR_T probes π and #. SU matches and values.
π- □	π-	□	$π_{1/2/3}$	π-	≠	ø	
#- □	#-	□	$\#_{SG/PL}$	#-	≠	ø	

Likewise, in unaccusative intransitives, nothing rules out the possibility that the lower AGR is controlled by the logical DO. Because English AGR_V is morphologically null, these instances of 'double agreement' are vacuous. Because intransitive AGR_V is unable to assign Case, there is no affect on subsequent AGR relations.

Absence of agreement displacement/second-cycle effects is also attested in less familiar languages, for example Abkhaz (North Caucasian), and Choctaw (Muskogean) shown below. Davies (1986) shows that Choctaw allows agreement with multiple arguments (subjects, objects and various applicatives/IOs). Choctaw has three core sets of agreement affixes, each of which I assume corresponds to a distinct AGR head, as in (10).[5] The data below show that NPs of all φ-specifications are able to value on the 1st cycle of Agree and so there is no agreement displacement triggered by φ permutations. (11) (a) and (b) show that SU and DO agreement can co-occur, so each must reflect a

[5] Davies labels each of the AGR types by Case (i.e. nominative agreement, accusative agreement, dative agreement) however the actual relation between Case and agreement in the language is more intricate than this labeling system might lead one to believe (as Davies shows in great depth). Crucially, for us, these intricacies are linked to Case/argument structure properties of individual verbs or verb classes; they are not related to φ-syntax.

separate probe, the goal of which successfully values on the 1st cycle. (I abstract away from the difference in morpheme order in (a) and (b); what is important is that the AGR relations are controlled independently of one another). Likewise, the fact that DO and IO agreement co-occur in (c) shows that they must enter into separate AGR relations. Note that there is an overt morphological reflex for 3rd person IO agreement (c), but not for 3rd person DO agreement (d); I take this to be a superficial morphological fact. The important point is that 3rd persons in all cases are interveners, they do not allow a probe to extend past them, nor do they trigger second-cycle effects.

10) Choctaw AGR distribution

… $AGR_{T/NOM}$… SU … $AGR_{V/DAT}$… IO … $AGR_{V/ACC}$… DO

11) Choctaw (Davies 1986:2-3)

a.

chi-	bashli	**-li**	-tok
2(ACC)	cut	1(NOM)	PST

'I cut you.'

b.

ano	**is-**	**sa-**	hottopali	-tok
I	2(NOM)	1(ACC)	hurt	PST

'You hurt me.'

c.

alla	towa	**ish-**	**i-**	pila	-tok
child	ball	2(NOM)	3(DAT)	throw	PST

'You threw the ball to the child.'

d.	an-	at-	o	iskali	**chim-**	a:	**-li**	-tok
	I	NOM	CONTR	money	2(DAT)	give	1(NOM)	PST

'I gave the money to you.'

e.	ofi	**is-**	**sami**	lhioli	-tok
	dog	2(NOM)	1(BEN)	chase	PST

'You chased the dog for me.'

The absence of second-cycle effects in Choctaw arises, by hypothesis, because all π and # AGR features are minimally specified, as are all default values of interpretable features, thus obviating second-cycle effects because match and value are satisfied on the 1st cycle of Agree, no matter what the □-configuration of the highest accessible NP.

12)

key:	≠ *no Match*		□ *Match, no Value*			□ *Match, Value*		
$AGR_{T/NOM}$…		SU …	$AGR_{V/DAT}$…		IO …	$AGR_{V/ACC}$…		DO
π-	□	$\pi_{1/2/3}$	π-	□	$\pi_{1/2/3}$	π-	□	$\pi_{1/2/3}$
#-	□	$\#_{SG/PL}$	#-	□	$\#_{SG/PL}$	#-	□	$\#_{SG/PL}$

We can contrast English and Choctaw with a language like Dakota, where no default values are interveners, and agreement displaces between SU and DO as a consequence. Hence the three-way ambiguity in (13) and the four-way ambiguity in (14):[6]

[6] Note that (13) cannot have a meaning where the 2nd person is the agent because a second-cycle effect in the morphology disambiguates 2nd person DOs and SUs. Likewise, (9) cannot mean 'I bind them'. I assume this is because a second-cycle effect disambiguates 1.sg DOs and SUs.

13) ni-castaka-pi Dakota (Buechel 1939)

2-strike-PL

'He strikes you.all'

'They strike you'

'They strike you.all'

14) un-kaska-pi

1-bind-PL

'He binds us.'

'They bind me.'

'We bind him(inan.).'

'We bind them(inan.).'

In order to properly show the derivation of a language like Dakota, the properties of uninterpretable must be made concrete, in particular the distribution of in the clause. I turn next to a discussion of this property, which is a crucial point of variation across agreement systems.

3.2 The distribution of uninterpretable

Recall that the conventional assumption is that unintepretable are distributed on the core functional projections (v^0, T^0) as in:

15)	T^0	…	SU	…	v^0	…	DO
	π-				π-		
	#-				#-		

Abstracting away from context-sensitive agreement effects, the closest NP goal in the domain of AGR_v is DO and the closest NP goal in the domain of AGR_T is SU, making T^0 the locus of subject agreement and v^0 object agreement. In such a φ-configuration, there are two AGR domains in the clause (sister of T^0 and sister of v^0) each correlated with a distinct functional complex (in the sense of Chomsky 1995).

I argue here that this is just one of the possible φ-distribution patterns found across languages. Choctaw, we saw, had a third AGR domain introduced within *v*P, by hypothesis a consequence of the applicative *v* hosting uninterpretable φ. We will now see that there also exists a robust class of languages in which there is just one AGR domain, i.e. just one uninterpretable φ-set is introduced in a clause. I refer to these henceforth as single-φ languages. It is in these languages that we most clearly see the effects of agreement displacement, etc., because there is a unique π-AGR and #-AGR, for (potentially) multiple arguments. Thus, whereas a language like Choctaw has a one-to-one ratio of AGR heads to arguments, in polyadic clauses of a single-φ language there is a one-to-many ratio between AGR heads and arguments. (This is what results in there seeming to be 'competition' between arguments with respect to agreement). Sticking with the assumption that the core AGR heads in a clause are v^0 and T^0, then we might expect to find two classes of single-φ languages: those with a high φ-set, and those with a low φ-set. (cf. Levin and Massam 1985, Chomsky 1993, Bobaljik 1993, Hale 2002).

16) High-φ and Low-φ

A.	high-φ	T^0 π- #-	…	SU	…	v^0	…	DO
B.	low-φ	T^0	…	SU	…	v^0 π- #-	…	DO

In a high-φ language we expect the SU to be the usual controller of agreement, with DO being a controller *only* if the SU cannot match/value. In a low-φ language, on the other hand, we expect the DO to be the usual controller, with the SU controlling agreement only if the DO cannot match/value. We will see that Algonquian, Karok and Dakota are lo-φ languages. Crucially, I have not found any languages that are consistent with the expected high-φ pattern.[7] While the reasons for this are largely speculative at this point, my working hypothesis is that this reflects an imperative that functional heads entering the derivation must select AGR-features as soon as possible. Thus, it is impossible for v^0 — being a potential AGR head — to enter the derivation without any uninterpretable φ (assuming that at least one set of AGR features is available; presumably if a language does not have AGR features than a φ-bare v^0 should be possible (cf. Massam 2002). If there is only one φ-set available in the numeration, φ-bare T^0 is permitted. The descriptive generalization (17) is consistent with this pattern:

17) No functional head is required to be an AGR head, but if there are AGR features available in the numeration, then they must enter the derivation as soon as possible (see also Bejar and Rezac in prep).[8]

I will return to (17) in section 3.5. In the next section I show how Nishnaabemwin (Algonquian) illustrates the low-φ pattern.

Note: that In the appendices to this chapter, complete TCP configurations for all the φ-distribution types discussed here are schematized. I will refer to this throughout this chapter when necessary. One fact that the appendices make clear is that each of the φ-distributions derives a unique set of TCP outcomes. More accurately, each of the φ-distributions allows (in principle) the same set of TCP outcomes, but the configuration of arguments that triggers each TCP pattern is distinctive (although there is convergence at various points).

3.3 Low-φ: Nishnaabemwin (Algonquian)

I begin with an analysis of the person features. Consider, to start, the inventory of personal pronouns:[9]

---

[7] It is tempting to suppose that perhaps languages like English and Spanish are in fact high-φ languages with no agreement displacement for the reasons discussed above. I will show that this cannot be the case in Chapter 5.

[8] This makes the prediction that in languages where C^0 is an AGR head, T^0 and v^0 must also be AGR heads. I have not had a chance to seek out possible counterexamples.

[9] In Nishnaabemwin, verbal agreement affixes are the default means of realizing pronominal arguments (Valentine 2001:609-610). The pronouns themselves only surface in focus constructions (i) or in elliptical fragments where the verb is missing (ii).

18) Personal pronouns (Valentine 2001:122)[10]

Abbreviation	Short Form	Long Form	Categories (Descriptive)	Translation
1	nii	niin	First Person Singular	'I'
2	gii	giin	Second Person Singular	'you'
3	wii	wiin	Third Person Singular	'he, she, it (anim.)'
1p	niinwi	niinwin	First Person Plural Exclusive	'we (but not you)'
21	giinwi	giinwin	First Person Plural Inclusive	'we (including you)'
2p	giinwaa		Second Person Plural	'You (all)'
3p	wiinwaa		Third Person Plural	'they'

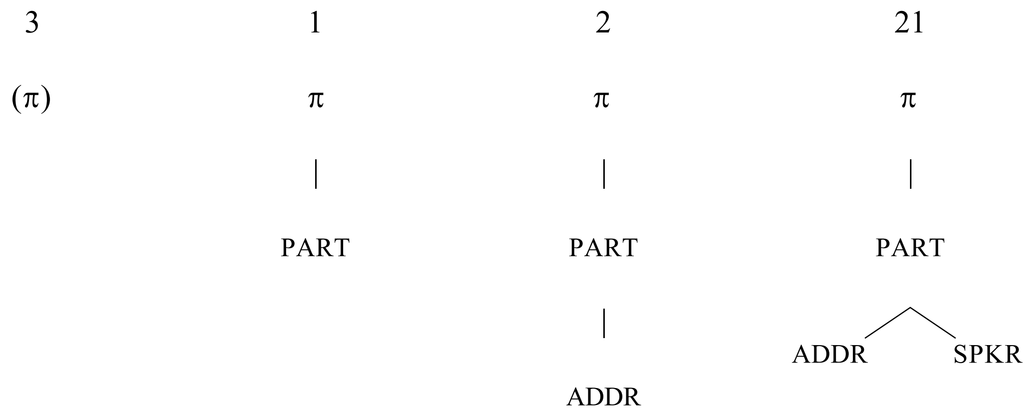
The core contrasts in the inventory of personal pronouns are between three persons (1, 2, 3) and two numbers (Sg, Pl) plus an additional contrast between 1st person Exclusive and Inclusive forms, which I will treat here as a person contrast (following H&R) that necessarily has a nonsingular interpretation. On the basis of the pronominal inventory alone, the interpretable □-specifications for person are plausibly the following:

(i) Mii maanda sa ge-waabndameg **giinwaa** waabizhaayeg
AUX this 2-will.see you.PL 2.pl.who.will.come.
'Such are things you will see, **you** who will come there.' Valentine 2001:610

(ii) Aapii dash **giin** gwiiwkwaaan?
AUX where you your.hat
'And you, where is **your** hat? Valentine 2001:609

[10] According to Valentine the contrast between short and long forms is not grammatical but merely a matter of inter-speaker variation (Valentine 2001:122-123).

19) interpretable π-specifications:



Note that there is no basis yet for determining whether 3rd person is totally underspecified or minimally [π]. The π-agreement prefixes for independent indicative transitive verbs with animate objects are given in (20), followed by a subset of the paradigm showing all possible cross-classifications of person arguments (21).

20) person prefixes (Valentine 2001:269-274)

1	n(i)-	
2	g(i)-	
3	w/o-	(=default π)

21) Sub-Paradigm

waabm 'to see'	3	1	2
3	w-waabm-aa-n 3-see-DIR.TH-*n* 'He sees him'	n-waabm-ig(w) 1-see-INV.TH 'He sees me'	g-waabm-ig(w) 2-see-INV.TH 'He sees you'
1	n-waabm-aa 1-see-DIR.TH 'I see him'	—	g-waabm-in 2-see-INV.TH 'I see you'
2	g-waabm-aa 2-see-DIR.TH 'You see him'	g-waab-am 2-see-DIR.TH 'You see me'	—

The descriptive generalization that is most commonly made about this pattern, which is robust throughout the Algonquian languages, is that the choice of prefix is controlled by the following hierarchy:[11]

22) 2 > 1 > 3

Regardless of grammatical function, the argument that is highest in the hierarchy will control π-agreement. If the controller is the subject, then the root will appear with what is known as a *direct* theme suffix. If the controller is object, then the root will appear with a so-called *inverse* theme suffix.

[11] The hierarchy actually covers a wider range of categories, however I have abridged them to just those that are relevant at this point in the discussion.

The formal status of such hierarchies is an unresolved question. Under the analysis to be proposed here, the question disappears. Person Hierarchy (PH) effects are just a special case of agreement displacement that arises as a consequence of match and value under the conditions given in (23)-(25):

23) 3rd person is totally underspecified

24) Specification of uninterpretable π is:

π-

|

PART-

|

ADDRESSEE-

25) There is just one uninterpretable π in a clause, and it is low:

… T … SU … *v* … DO

π-

Only a 2nd person DO can value the probe in (58). 1st and 3rd persons do not satisfy the entailment condition on value, as they are proper subsets of the probe. So, when the DO is 2nd person it controls agreement. Otherwise value fails and a second cycle of Agree is triggered. Due to partial default agreement after the failed first cycle, conditions on value are less strict for the second cycle. If the SU is 1st person, it can match and value. If the SU is 3rd person (totally underspecified) then it cannot intervene, making the DO the highest accessible argument. If the DO is 1st person, it will match and value. Otherwise, default (3rd person) agreement will occur. The derivation for each of the direct cases is given in (26)-(29). The derivations for the inverse cases follow in (30)-(32).

26)

2>3

key: ≠ *no Match* □ *Match, no Value* □ *Match, Value*

	…	SU	…	*v*	…	DO	•π-probes, no match found •Partial default agreement
		π		π-	≠		
		PART		~~PART-~~			
		ADD		~~ADD-~~			
v	…	SU	…	*v*	…	DO	•projection of *v* •π-probes, second-cycle •match and value by SU •predicted spell-out: *g-*
π-	□	π		π-	≠		
		PART		~~PART-~~			
		ADD		~~ADD-~~			

27)

2>1

key: ≠ *no Match* □ *Match, no Value* □ *Match, Value*

	…	SU	…	*v*	…	DO	•π-probes •DO matches but can't value •Partial default agreement
		π		π-	□	π	
		PART		~~PART-~~		PART	
		ADD		~~ADD-~~			
v	…	SU	…	*v*	…	DO	•*v* projects •π-probes, second-cycle •match and value by SU •predicted spell-out: *g-*
π-	□	π		π-	□	π	
		PART		~~PART-~~		PART	
		ADD		~~ADD-~~			

28)

1>3							
key:	≠ *no Match*			□ *Match, no Value*			□ *Match, Value*
	…	SU	…	*v*	…	DO	•π-probes, no match found •Partial default agreement
		π		π-	≠		
		PART		~~PART-~~			
				~~ADD-~~			
v	…	SU	…	*v*	…	DO	•*v* projects •π-probes, second-cycle •match and value by SU •predicted spell-out: *g-*
π	□	π		π-	≠		
		PART		~~PART-~~			
				~~ADD-~~			

29)

3>3							
key:	≠ *no Match*			□ *Match, no Value*			□ *Match, Value*
	…	SU	…	*v*	…	DO	•π-probes, no match found •Partial default agreement
				π-	≠		
				~~PART-~~			
				~~ADD-~~			
v	…	SU	…	*v*	…	DO	•*v* projects •π-probes, second-cycle •No match found; Partial Dflt •predicted spell-out: w- (dflt)
π	≠	π		π-	≠		
				~~PART-~~			
				~~ADD-~~			

30)

3>2							
key:	≠ *no Match*			□ *Match, no Value*			□ *Match, Value*
	…	SU	…	*v* π- PART- ADD-	… □	DO π PART ADD	•π-probes •DO matches and values
v	…	SU	…	*v* π- PART- ADD-	… □	DO π PART ADD	•projection of *v* does not include active π-probe • predicted spell-out: *g-*

31)

1>2							
key:	≠ *no Match*			□ *Match, no Value*			□ *Match, Value*
	…	SU π PART	…	*v* π- PART- ADD-	… □	DO π PART ADD	•π-probes •DO matches and values
v	…	SU π PART	…	*v* π- PART- ADD-	… □	DO π PART ADD	•projection of *v* does not include active π-probe • predicted spell-out: *g-*

32)

3>1							
key:	≠ *no Match*			□ *Match, no Value*			□ *Match, Value*
	…	SU	…	*v*	…	DO	•π-probes •DO matches but cannot value •Partial default deletes [ADD-]
				π-	□	π	
				~~PART-~~		PART	
				~~ADD-~~			
v	…	SU	…	*v*	…	DO	•*v* projects •π-probes, second-cycle •SU cannot intervene • match and value by DO •predicted spell-out: *n-*
π-				π-	□	π	
				~~PART-~~		PART	
				~~ADD-~~			
□							

The sub-paradigm (21) excludes cases with a 1st person inclusive argument. The □-specification adopted in (19) for 1st person inclusives predicts that these should pattern with 2nd persons in being able to satisfy both conditions on match and on value. This prediction is born out in that a 1st person inclusive argument always controls π-agreement (33)(in fact, this relation is even spelled out with the *g-* prefix of 2nd person agreement). Note that 1st person inclusive arguments can only co-occur with 3rd person arguments. Reflexive argument combinations (1>1, 2>2) are prohibited; because the 1st person inclusive contains both 1st person and 2nd person combinations, the restriction against reflexive combinations disallows combinations of the 1st person inclusive with either 1st or 2nd person co-arguments.

33) a. g-waabm-aa-naan 21>3

2-see-DIR.TH-*naan*

'We(incl.) see him'

b. g-waabm-igo-naan 3>21

2-see-INV.TH-*naan*

'He sees us(inclusive)'

3.4 Nishnaabemwin number agreement

Like π-agreement, #-agreement in Nishnaabemwin is low. To appreciate this, however, the core #-agreement system must be differentiated from the suffixal gender/obviation system which is also contrastive for number under certain conditions. Fortunately, this confound only arises in clauses with 3rd person arguments (these are referred to as *nonlocal* clauses in the Algonquian literature). For present purposes I restrict my attention to *local* constructions, where all arguments are participants and the core #-agreement pattern is transparent. Consider the following subparadigm:[12]

---

[12] Note the special form for 1st person plurals. 1st person plurals cross-linguistically tend to have special forms. Reasons for this are obscure, but they may have to do with the special heterogeneous nature of this category (i.e. 1.pl does not mean 'more than one me'). See discussion in H&R 1998.

34) local direct agreement in the indicative (positive) transitive animate paradigm (Valentine 2001)

2>1	SG.DO	PL.DO
SG.SU	**g**-waabm-[i] 2-see-[TH] 'you see me'	**g**-waabm-i-**min** 2-see- TH-1.PL 'you see us'
PL.SU	**g**-waabm-i-**m** 2-see-TH-PL 'you.all see me'	**g**-waabm-i-**min** 2-see- TH-1.PL 'you.all see us'

The core number agreement pattern can be described as follows: the DO blocks the SU from controlling number, unless the DO is singular. In Appendix 3 it can be seen that this intervention pattern is uniquely characteristic of low-□. It is only for low-□ that patterns D and A (defined in Appendix 1) — both of which include the TCP <#-AGR, SU> — are restricted to □-configurations where the DO is singular. Note that this pattern holds true across tense and mode permutations:

35) local direct agreement across transitive animate paradigms (Valentine 2001:Ch.7)

	SG.DO		PL.DO	
SG.SU	g-waabm-i-naa-ban	Pret., Pos.	g-waabm-i-min-naa-ban	Pret., Pos.
	g-waabm-i-naa-dig	Dub., Pos.	g-waabm-i-min-naa-dig	Dub., Pos.
	g-waabm-i-siin	Neut., Neg.	g-waabm-i-sii-min	Neut., Neg.
	g-waabm-i-siin-aa-ban	Pret., Neg.	g-waabm-i-sii-min-aa-ban	Pret., Neg.
	g-waabm-i-siin-aa-dig	Dub., Neg.	g-waabm-i-sii-min-aa-dig	Dub., Neg.

PL.SU	g-waabm-i-mwaa-ban Pret., Pos. g-waabm-i-mwaa-dig Dub., Pos. g-waabm-i-sii-m Neut., Neg. g-waabm-i-sii-mwaa-ban Pret.,Neg. g-waabm-i-sii-mwaa-dig Dub., Neg.	g-waabm-i-min-naa-ban Pret., Pos. g-waabm-i-min-naa-dig Dub., Pos. g-waabm-i-sii-min Neut., Neg. g-waabm-i-sii-min-aa-ban Pret., Neg. g-waabm-i-sii-min-aa-dig Dub., Neg.

Note that on the basis of this pattern alone we cannot determine the specifications for interpretable and uninterpretable #-features. Clearly, a singular DO creates the conditions for a second cycle of #-agreement, however this could arise either because interpretable # is totally underspecified and unable to match, or because uninterpretable # is more highly specified than (minimally specified) singular, in which case match succeeds but value fails, triggering a second cycle. I will address this problem shortly. In the derivations that follow, I simply revert to labeling the arguments according to the descriptive categories (Sg, Pl) and I assume there is no match with singular NPs, for the time being. The derivations for this intervention pattern are shown in (36)-(38).

36)

Sg>Sg key: ≠ *no Match* □ *Match, no Value* □ *Match, Value*							
	…	SU	…	*v*	…	DO	•[#-] probes, no match found
		Sg		#-	≠	Sg	
v	…	SU	…	*v*	…	DO	•projection of *v* •[#-] probes, second-cycle
#-	≠	Sg		#-	≠	Sg	•no match found •default agreement predicted

37)

Sg/Pl>Pl key: ≠ *no Match* □ *Match, no Value* □ *Match, Value*							
	…	SU	…	*v*	…	DO	•[#-] probes •DO matches and values
		Sg/Pl		#-	□	Pl	
v	…	SU	…	*v*	…	DO	•projection of *v* does not include [#-]
		Sg/Pl		#-	□	Pl	•NB:number of SU irrelevant

38)

Pl>Sg							
key:	≠ *no Match*			□ *Match, no Value*			□ *Match, Value*
	…	SU	…	*v*	…	DO	•[#-] probes, no match found
		Pl		#-	≠	Sg	
v	…	SU	…	*v*	…	DO	•projection of *v* •[#-] probes, second-cycle •SU matches and values
#-	□	Pl		#-	≠	Sg	

A window on the specification of interpretable and uninterpretable [#] is opened in the inverse configurations of the transitive animate paradigms, shown below:

39) local inverse agreement in the indicative (pos.) transitive animate paradigm (Valentine 2001)

1>2	SG.DO	PL.DO
SG.SU	**g**-waabm-in 2-see- INV.TH 'I see you.'	**g**-waabm-in-**im** 2-see-INV.TH-2.PL 'I see you.all.'
PL.SU	**g**-waabm-igoo 2-see-INV.TH 'We see you.'	**g**-waabm-igoo-**m** 2-see- INV.TH -2.PL 'We see you.all.'

Unlike in the direct pattern, here even if the DO is singular, a plural SU cannot control number agreement in the inverse. Like the direct pattern, this one too is robust across tense and mode permutations, although there is some variation across dialects which I will not deal with here (see Ch. 4). I show in Chapter 4 that the absence of SU-controlled #-agreement in the inverse arises because the π-AGR relation with the DO (which is what defines the inverse as such) triggers movement of the DO to a position that is superior to the SU, creating an intervention effect that blocks the SU from being an accessible goal for the second-cycle of #-agreement. The DO can only be an intervener in this configuration if it is able to match, thus it cannot be the case that singular is totally underspecified. The fact that singular matches but cannot value tells us, furthermore, that the specification for uninterpretable [#-] must be a proper superset of the specification for interpretable singular. Minimally, then, uninterpretable [#-] must be [#-[GROUP-]. The derivation of the 1.pl>2.sg inverse patterns is shown below:

40)

1.Pl>2.Sg ≠ *no Match* □ *Match, no Value* □ *Match, Value*	
… SU … *v* … DO $\pi_{[PART]}$ π- □ $\pi_{[PART[ADD]]}$ $\#_{[GROUP]}$ #~~[GROUP-]~~ □ #	•[π-] probes •DO matches & values π-AGR •[#-] Probes •DO matches but fails to value •Partial dflt deletes [GROUP]
v … DO … SU …*v* … t_{DO} $\pi_{PART[ADD]]}$ $\pi_{[PART]}$ ~~π-~~ □ $\pi_{[PART[ADD]]}$ #- □ # #~~[GROUP]~~ #- □ #	•DO displaces (EPP) to [Spec, *v*P] •projection of *v* no longer includes π-probe, just [#-] •[#-] probes (2nd cycle) •DO intervenes, blocking AGR-relation with SU.

Summary: person hierarchy effects and number agreement asymmetries are accounted for. The surprising result is that agreement with the DO (considered to be 'special', i.e. inverse) is in fact the 'default', with SU-agreement being a second-cycle effect. In this respect Nishnaabemwin agreement is ergative, with transitive objects and intransitive subjects entering into the same AGR relation. The highly specified probe (which privileges 2nd person agreement) introduces a confound that obscures this pattern, but the obfuscation is principled and predictable.

3.5 Split □

We have seen three □-distribution patterns so far: double-□, triple-□ and single (low-□). I will now introduce a fourth, which is a variant of the single-□ class. Generalization (17) prohibits high-□ languages, however there is a class of languages with just one □-set where the individual features of that set are split across more than on AGR head. Such languages have a single □-set, but two AGR domains, one high and one low. In principle, there are two logical split-□ patterns that we might expect to find, assuming a □-set that consists of person and number AGR features.

41) Split-□ set

split-□ I	T^0	…	SU	…	v^0	…	DO
	#-				π-		
split-□ II	T^0	…	SU	…	v^0	…	DO
	π-				#-		

In the first type, the expectation is that person agreement will only be controlled by the SU if the DO fails to match or value π-AGR, and number agreement will only be controlled by the DO if the SU fails to match or value #-AGR. In the second type, the reverse is expected: number agreement should be controlled by the SU only if the DO fails to match or value #-AGR, and person agreement should be controlled by the DO only if the SU fails to match or value π-AGR.[13]

Notably, as before, only one of the two logical possibilities appears to be attested: split-ϕ I. The chance remains that this reflects nothing more than a gap in the sample, however until evidence to the contrary is encountered, I take this gap to reflect a real constraint on the distribution of AGR features. The following refinement on (17) captures both the absence of high-ϕ and split-ϕ II:

42) No functional head is required to be an AGR head, but if there is a π-AGR feature available in the numeration, then it must enter the derivation as soon as possible

The existence of split-ϕ (I) languages (henceforth just split-ϕ) raises interesting questions with respect to the idea of a CFC. In the next section I show that Georgian agreement reflects the split-ϕ pattern.

3.6 Georgian

In the previous chapter I introduced selected pieces of data from Georgian in order to illustrate the empirical motivation for the proposals being developed there. I now provide a detailed analysis of the Georgian agreement system. I demonstrate that a unified picture of agreement and CSA effects within and across conjugations is possible if we posit the properties given in (43)-(45) for the Georgian ϕ-system:

---

[13] This description assumes the availability of agreement displacement and second-cycle effects. If these are obviated (as in Choctaw, English, etc.) then the generalization should be for split-ϕ I that person agreement is always with the DO and number agreement always with the SU (with the reverse expected for split-ϕ II).

43) Georgian interpretable φ-specifications:

person:	3	2	1
	π	π	π
		|	|
		PART	PART
			|
			SPEAKER
number:	SG	PL	
		GROUP	

44) uninterpretable φ-specification:

Person Probe	[π-[PART-]]
Number Probe	[#-]

45) uninterpretable φ-distribution: single φ system with split φ

	high		low	
…	T	…	*v*	…
	#		π	

These specifications combine to yield the following predictions about TCP patterns in a transitive clause (see also appendices 1-3):

46) TCP Predictions

A. The SU controls both person and number agreement (47)) when the DO is 3rd person and the SU is a plural participant

B. The internal argument controls both person and number agreement (48) when the DO is a plural participant and the SU is singular

C. The external argument controls person agreement, and the internal argument controls number agreement (49) when the DO is 3rd person plural and the SU is a singular participant.

D. The internal argument controls person agreement, and the external argument controls number agreement (50) when the DO is a participant (singular or plural) and the SU is plural.

47)

<table>
<tr><td colspan="8">1/2.Pl>3Sg
≠ no Match □ Match, no Value □ Match, Value</td></tr>
<tr><td>T

#</td><td>…</td><td>SU

$\pi_{\text{[PART]}}$

$\#_{\text{[GROUP]}}$</td><td>…</td><td>v

$\pi_{\cancel{\text{[PART]}}}$ □</td><td>…</td><td>DO

π</td><td>•[π-] probes
•DO matches ,but fails to value π-AGR
•Partial default agreement deletes uninterpretable [PART]</td></tr>
<tr><td>T

□ ⟶</td><td>v …

π □</td><td>SU

$\pi_{\text{[PART]}}$

$\#_{\text{[GROUP]}}$</td><td>…</td><td>v

$\pi_{\cancel{\text{[PART]}}}$ □</td><td>…</td><td>DO

π</td><td>•v projects with uninterpretable π and probes 2nd cycle
• SU matches and values
• T probes # and SU matches and values</td></tr>
</table>

48)

1/2/3.sg>1/2.pl ≠ *no Match* □ *Match, no Value* □ *Match, Value*	
T … SU … *v* … DO # $\pi_{([\text{PART}])}$ $\pi_{[\text{PART}]}$ □ $\pi_{[\text{PART}]}$ $\#_{[\text{GROUP}]}$	•[π-] probes •DO matches and values
T … SU … *v* … DO $\pi_{([\text{PART}])}$ $\pi_{[\text{PART}]}$ □ $\pi_{[\text{PART}]}$ # □ ——→ $\#_{[\text{GROUP}]}$	•[#-] probes •DO matches and values

49)

1/2.sg>3.pl ≠ *no Match* □ *Match, no Value* □ *Match, Value*	
T … SU … *v* … DO # $\pi_{[\text{PART}]}$ π~~[PART]~~ □ π $\#_{[\text{GROUP}]}$	•[π-] probes •DO matches ,but fails to value π-AGR •Partial default agreement deletes uninterpretable [PART]
T *v* … SU … *v* … DO π □ $\pi_{[\text{PART}]}$ π~~[PART]~~ □ π # □ ——→ [GROUP]#	•*v* projects with uninterpretable π and probes 2nd cycle • SU matches and values • T probes # and DO matches and values

50)

1/2/3.pl>1/2.sg/pl							
≠ *no Match*		□ *Match, no Value*			□ *Match, Value*		
T	…	SU	…	*v*	…	DO	•[π-] probes •DO matches and values]
#		$\pi_{([\text{PART}])}$		$\pi_{[\text{PART}]}$	□	$\pi_{[\text{PART}]}$	
		#				(#)	
T	…	SU	…	*v*	…	DO	•[#-] probes •SU matches and values
		$\pi_{([\text{PART}])}$		$\pi_{[\text{PART}]}$	□	$\pi_{[\text{PART}]}$	
#	□	#				#	

We will see that the basic Georgian agreement pattern confirms these predictions in all but one case (C), where independent factors intervene.

The interpretable □-specifications in (43) are (for the most part) those predicted by the feature-theory of H&R 2002 for a language with the pronominal inventory that Georgian has, with three persons (1st, 2nd, 3rd) and two numbers (SG, PL). The complete pronominal inventory, including case contrasts, is given below.

51) *Georgian pronouns* (Hewitt 1995:76-77)

<table>
<tr><th></th><th>3sg</th><th>3pl</th><th>2sg</th><th>2pl</th><th>1sg</th><th>1pl</th></tr>
<tr><td>nominative</td><td>is</td><td>isini</td><td rowspan="7">šen</td><td rowspan="7">tkven</td><td rowspan="7">me</td><td rowspan="7">čven</td></tr>
<tr><td>accusative</td><td rowspan="2">ma-s(a)</td><td rowspan="6">ma-t(a)</td></tr>
<tr><td>dative</td></tr>
<tr><td>ergative</td><td>ma-n</td></tr>
<tr><td>genitive</td><td>m-is(a)</td></tr>
<tr><td>instrumental</td><td>m-it(a)</td></tr>
<tr><td>adverbial</td><td>ma-d(a)</td></tr>
</table>

Note in (43) the asymmetry between the representation of default values for person and number. The least specifed (3^{rd}) person is relatively, but not totally, underspecified. The least specified (sg) number, on the other hand, is totally underspecified. The empirical basis for this choice is that singular arguments appear to be altogether transparent to match with respect to the #-probe, whereas 3^{rd} person arguments are interveners with respect to the π-probe.

Crucially, person agreement and number agreement have clearly contrasting locality domains. This is what motivates the □-split. For person agreement, the (now familiar) generalization is that person agreement is controlled by the object, unless the object is 3^{rd} person (i.e. fails to value), in which case the controller may be the subject. This reveals that the specification for uninterpretable π must be [π[PART]] as in (44) because this unites 1^{st} and 2^{nd} persons, to the exclusion of 3^{rd} persons. In (52) a complete agreement paradigm is shown, with person agreement highlighted in the data, and the controller highlighted in the translation. Comparison between the framed column (3rd person objects) and the remainder of the paradigm should clearly demonstrate the accuracy of the descriptive generalization.

52) Present inflection: *xedav* 'see'; arranged by person

Subject	Object				
	1sg	1pl	2sg	2pl	3
1sg	—	—	g-xedav 2-see I see you	g-xedav-t 2-see-PL I see you(pl)	v-xedav 1-see I see him/them
1pl	—	—	g-xedav-t 2-see-PL we see you	g-xedav-t 2-see-PL we see you(pl)	v-xedav-t 1-see-PL we see him/them
2sg	m-xedav 1-see you see me	gv-xedav 1.PL-see you see us	—	—	ø-xedav 2-see you see him/them
2pl	m-xedav-t 1-see-PL you(pl) see me	gv-xedav-t 1. PL-see-PL you(pl) see us	—	—	ø-xedav-t 2-see-PL you(pl) see him/them
3sg	m-xedav-s 1-see-TNS he sees me	gv-xedav-s 1. PL-see-TNS he sees us	g-xedav-s 2-see-TNS he sees you	g-xedav-t 2-see-PL he sees you(pl)	ø-xedav-s 3-see-TNS he sees him/them
3pl	m-xedav-en 1-see-TNS they see me	gv-xedav-en 1. PL-see-TNS they see us	g-xedav-en 2-see-TNS they see you	g-xedav-en 2-see-TNS they see you(pl)	ø-xedav-en 3-see-TNS they see him/them

By locality, this pattern can only arise if (the first cycle of) person agreement occurs at a point in the derivation where the DO is the closest potential goal, as in the schematizations in (47).If [π-] were high, the (incorrect) prediction would be that whenever the SU has the potential to satisfy conditions on match and value, it should do so. This would mean that in the rows with 1st and 2nd person subjects the subject should always be the controller, the object being able to control π-agreement only when the subject was 3rd person. This is clearly not the case, as can be seen by the perfect pattern of object control in all but the framed cells above.

However, although this prediction is incorrect for π-agreement, it *is* correct with respect to the #-agreement marker -*t*. Whenever the external argument is plural it controls –*t* (53)B); the internal argument only controls –*t* if the external argument is not plural (53)A). Note that the distribution of –*t* is not entirely regular: it does not surface when the sole plural argument is a 1st person object or a 3rd plural object, nor does it surface with a 3.pl subject.[14] However within the set of forms that does exhibit –*t* morphology, the generalization holds.

---

[14] For reasons that are not clear to me, 3rd person objects in standard Georgian cannot control number agreement. The generalization would appear to be that failure to value low-□ prohibits the 3rd person object from entering into an AGR-relation with a higher □, even though conditions on match might seem to be satisfiable. Other languages (e.g. Dakota) manifest similar behaviour. Historically, 3.pl object agreement was possible in Georgian, and certain dialects still allow this (Hewitt 1995). Crucially for us, 3.pl objects in these dialects can only control agreement when the subject is singular, as predicted by the split-□ analysis.

53) Present inflection: *xedav* 'see'; arranged by number (singular forms in grey text)

Subject	Object					
	1sg	2sg	3	1pl	2pl	
1sg	—	g-xedav 2-see I see you	v-xedav 1-see I see him/them	—	g-xedav-t 2-see-PL I see you(pl)	
2sg	m-xedav 1-see you see me	—	ø-xedav ø-see you see him/them	gv-xedav 1.PL-see you see us	—	A
3sg	m-xedav-s 1-see-TNS he sees me	g-xedav-s 2-see-TNS he sees you	ø-xedav-s see-TNS he sees him/them	gv-xedav-s 1. PL-see-TNS he sees us	g-xedav-t 2-see-PL he sees you(pl)	
1pl	—	g-xedav-t 2-see-PL we see you	v-xedav-t 1-see-PL we see him/them	—	g-xedav-t 2-see-PL we see you(pl)	
2pl	m-xedav-t 1-see-PL you(pl) see me	ø-xedav-t see-PL you(pl) see him/them	—	gv-xedav-t 1. PL-see-PL you(pl) see us	—	B
3pl	m-xedav-en 1-see-TNS they see me	g-xedav-en 2-see-TNS they see you	ø-xedav-en see-TNS they see him/them	gv-xedav-en 1. PL-see-TNS they see us	g-xedav-en 2-see-TNS they see you(pl)	

It could be objected that in B where both subject and object are plural, the controller of *–t* is ambiguous, for example in a 2.pl>1.pl configuration how do we know it is the subject that controls *–t*? Independent support for the claim that the controller must be the subject can be found in constructions with 3rd person subjects. The tense suffixes *–s* and *–en* appear only with 3[rd] person subjects (singular and plural, respectively). These markers are traditionally included as part of the agreement paradigm, but they occupy a different morphological slot from the default plural suffix –t, and they differ from the other agreement markers in that they are synthetic with tense, as can be seen from the fact that they vary with the tense/aspect/mood of the verb:

54)

	3.sg subject markers	3.pl subject markers
present, future	-s	-en
imperfect/conditional	-a	-nen
aorist	-o	-es

What is instructive about the 3rd person subjects, for our purposes, is their interaction with plural –t agreement. When the 3rd person subject is plural, –t agreement is blocked (55).[15] This supports the claim that when the subject is plural, number agreement with the object is impossible (leaving aside the obvious exception of 1.pl object agreement).

55) a. m-xedav-en 'they see me'

b. gv-xedav-en 'they see us'

c. g-xedav-en 'They see you(sg)/(pl)'

Note also that when the sole plural argument is a 1st person object (56) there is no *–t* doubling the plural agreement expressed by the 1.Pl prefix gv-. This suggests that in the ambiguous case of 2pl>1pl repeated here as ((56)b, it must be the subject that is controlling *–t*.

---

[15] Note:Most previous analyses (Anderson 1992, Harris 1981, etc.) assume that both suffixes are present underlyingly.

56) a. gv-xedav 'You see us'

1.PL-see

b. gv-xedav-t 'You(pl) see us'

1.PL-see-PL

3.7 Second-cycle effects in Georgian

Failure of 3rd persons to value triggers a second cycle of π-Agree, which in Georgian has a morphological consequence referred to here as the *m-/v-* alternation (see also chapter 2). As in Algonquian, all but the root of the probe is deleted. After the SU is Merged and *v* projects (effectively re-merging) the unvalued uninterpretable person feature (now just π) triggers a second cycle of Agree, this time with SU in its search domain. Given the reduced entailment criterion for value, any NP with π (i.e. 1st, 2nd or 3rd person) is now able to control π-agreement.

57) □ = Match + Value □ =Match

T	[*v*		$SU_{3rd/2nd/1st}$	[*v*	DO_{3rd}…]
	π-	□	π	π- □	π
			|	|	
			(PART)	~~PART~~	

The mechanisms by which the morphological Vocabulary Insertion operation recognizes the distinction between first-cycle and second-cycle agreement are discussed in the next chapter.

3.8 Permutations on argument structure in Georgian

In the above sections we have seen how the specification and distribution of (un)interpretable φ-features interacts with the entailment conditions on match and value to derive conjugation-internal CSA effects. I will now show the interaction between the proposed φ-system for Georgian and argument structure permutations in the language.

3.8.1 Triadic predicates: In Georgian, a second *v*P-internal argument can be introduced by an applicative head, as in double-object constructions (DOCs) and causatives. It is the IO — not the DO — of the DOCs in (59) that controls π-agreement in these constructions, because it is the first matching goal in the domain of *v* (58). The same is true for the IO in causative constructions (60).

58)	T	[SU	[*v**	[IO	[v_{appl}	[V	DO]]]]]
	[#-]		[PART-]				

59) a. **m**-acukeb-s am cigns
1-give-PRES this book
He is giving **me** this book

b. **g**-acukeb-t am cigns
2-give-PL this book
He is giving **you**(PL) this book

c. **ø**-acukeb-s tavis cign-s cems stent-eb-s
DFLT-give-PRES self's book-DAT my student-PL-DAT
He is giving his book to **my students** (Harris 1981:213-214)

60)	šen	me	mi-**m**-ačvie	sixarbe-s
	you(ERG)	me(NOM)	PreV-**1**-CAUS.accustomed*	greed-DAT

'You made **me** accustomed to greed.' (Hewitt 1995:557)

If the IO is 3rd person this triggers a second-cycle of π-agreement, which is controlled by the SU.

61) second-cycle data for DOCs:

ma-v-eca

PreV-1-give*

'I gave something to someone.' (Hewitt 1996: 106

Number agreement in triadic predicates is, as usual, preferentially controlled by the SU, with a lower controller possible only if the SU is singular.

3.8.2 monadic predicates: Unergatives, unaccusatives, and passives of transitives all exhibit the same agreement in Georgian (though not the same Case properties: unergative subjects have ergative Case, whereas unaccusative and passive subjects are nominative). Specifically, all three manifest second-cyle (*v*-set) morphology. This is in fact the expected agreement for unergatives, given a view such as Hale and Keyser's (2002). There should be no first-cycle π-AGR morphology because the sole argument in the verbal complex is merged in [Spec, *v*P] and is thus not in the domain of π-probe until the second cycle of Agree. In effect, agreement on unergative predicates should be the same as agreement in transitive clauses with 3rd person objects, which is in fact the case:

62)	T	[SU	[*v**	[V	3/ø]
	[#-]		[PART-]		

63) *transitive with 3rd person object* *unergative* (Hewitt 1996:160)[16]

a.	v-xedav 1-see I see him/them	v-it'are 1-cried* I cried
b.	v-xedav-t 1-see-PL we see him/them	v-it'are-t 1-cried*-PL we cried
c.	ø-xedav 2-see you see him/them	ø- it'are 2- cried* you cried
d.	ø-xedav-t 2-see-PL you(pl) see him/them	ø- it'are -t 2- cried* -PL you(pl) cried
e.	ø-xedav-s 3-see-TNS he sees him/them	ø- it'ar-a 3- cried* -TNS he cried
f.	ø-xedav-en 3-see-TNS they see him/them	ø- it'ar -es 3- cried* -TNS they cried

However with passives and unaccusatives we would expect the sole argument to value on the 1st cycle, under the standard assumption that the base position of such arguments is the same as that of logical DOs. But this is not the case. In fact they agree in exactly the same way as unergatives and transitives with 3rd person objects:

[16] The * in the gloss signifies that the glossed form is actually morphologically complex (in irrelevant ways) and I have not broken it down completely.

64) *passive* (Hewitt 1995:281) / *unaccusative* (Hewitt 1996:130)

	passive	unaccusative
a.	[v]-ič'irebi 1-caught.PASS* [I] am caught	[v]-ikceva 1-fall* [I] fall over
b.	[v]- ič'irebi -[t] 1- caught.PASS*-PL [we] are caught	[v]-at'are-[t] 1- fall*-PL [we] fall over
c.	[ø]- ič'irebi 2- caught.PASS* [you] are caught	[ø]- at'are 2- fall* [you] fall over
d.	[ø]- ič'irebi -[t] 2- caught.PASS*-PL [you(pl)] are caught	[ø]- at'are -[t] 2- fall*-PL [you(pl)] fall over
e.	[ø]- ič'ireb-a 3- caught.PASS*-TNS [he] are caught	[ø]- at'ar-a 3- fall*-TNS [he] fall over
f.	[ø]- ič'ireb -an 3- caught.PASS*-TNS [they] are caught	[ø]- at'ar -es 3- fall*-TNS [they] fall over

Given this pattern I must assume that the DO in passives/unaccusatives cannot value the π-AGR relation. I can only speculate at this point as to the reason for this. This fact may be related to the inability of unaccusative/passive *v* to assign Case to an argument in its AGR-domain. One possible solution to explore is an analysis along the lines of Bejar 2000a, where a second-cycle of π-AGR is triggered by head-movement to T^0. This approach presupposes that by virtue of being adjoined to T^0, a category with the ability to assign Case, the π-AGR relation will be satisfied on the second cycle.

3.8 Georgian Inverse agreement

The so-called Georgian inverse is traditionally seen as triggering a separate agreement pattern. In the current framework, the inverse is more transparently labeled a dative-nominative construction (DNC), where the dative functions in every respect as subject. DNCs arise in psychological predicates, in the evidential/pluperfect conjugations of all verb classes except unaccusatives, and in the passive of ditransitives

Previous analyses of the inverse have assumed either inverted grammatical functions, such that the dative is an object and the nominative the subject (Harris 1981) or language-specific morphological mechanisms to account for the aberrant pattern (e.g. Anderson 1992, Halle&Marantz 1993, Carmack 1997). I will argue that this is just a case of agreement displacement, under conditions that should by now seem familiar.

A complete inverse agreement paradigm is shown below. The first point to note is that, contrary to the core agreement pattern, the π-agreement prefix is uniformly controlled by the higher (dative) argument, unless the dative is 3rd person, in which case the controller is a lower argument (as in the highlighted row at the bottom of the table).

65) Evidential: *u-naxav-s* 'see'

Subject	Object				
	1sg	1pl	2sg	2pl	3
sg	—	—	m-inaxav-xar 1- EV.see-2.AUX I have seen you	m-inaxav-xar-t 1- EV.see-2.AUX - PL I have seen you(pl)	m-inaxav-s 1- EV.see-TNS I have seen him/them
pl	—	—	gv-inaxav-xar 1. PL - EV.see-2.AUX we have.see you	gv-inaxav-xar-t 1. PL -EV.see-2.AUX -PL wehave seen you(pl)	gv-inaxav-s 1. PL -EV.see - TNS we have seen him/them
sg	g-inaxav-var 2-EV.see-1.AUX you have seen me	g-inaxav-var-t 2- EV.see-1.AUX-PL you have seen us	—	—	g-inaxav-s 2- EV.see - TNS you have seen him/them
pl	g-inaxav-var-t 2- EV.see-1.AUX - PL you(pl) have seen me	g-inaxav-var-t 2- EV.see-1.AUX - PL you(pl) have seen us	—	—	g-inaxav-t 2- EV.see -Pl you(pl) have seen him/them
sg	v-unaxav-var 1- EV.see-1.AUX he has seen me	v-unaxav-var-t 1-EV.see-1.AUX - PL he has seen us	ø-unaxav-xar 2-EV.see-2.AUX he has seen you	ø-unaxav-xar-t 2- EV.see-2.AUX he has seen you(pl)	ø-unaxav-s 3-EV.see - TNS he has seen him/them
3pl	—	—	—	—	(u-naxav-t)[17] EV.see - PL They have seen him/them

There are various noteworthy differences between the inverse and basic agreement patterns which are outside the scope of the analysis. For example, I will not discuss the compound nature of the inverse: the incorporated auxiliaries *-xar* and *-var* (these are the 1st and 2nd person forms of the verb 'to be'). Nor will I examine the tense suffixes whose distribution is here restricted to clauses with 3rd person objects (in the core agreement pattern their distribution is restricted to clauses with 3rd person subjects). My focus is on the property that has received the most attention in the literature, and the one that is considered to be the main characteristic of the inverse pattern: that π-agreement is characteristically controlled by the subject rather than the object in the inverse pattern. If we hold constant the distribution of uninterpretable in the clause, by locality this pattern can only emerge if the dative subject is the first potential match in the domain of *v*. In other words, dative subjects must have a lower base position than other subjects.

66) Base structure non-dative subject constructions

... [SU [v* [V DO]]]

π-

67) Base structure for DNCs

.... [v* [DAT [v_{appl} [V DO]]]]

π-

This has previously been proposed for Georgian on independent grounds (Marantz 1989, McGinnis 1997). Dative arguments in Georgian exhibit structural parallels that appear to be independent of their actual thematic role, i.e. whether the dative is a subject (psych-verbs, evidentials) or an IO (applicatives, etc.). First of all, dative subjects are in complementary distribution with other datives, suggesting that there is a unique base structural position for all dative arguments Ditransitive psych-verbs are simply unattested (this is true cross-linguistically). I refer the reader to Marantz 1989, McGinnis 1997 (see p. 5 in particular) and McGinnis (1995) for further discussion of this matter.

---

[17]This form is an innovation. For many Georgian speakers, inverted 3rd person subjects do not show plural marking (Carmack 1997).

Furthermore the word order of DNCs is freer than that of the regular transitives, suggesting that there is indeed an extra position in the former construction. In (68)b, which has the same meaning has (68)a, the object anaphor cannot displace to the left of the embedded non-dative subject, but in (69)b— same meaning as (69)a — the embedded subject is a dative, and displacement of the object anaphor is possible.

68) a. Nino-s unda [rom Lali-m naxos tavisi tavi televisor-shi]

N.-DAT wants that Lali-ERG see(opt) self-NOM television-on

'Nino wants Lali to see herself on TV'

b. *Nino-s unda [rom tavisi tavi Lali-m naxos televisor-shi

N.-DAT wants that self-NOM L.-ERG see(opt) television-on

69) a. Nino-s unda [rom Lali-s achuenos tavisi tavi televisor-shi]

N-DAT wants that L.-DAT watch(opt) self-NOM television-on

'Nino wants Lali to watch herself on television' (field notes)

b. Nino-s$_i$ unda [rom tavisi tavi$_{i/j}$ Lali-s$_j$ achuenos televisor-shi

N-DAT wants that self-NOM L.-DAT watch(opt) television-on

I will now show that the special argument structure of DNCs is sufficient to derive the fundamental difference between core agreement and inverse agreement; v* is still the locus of π-AGR and T the locus of #-AGR. There is thus really only one agreement system in Georgian, but CSA effects arise from locality, match and value due to the unique argument structure permutation in DNCs.

70) core agreement pattern – [Spec, v*] is a thematic position

T...[SUBJ ... v* ... [IO ... *v* ... V ... OBJ]

#- π-

71) inverse agreement — [Spec, v*] is non-thematic

T …v*...[DAT …*v* … [V …OBJ]

#- π-

As in the basic agreement pattern, conditions on match and value conspire to derive the details of the paradigm. When the dative has both π and #, it controls both agreement relations (by locality).

72) (you) mo-**g**-cons-**t** pelamusi

You-DAT preverb-2-like-pl Pelamusi

'You(pl) like Pelamusi.' (Harris)

73)

≠ *no Match* □ *Match, no Value* □ *Match, Value*	
T … *v* … DAT … DO # $\pi_{([PART])}$ □ $\pi_{[PART]}$ $\pi_{([PART])}$ # #	•[π-] probes •DAT matches and values
T … *v* … DAT … DO $\pi_{([PART])}$ □ $\pi_{[PART]}$ $\pi_{([PART])}$ # □ ⟶ # #	•[#-] probes •DAT matches and values

If the subject is a singular participant, then it will control π-agreement, but it will not intervene in #-agreement; #-probe will see past it and can be controlled by lower plural argument.[18]

74) a. m-inaxavar-**t**

1-EV.see-Pl

'I have seen you(**pl**).'

b. mo-consxar-**t**

preV-like-pl

'He likes you(**pl**).'

75)

≠ *no Match* □ *Match, no Value* □ *Match, Value*	
T … *v* … DAT … DO # $\pi_{([PART])}$ □ $\pi_{[PART]}$ $\pi_{([PART])}$ #	•[π-] probes •DAT matches and values
T … *v* … DAT … DO $\pi_{([PART])}$ □ $\pi_{[PART]}$ $\pi_{([PART])}$ # □ ⟶ #	•[#-] probes •DO matches and values

[18] Note that in (65) a 1.pl>2.pl form does not behave in the expected way: the plural SU should block #-agreement with the lower object, but unexpectedly the plural object controls the *–t* plural marker here. In Chapter 4 I will show that this is explained by an interaction between goal deactivation and a special structural property of 1[st] person plurals (across languages).

If the dative is not a participant then it will intervene but not control π-probe. A second-cycle effect occurs here (we will see in chapter 4 that this is enabled by extraction of the dative out of the search space for π-probe), allowing π-agreement to be controlled by the lower NP.

76) a. **v**-unaxav-var

1-EVIDsee-1

'He has seen **me**.'

b. mo-**v**-consvar

preverb-1-like

'He likes **me**.' (Harris)

77)

≠ *no Match* □ *Match, no Value* □ *Match, Value*	
T … *v* … DAT … DO # π[~~PART~~] □ π π([PART]) (#) (#)	•[π-] probes •DAT matches, but doesn't value •Partial dflt agreement deletes [PART]
T …DAT *v* … *t* … DO # π π[~~PART~~] □ π[PART]	•DAT displaces •On second-cycle DO is highest accessible NP; π-probes and DO matches, values[19]

[19] Note that this is a different second-cycle configuration. Instead of waiting for projection, the second-cycle occurs from the same occurrence of *v*. This may be problematic, unless it can be shown that this order of operations follows from some independent factor (a similarly ordered second cycle is assumed in Chomsky's 2001 analysis of Icelandic nominative Case assignment past dative interveners). Alternatively, it might do less damage to the system to explore the original option I proposed for such cases (Bejar 2000) which involved abstract head movement of *v* to T.

3.10 φ-distribution summary

A cross classification of the φ-distribution parameter with the CSA outcome of the φ-specification parameters is shown below, with the languages we have seen so far indicated.

78)

	CSA languages (see A/B interaction)	non- CSA languages (see A/B interaction)
split-φ	Georgian,	
low-φ	Algonquian	
double-φ		Germanic, Romance
triple-φ		Choctaw

The gaps in all but one case can be filled. I briefly review some candidates here:

A. double-φ (CSA) = Iroquoian (Mithun 1991, Dyck 1992, Barrie 2002). Iroquoian languages, are double-φ. There is full 'portmanteau' agreement for with both subjects and objects in canonical transitive clauses. There are also two series of markers for intransitive verbs: agent markers and patient markers. Transitive verbs with a 3rd person singular neuter object act like intransitive verbs with respect to the choice of agreement marker, which suggests a CSA effect for underspecified π-attributes:

B. triple-φ (CSA) = Basque ergative displacement (Laka 1993, Arregi 1999, Rezac 2003). Basque has three-way agreement for person and number with ergative, absolutive and dative arguments. In non-present tense, Basque exhibits a CSA effect for π: if the absolutive argument is 3rd person then the agreement that normally tracks the absolutive will instead track the ergative argument.

C. low-□ (non-CSA)= any canonical ergative agreement pattern, with AGR controlled by DO and intransitive S (this follows from low-□ without CSA in that there will only be a second-cycle with expanded search domain if the verb is unergative; if is unaccusative there will be agreement with the DO (but still this agreement with intransitive S). e.g. ergative agreement pattern in Abkhaz, where same series of markers indexes agreement with transitive DOs and all intransitive S).

I have not yet found evidence for a non-CSA language with split □. This would be a language with (rigid) person agreement with DO and number agreement with SU in transitive clauses. It is not clear to me why this should be a gap.

Note that CSA effects in double-□ and triple-□ languages are marginal. They are restricted to just one □-feature (person in both Mohawk and Basque) and that only in certain tense/aspects. It seems intuitive that this should be the case, just because the complexity of a maximally CSA-affected system with double- or triple-□ would be quite tremendous (see Appendix 4 for a schematization of what the double-□ scenario would be). Thus it would seem that there must be a functional appeal to limits on complexity, but the correct way to articulate this is far from clear.

3.11 Distributing uninterpretable φ

Properties of uninterpretable φ constitute a powerful locus of variation in the proposed system. First, uninterpretable φ-specification is directly responsible for establishing the specific conditions on match and value, since both operations are evaluated relative to the probe. The more specific the probe, the wider the set of second-cycle configurations will be. The less specific the probe, the smaller the set of configurations will be. Second, the distribution of uninterpretable φ establishes initial AGR domains with important consequences for locality. DOs are privileged controllers with respect to low AGR heads, and SUs with respect to high AGR heads, and we have seen that there are various high and low configurations available, subject to the restriction that uninterpretable π-features merge as early in the derivation as possible.

The fascinating question is what, if anything, determines uninterpretable φ-specification and distribution. Unlike variation in the specification of *interpretable* φ, which seems to be at least indirectly correlated with morphological inventories (H&R), variation in *uninterpretable* features is not clearly correlated with any independent property of a language. However, this is a question that has hardly been explored.

For Chomsky (1995, 2000, 2001) the presupposition concerning features is that UG makes available a finite inventory from which a selection is made by the language learner on the basis of primary linguistic data (PLD). By hypothesis, this selection of FFs establishes the building blocks for the assembly of lexical items (LIs), which in turn are terms in the computation. Thus, cross-linguistic differences in the construction of a feature inventory are the basis of variation in the output of the computation, the computational system itself being held to be uniform. There seem to be crucial differences between the selection of interpretable and uninterpretable FFs. Recall from chapter 1 that selection of interpretable FFs can be seen as a form of grammaticalization of contrasts/categories detectable from surface PLD; thus variation in language-specific choices of these affects (within limits) the slicing up of conceptual space, or more precisely, the choice of which surface contrasts relevant (or at least visible) to the computation.[20] Variation in the selection of uninterpretable FFs has a very different consequence. By definition, uninterpretable FFs trigger syntactic operations, so variability in the set of FFs necessarily introduces variation in the set of operations of the computation. The strong minimalist thesis which holds the locus of cross-linguistic variation to be lexical is not so clearly compatible with this possibility. The notion of a uniform computation in the MP view – simplifiying a bit – included uniformity of operations, with parameterization being restricted to whether operations occur covertly or overtly. The position I am taking here is very different from this, but not different in spirit from previous proposals to parameterize the distribution and properties of functional projections across languages thereby introducing variability in the set of operations of the computation (See for example Levin and Massam 1985, Bobaljik 1993, Rice and Saxon 1994, among others). The difference between these proposals and the proposals here is largely just a reflection of the evolution of the framework, which currently allows manipulation of the set of operations that are triggered at any given point in the functional structure, without manipulation of functional structure itself.

[20] In principle, not all contrasts/categories that might be gleaned from the surface morphology of the PLD need be reflected in the set of interpretable FFs.

To conclude, the uniformity of the computation lies in the nature of the operations available (not individual instances of these) and the principles that constrain these (locality, cyclicity, etc.). Thus, fine-grained variation in the selection of AGR features, and variability in their distribution — while having with direct consequences for the derivation (which presumably form the basis for the language learners' selection of one option over another) — must be considered to be 'lexical' in some sense that is not clear. Many questions remain unanswered at this point. For instance, is the inventory of uninterpretable features selected from amongst the interpretable ones (or vice versa), or are the two sets arrived at independently? What determines the distribution of AGR features on LIs? In the MP uninterpretable features are not considered to be inherent in LIs, but rather optional features that are 'selected' by an LI as it enters the numeration. The restriction on early selection of π-agr suggests that this selection process cannot be entirely free. How this is to be explained is a completely open question. One possibility is that the computation itself assembles LIs by Merge, according to principles that can only be guessed at. Entailment relations are a natural starting point for if we restrict our attention to the assembly of interpretable FFs on LIs, however it seems incoherent to suggest that entailment relations could have anything to say regarding the assembly of FFs into LIs on which they are *not* interpretable.

If the specification of interpretable □ can be construed as 'grammaticalization' (see chapter 1), then specification of uninterpretable □ is *hypergrammaticalization*. The selection of uninterpretable □ determines the set of AGR operations that will be triggered, and establishes the criteria by which conditions on match and value are evaluated. It is beyond the scope of this thesis to provide a serious proposal regarding what determines hypergrammaticalization. Simplistic hypotheses are easily disproven. For example, one might suppose that perhaps the most highly specified value available (e.g. [π[PART[ADDRESSEE]] for person) would be selected, or alternatively the least specified (e.g. [π] for person]). This just cannot be the case. In some languages (Germanic, Romance) the latter scenario must be the case, whereas in others (Algonquian, Kashmiri) the former seems closer to the truth. I will assume the working hypothesis that uninterpretable feature specifications must be a subset of the interpretable ones in the language, and that the selection is made once and for all during acquisition. In other words, numerations do not freely select any feature specification from amongst those made available by the inventory of the language. If the uninterpretable person feature that has been selected for a language L is [π[PART]], then all uninterpretable person features will have this specification in the numeration. There will not be competing derivations where one numeration selects a [π[PART]-probe and another just a simple [π]-probe, for example.[21]

There are almost certainly more possibilities than have been discussed here here. In principle, if the selection of uninterpretable □ can be any subset of the interpretable □ than the logical possibilities are tremendous. It seems fair to say that not all logical possibilities are available, but given that there is not even a clear picture of what the range of possibilities are, it is simply premature to consider what conditions license or constrain them.

---

[21] This is hardly a necessary assumption. One could imagine that uninterpretable □-could enter the

Appendix 1

TCPs	Split-□	High-□	Low-□
π-SU #-SU **A**	______ ______ # SU π DO SU: π, # DO: (#)	______ ______ π, # SU DO SU: π, # DO: (π), (#)	______ ______ SU π, # DO SU: π, # DO: ø
π-DO #-DO **B**	______ __________________ # SU π DO SU: (#) DO: π, #	____________ ____________ π, # SU DO SU: ø DO: π, #	______ ______ SU π, # DO SU: (π), (#) DO: π, #
π-SU #-DO **C**	______ __________________ # SU π DO SU: π DO: #	______ ____________ π, # SU DO SU: π DO: (π), #	______ ______ SU π, # DO SU: π, (#) DO: #
π-DO #-SU **D**	_____ _____ # SU π DO SU: (π), # DO: π, (#)	____________ ______ π, # SU DO SU: # DO: π, (#)	______ ______ SU π, # DO SU: (π), # DO: π

---

numeration as specified 'as necessary' for convergence. I will not explore this avenue here.

π-DO #-dflt **d1**	______ ____________________ø # SU π DO SU: (π) DO: π	______________ ______________ø π, # SU DO SU: ø DO: π	______ ______ SU π, # DO SU: (π), # DO: π
π-SU #-dflt **d2**	______ ____________________ø # SU π DO SU: π DO: ø	______ ______________ø π, # SU DO SU: π DO: (π)	______ ______|______ø SU π, # DO SU: π DO: ø
π-dflt #-DO **d3**	______|______ø ____________________ # SU π DO SU:ø DO: #	______________ø ______________ π, # SU DO SU: ø DO: #	______|______ø ______ SU π, # DO SU: ø DO: #
π-dflt #-SU **d4**	______|______ø ________ # SU π DO SU: # DO: ø	______________ø ________ π, # SU DO SU: # DO: (#)	______|______ø ______ SU π, # DO SU: # DO: ø
π-dflt #-dflt **d5**	______|______ø ____________________ø # SU π DO SU: ø DO: ø	______________ø ______________ø π, # SU DO SU: ø DO: ø	______|______ø ______|______ø SU π, # DO SU: ø DO: ø

Appendix 2: Comparison of TCP configurations by □ (organized by SU specs)

SU		DO		Split	High	Low
π	#	π	#	D	A	B
π	#	π	ø	D	A	D
π	#	ø	#	A	A	C
π	#	ø	ø	A	A	A
ø	#	π	#	D	D	D
ø	#	π	ø	D	D	D
ø	#	ø	#	d4	d4	d3
ø	#	ø	ø	d4	d4	d4
π	ø	π	#	B	C	B
π	ø	π	ø	d1	d2	d1
π	ø	ø	#	C	C	C
π	ø	ø	ø	d2	d2	d2
ø	ø	π	#	B	B	B
ø	ø	π	ø	d1	d1	d1
ø	ø	ø	#	d3	d3	d3
ø	ø	ø	ø	d5	d5	d5

Appendix 3: Comparison of TCP configurations by □ (organized by DO specs)

SU		DO		Split	High	Low
π	#	π	#	D	A	B
ø	#	π	#	D	D	B
π	ø	π	#	B	C	B
ø	ø	π	#	B	B	B
π	#	ø	#	A	A	C
ø	#	ø	#	d4	d4	d3
π	ø	ø	#	C	C	C
ø	ø	ø	#	d3	d3	d3
π	#	π	ø	D	A	D
ø	#	π	ø	D	D	D
π	ø	π	ø	d1	d2	d1
ø	ø	π	ø	d1	d1	d1
π	#	ø	ø	A	A	A
ø	#	ø	ø	d4	d4	d4
π	ø	ø	ø	d2	d2	d2
ø	ø	ø	ø	d5	d5	d5

Appendix 4: TCP configurations for double □-set (organized by DO specs)

SU	DO	High	Low	Diagrams
π #	π #	A	B	T SU v DO π—— π π—— π #—— # #—— #
ø #	π #	d4	B	T SU v DO π- ø π—— π #—— # #—— #
π ø	π #	d2	B	T SU v DO π—— π π—— π #- ø #—— #
ø ø	π #	d5	B	T SU v DO π- ø π—— π #- ø #—— #
π #	ø #	d4	C	T SU v DO π- π —— π- ø #—— # #—— #
ø #	ø #	d4	d3	T SU v DO π- ø π- ø #—— # #—— #
π ø	ø #	d5	C	T SU v DO π- π —— π- ø #- ø #—— #
ø ø	ø #	d5	d3	T SU v DO π- ø π- ø #- ø #—— #
π #	π ø	d2	D	T SU v DO π—— π π—— π #- # —— #- ø
ø #	π ø	d5	D	T SU v DO π- ø π—— π #- # —— #- ø
π ø	π ø	d2	d1	T SU v DO π—— π π—— π #- ø #- ø
ø ø	π ø	d5	d1	T SU v DO π- ø π—— π #- ø #- ø
π #	ø ø	d5	A	T SU v DO π- π —— π- ø #- # —— #- ø
ø #	ø ø	d5	d4	T SU v DO π- ø π- ø #- # —— #- ø
π ø	ø ø	d5	d2	T SU v DO π- π —— π- ø #- ø #- ø
ø ø	ø ø	d5	d5	T SU v DO π- ø π- ø #- ø #- ø

Chapter 4: A Few Notes on Morphology

4.0 Introduction

In this chapter I consider certain morphological questions relevant to the morphology-syntax interface. As noted in the Chapter 1. I assume a realizational theory of morphology with postsyntactic vocabulary insertion, along the lines of that proposed by Halle & Marantz (1993). Recall that a formal distinction is posited between lexical items (LIs), which are terms in the computation, and vocabulary items (VIs) which constitute the actual lexical material of a language (what would traditionally be considered morphemes). By hypothesis, the list of vocabulary items in a grammar is indexed to a set of morphological features, which are matched to the FFs that are organized as terminal elements in in the output of the syntactic component. Halle & Marantz assume a post-syntactic module of grammar that is a morphological component, where purely morphological operations may apply (etc. node insertion, node deletion, fusion, merger, fission, impoverishment). I take these operations to be the 'stuff' of morphology proper. For the most part, these will not concern us here. I will assume that the vocabulary insertion operation occurs late in the morphological component.

The first issue I address here is a technical problem for vocabulary insertion. We saw in chapters 2 and 3 that Georgian second-cycle agreement affects the choice of vocabulary item with which the AGR relation is spelled out. In section 4.1 I consider the implications of this for our understanding of the vocabulary insertion operation.

In section 4.2 I address the issue of exceptionality in morphology, a pervasive feature of most morphological systems. I present the Karok (Hokan) verbal agreement system, which patterns as a low- language up to a point. The point is a simple one: the identification of exceptionality is intimately tied to expectations of what would be 'normal'. We cannot ask questions about exceptionality (e.g. are there limits to what is and is not vulnerable to idiosyncracy; are some morphosyntactic environments more susceptible to exceptionality than others?) in the absence of a base-line hypothesis about the core properties of a system.

This point is made more persuasively in section 4.3 where I argue that certain apparent idiosyncracies in Erza Mordvinian can actually shown be shown to pattern systematically if the underlying morphosyntactic structures are identified according to the proposals in this thesis.

4.1 Spelling out the *m/v* alternation

We saw that second-cycle π-agreement in Georgian correlates with a morphological split between *m*-set and *v*-set person agreement markers. Crucially, this means that the vocabulary insertion operation must be sensitive to the contrast between 1st and 2nd cycle insertion sites. This is unexpected if the morphological features indexed with the respective morphological strings are simply those shown below:

1)

	m-set			*v-set*	
m		1	v		1
g		2	ø-/x-		2
			ø		3

Likewise, there is no obvious secondary feature that could be appealed to ensure that only m-set markers are inserted in first-cycle AGR sites, and v-set in second-cycle ones.

The absence of a strict correlation between Case or grammatical function and agreement features in Georgian makes this untenable. For example, the m-set markers cross-reference SU, DO and IO arguments (seen in terms of grammatical function), and nominative, accusative and dative arguments (seen in terms of Case). The *v*-set markers in cross-reference SU and DO arguments (second-cycle agreement with IOs being simply impossible given the φ-specifications and distribution of Georgian), and nominative, ergative and accusative cases.[1] Clearly they are not in complementary distribution with respect to Case and grammatical function. Yet, differentiation must be possible, as insertion of one set in lieue of the other is ungrammatical.

2) a. v/*m-xedav
1-see
I see him/them

b. v/*gv-xedav-t
1-see-PL
we see him/them

c. ø/*g-xedav
2-see
you see him/them

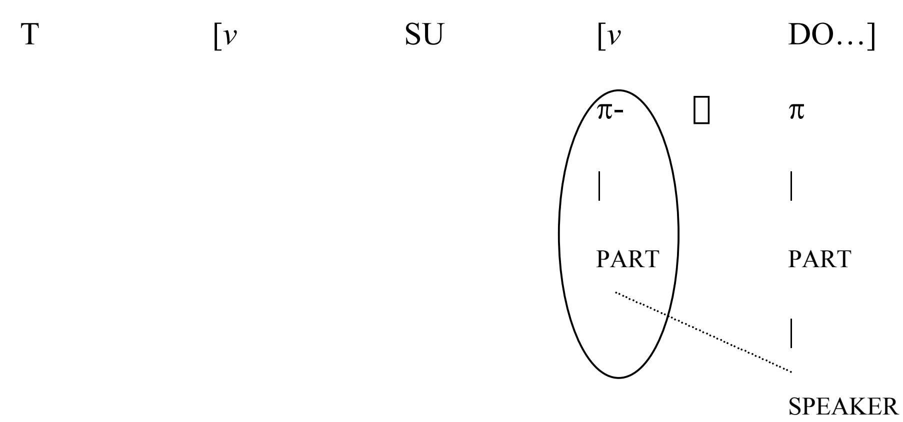
d. ø/*g-xedav-t
2-see-PL
you(pl) see him/them

To see why *m-, gv-,* and *g-* are blocked by *v-* and *ø-* , we must compare the feature specifications for these vocabulary items as well as the specification of their respective

[1] See analysis of Georgian in section 3.6 of the previous chapter.

insertion sites. Given the theory of Vocabulary Insertion (Halle and Marantz 1993, Harley & Noyer 1999, Embick 1997, and others), there are two possibilities at the outset: either the blocked set is simply incompatible with the insertion site (meaning features of the blocked vocabulary items are a superset of the features at the insertion site), or both sets are equally compatible and there is vocabulary competition, but the vocabulary items of the blocked set are less highly specified than their competitors, and so lose the competition. Nothing in my analysis hinges on which of these solutions is correct; what is important to me is that one of these (or alternatively, some counterproposal) must be plausible. Let us consider the first hypothesis: that the vocabulary items in the blocked set are not compatible with the insertion site. To begin, consider the (circled) insertion sites for the first cycle person affixes shown in (3) and (4). Note that there is no insertion site that corresponds to a 3rd person 1st-cycle affix because in the environment where this might arise, a second-cycle is always triggered.

3) *insertion site for* m- *(controlled by 1st person goal)*



4) *insertion site for* g- *(controlled 2nd person goal)*

T	[*v*	SU	[*v*		DO…]
			π-		π
			|		|
			PART		PART

The insertion sites in (3) and (4) are contrastive in that the former has extra structure built in the value operation, and the latter does not. We could, as a first approximation, say that these 1st-cycle VIs are inserted by the following vocabulary insertion rules:

5) *m*-set

m [π[PART[SPEAKER]]]

g [π[PART]]

However this does not resolve our problem, it just restates it in terms of the actual FFs at the insertion site. Thus, the *v*-set markers for 1st and 2nd person controllers could be argued to be inserted by identical rules.

Let us now take a look at the second-cycle insertion sites. Recall that these occur just when a previous cycle of Agree has failed (which in Georgian is whenever first accessible NP of the original π-AGR domain fails to match or value). The relevant contexts of the (circled) insertion sites for the second-cycle affixes are shown in (6) and (7):

6) *insertion site for* ø- *(controlled by 2nd or 3rd person SU)*

T	[*v*		SU	[*v*		DO…]
	π-	→	π	π-	→	π
			|	|		
			(PART)	~~PART~~		

7) *insertion site for* v- *(controlled 1st person SU)*

T	[*v*		SU	[*v*		DO…]
	π-	→	π	π-	→	π
			|	|		
			PART	~~PART~~		
			|			
			SPEAKER			

If we allow for the feature specification of agreement affixes to distinguish between the feature of the probe (i.e. the feature specification that establishes conditions on match) and feature values (i.e feature structure built as a result of successful value), then we have a clear mechanism for blocking the unattested insertion of the second-cycle affixes in the first-cycle insertion sites of (3) and (4) as well as the unattested insertion of the first-cycle affixes in the second-cycle insertion sites of (6) and (7). In the following list of vocabulary insertion rules I distinguish the feature structure of the probe from the feature structure built by value by italicizing and underlining the latter. The distinction has no status for interpretable features, since these do not trigger Agree.

8) first-cycle affixes (m-set) — second-cycle affixes (v-set)

gv- [π[PART*/SPEAKER.GROUP/*]]

m-	[π[PART/*SPEAKER*/]]	v-	[π/*PART[SPEAKER*/]
g-	[π[PART]]	Ø-	[π] or elsewhere

The vocabulary items in (8) are fully contrastive; m-set rules insert vocabulary in [π[PART]] AGR categories, and *v-set* rules insert vocabulary in [π] AGR categories.

Consider what will happen if we attempt to insert these items in any slot but the correct one. There are two questions to consider: (1) will the VI match any insertion site other than the attested one; (2) what will the result of vocabulary competition be for any insertion site other than the attested one? I assume the standard criterion for matching: the features of the VI must be a subset of the features of the LI insertion site. To this standard assumption, I add the following criterion:

9) Matching (in the sense relevant to vocabulary insertion) is evaluated for inherently specified feature structure and value-built feature structure independently.

Consider just the evaluation of inherent feature specifications. No m-set marker will match a v-set environment, as the inherent feature specification of any m-set vocabulary item is a superset of any v-set insertion site. However, in principle, *v-set* markers *are* eligible for insertion in m-set environments, because their inherent feature specifications have the required subset of features. Thus, incorrect m-set insertion is eliminated by conditions on matching (nb: not to be confused with the conditions on match proposed for Agree!), but incorrect v-set insertion must be eliminated by vocabulary competition. Consider the competition for insertion in (3). The matching items are listed below. The winner is correctly predicted to be *m-* because it is the VI with the best fit.

10) m- [π[PART*[SPEAKER]*]]

v- [π*[PART[SPEAKER]*]

g- [π[PART]]

ø- [π] or elsewhere

Note that given (9), *v-* in (10) is in fact eliminated independently because its inherent specification is a superset of the value specification of the insertion site. Thus competition is between *m- , g- and ø-,* the first two being m-set forms and the latter a *v*set form (*m-* being the best fit)

An alternative mechanism for ensuring only the attested vocabulary insertion pattern would be to posit a conditioning context that could only be met by one set, as in:

11)

first-cycle affixes		second-cycle affixes	
gv-	[π[PART[SPEAKER.GROUP]]]		
m-	[π[PART[SPEAKER]]]	v-	[π[PART[SPEAKER]]/ T __
g-	[π[PART]]	ø-	[π]/ T__

If the conditioning context for such suppletive allomorphy is strictly local, as has been proposed throughout the literature (Lieber 1980, Sproat 1985, Bobaljik 1995, Adger, Bejar and Harbour 2001), then it is only in the 2nd-cycle that the insertion context would count as being in the environment of T. Second-cycle affixes would block the 1st-cycle affixes in the local environment. Second-cycle affixes would be incompatible with the insertion point in the 1st-cycle insertion context.[2]

[2] Additional assumptions would be necessary here: we would have to assume that conditioning context 'out-specifies' the [group] feature, so we can't get gv- in 2nd cycle agreement with a 1.pl SU because

To reiterate: All else being equal, the details of the solution to this issue are not relevant to my analysis. What matters is that some plausible mechanism be available

4.2 □-Distribution and morphological irregularity

I now turn to the problem of irregularity in inflectional morphology, in particular suppletive allomorphy (where by 'suppletive' I mean not phonologically conditioned). Irregularity is a persistent property of morphological systems, and both a source of complexity and cross-linguistic variation in its own right. In chapter 1 it was noted that this dimension of complexity would not fall within the scope of the thesis. Indeed, it is not the purpose of this section to examine the properties of suppletion per se. The goal is to show that the proposals that have been developed thus far cast new light on the problem of suppletion. This is necessarily true for the very basic reason that exceptionality is only identifiable with respect to some set of expectations as to what would be 'regular'. Because the characterization of a language as split-□, low-□, etc. makes clear and specific predictions with respect to the expected shape of the agreement system (see Appendices 1-4), a new perspective is introduced with respect to what is to count as regular or irregular. This point is fairly banal, but important nonetheless, given that one of the subtlest challenges to morphological analysis is differentiating between morphological irregularity versus a morphological pattern that is complex but systematic.

---

the T_ environment counts as more specific than the extra feature [GROUP] in the competitor, and we can't get g- in 2nd cycle agreement with a 2 SU because T_ counts as more specific than [PART]. Also note that ø- cannot be an elsewhere form under this view.

4.3 Karok

The Karok (Hokan) indicative agreement paradigm is presented below (Bright 1957).

12) singular DOs plural DOs

	1	2	3	1pl	2pl	3pl
1.SU		nu-	ni-		ki-___-ap	ni-
	—	2	1	—	2.PL	1
2.SU	na-		ʔi-	kin-		ʔi-
	1	—	2	1.PL	—	2
3.SU	na-	ʔi-___-ap	ʔu-	kin-	ki-___-ap	ʔu-
	1	*2.PL*	3	1.PL	2.PL	3
1.pl.SU		nu-	nu-		ki-___-ap	nu-
	—	2	1.pl	—	2.PL	1.pl
2.pl..SU	ka-na-		ku-	kin-		ku-
	PL-1	—	2.pl	1.PL	—	2.pl
3.pl..SU	ka-na	ʔi-___-ap	kun	kin-	ki-___-ap	kun-
	PL-1	*2.PL*	3.pl	1.PL	2.PL	3.pl

If we put aside the highlighted column then, like Nishnaabemwin, Karok is a low-□ language. The low-□ pattern emerges as follows (see Appendices 1-3). A participant DO always controls π-agreement, and a plural DO always controls #-agreement. There can be π-agreement with the SU only if the DO is 3rd person. Thus, the rightmost column of each quadrant shows π-AGR controlled by the SU. There can be #-agreement with the SU only when the DO is singular. Thus, the black-type forms in the top right and bottom right quadrants are identical, because the plurality of SU never enters into the agreement. The 3rd person DO forms are in grey-type here because, just as in Georgian, 3rd person objects are for some reason unable to control plural number agreement. This is why the plurality of the SU *is* contrastive in just these cases.

The set of exponents for person are given below.

13)

	First cycle			*Second cycle*	
	A_1	B_1		A_2	B_2
1	na-,	kin-	1	ni-,	nu-
2	*nu-,ʔi-*	*ki-*	2	ʔi-,	ki-
3	N/A		3	ʔu-,	kun-
pl	-ap		pl	ka-	

Setting aside the plural morphemes *–ap, ka-*, the generalization is that in each cycle there is a singular and a plural variant for each person (NB: there is, by definition, no AGR relation with 3rd person on the 1st cycle, because it fails to value). Series A_1 affixes (first-cycle) occur when the DO is singular, and series B_1 when the DO is plural. Series A_2 (second-cycle) occurs when the SU is singular, and B_2 when the SU is plural. In fact, we can generalize across both cycles and say that Series A occurs when the π-controller is singular, and series B occurs when the π-controller is plural. Taking seriously the second merger of a head under projection, the fused person/number form for Series B affixes is predicted by the AGR configurations of a low-□ analysis. It is precisely in these configurations that valued π and # enter the morphological component on the same head. Consider first the B_1 affixes, which occur when the DO is a plural participant. Because π-

AGR and #-AGR are both low, DO will match and value both probes; features of the SU are irrelevant to this result. When *v* projects, it will be without AGR features, thus the output of syntax is an insertion site with both valued π and #; a natural candidate for a vocabulary insertion rule that makes references to both sets of features.

14)

	≠ no Match			*□ Match, no Value*			*□ Match, Value*
	…	SU	…	*v* $\pi_{[\text{PART}]}$- #-	… □ □	DO $\pi_{[\text{PART}]}$ #	DO controls π-AGR and #-AGR

Now consider the B_2 series, which occur when both π-AGR and #-AGR are controlled by the SU. In order for this configuration to arise in a low-□ language, it must be the case that the DO is neither plural, nor a participant. (In Karok, we can generalize this to all non-participant DOs, since 3^{rd} person plurals independently cannot control #-agreement). Here again, the output of syntax is an AGR head with both valued π and # (this time it is the projected occurrence of *v* instead of the original occurrence of *v*). This too is a natural candidate for insertion of a fused person/number vocabulary item.

15)

	≠ no Match			*□ Match, no Value*			*□ Match, Value*
	…	SU $\pi_{([\text{PART}])}$ #	…	*v* π[~~PART~~]- #-	… □ ≠	DO π	DO fails to value π-AGR and #-AGR. *v⁰ projects with both AGR features*
v π- #-	… □ □	SU $\pi_{([\text{PART}])}$ #	…	*v* π[~~PART~~]- #-	… □ ≠	DO π	SU values π-AGR and #-AGR

Notably, these are the only configurations in which a single AGR head will host both valued π and # features. In all others, one feature will project and the other will not. Abstracting away from the morphological operations that we know have the potential to disestablish isomorphy in the mapping between syntax and morphology, the most simplistic (and likely to be wrong) prediction is that in all other configurations, person and number will be realized by distinct vocabulary items. Amazingly, this is true outside of the exceptional 2nd person DO forms. Thus in 2>1 and 3>1 forms (bottom left quadrant) the plural and person markers are distinct.

The simplistic prediction fails for the 2.pl DO forms, which we predict should have fused person/number morphology. However here we are entering into the realm of exceptionality in the paradigm. Note first of all that in fact the singular and plural 2nd person prefixes in (13) *are* contrastive, so there is no grounds for saying that second/plural is not a fused prefix. The pattern only appears to be disrupted because the 2nd person DO morphology is more generally idiosyncratic. 2.sg DOs are spuriously spelled out by a 2nd cycle (polite because agentive?) prefix, and there is spurious plural morphology in the 3.sg>2.sg cell (polite tu/vous?). The addition of the *–ap* plural suffix in the 2.pl DO forms likewise fits in with the extra (politeness?) marking of 2nd person DOs.

What a close examination of Karok has shown is that, while we cannot necessarily account for irregularity, there are certain subtle regularities which correlate with structural properties of the output of the computation (fused versus non-fused forms). In the next section, however, we will see that likewise, the structural output of the computation *can* in fact be useful in the explanation of the distribution of irregularities.

4.4 Suppletion

A less obvious consequence of the proposed theory of Agree is that precise predictions are made with respect to the actual distribution of suppletion. This claim is contingent upon the hypothesis that the distribution of suppletion is narrowly constrained by syntactic structure in a manner to be made precise immediately. I follow the analysis of suppletion presented in Adger, Bejar and Harbour 2001 (henceforth ABH).

ABH define suppletion simply as the choice of a special morphological form in a specific context. They assume a realizational model of morphology along the lines of Halle and Marantz 1993, within which a natural hypothesis is that this choice is implemented by rules of vocabulary insertion which insert phonological content into morphosyntactic structure (the output of the computation and a morphololgical component). The general environment for suppletion is taken to be a complex syntactic word, i.e. a head-adjunction structure.[3] The insertion site itself is taken to be an X^0 element. The general structure of a suppletive rule is thus:

16) Vocabulary Insertion

/□/ □ F (| __ G) for F, G, morphosyntactic feature bundles

□, a phoneme string

The core claim made in ABH is that the environment defined by '|' in (16) is strictly local and can only mean *structurally adjacent to*.

17) Structural Adjacency

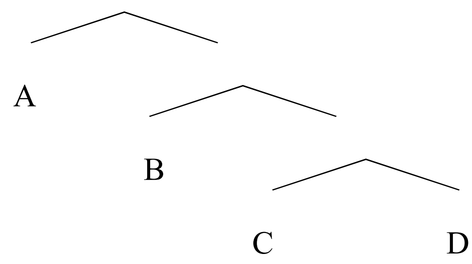
In a complex head, {C, {A, B}}, C is *structurally adjacent* to the terms A and B immediately contained by its sister (cf. Bobaljik 2000).

Thus, in (18) the terminal element B can be a context for insertion of a suppletive form targeting A, C or D. But A can only trigger suppletion on B.

---

[3]Many questions arise regarding the validity/ontology/distribution of such structures. I put these aside here.

18)



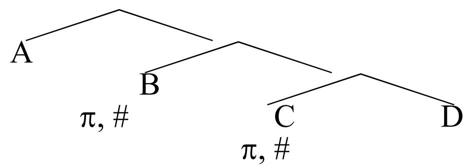
By hypothesis, allomorphy is blocked by intervening morphosyntactic structure. This is illustrated by the following verb forms from Kiowa. (a) and (b) show that negation triggers root suppletion. Kiowa also has a distributive morpheme (c); when it co-occurs with negation it intervenes between NEG and the root, thus disestablishing adjacency and prohibiting suppletion (d) and (e).

19)Kiowa (Harrington 1928, Watkins 1984, in ABH 2001)

	AGR	root	DISTR	NEG	
a.	á-	k'ú·			'they sit'
b.	á-	<u>kóp</u>		-gɔ̂·	'they don't sit'
c.	á-	kú·	-yɔ́-		'they sit about'
d.	á-	<u>kú·</u>	-yɔ́-	-gɔ̂·	'they don't sit about'
e. *	á-	<u>kóp</u>	-yɔ́-	-gɔ̂·	'they don't sit about'

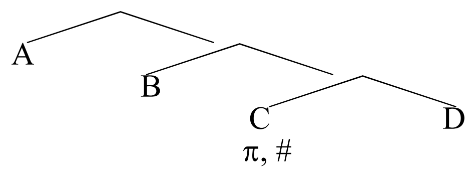
Crucially, permutations on the specification and distribution of □ directly affect the distribution of contexts in which suppletive relations might be established or disestablished. In (20) a suppletive relation is predicted to be possible between A and the higher □-set. In (21) A is not structurally adjacent to an AGR-head, so no suppletive relation between A and φ should be possible. In (22) a suppletive relation should be possible between A and #AGR, but not with π-AGR.

20)



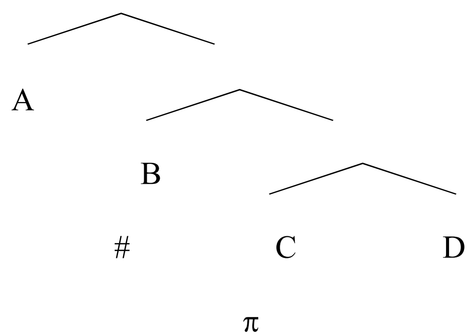
Double-φ

21)



Low-φ

22)



Split-φ

But this is only part of the picture. We further predict that suppletive relations will be circumscribed by the set of TCPs established in a particular derivation and in particular by second-cycle effects. The latter dramatically affect locality in a complex word by altering the distribution of φ. For example, the complex verb word in a low-φ language may have any of the φ-configurations in (23) depending upon whether there have been second-cycle effects. (Note that I take seriously here the proposal to treat the label of a projection as a head. The second *v* and T in (23) are just that. Since T is not an AGR-head, its label does not affect the system.)

23) a. [T [T [*v* [$v_{\pi,\#}$ V]]]] *no second-cycle of Agree*

b. [T [T [$v_{\#}$ [v_{π} V]]]] *second-cycle for #*

c. [T [T [v_{π} [$v_{\#}$ V]]]] *second-cycle for π*

d. [T [T [$v_{\pi,\#}$ [*v* V]]]] *second-cycle for π and #*

Thus the range of suppletive relations into which T^0 might enter with □ is dependent entirely on the □-syntax of Agree. This in turn is dependent upon the □-specification of arguments in the AGR-domain. The predictions are clear: there will be no T-□ suppletion if the DO controls both π and # (a); there might be T-# suppletion if the DO controls π and the SU controls # (b), and so on. Of course, the matter of which controller is seen to be triggering/targeted by a suppletive relation is dependent upon □-syntax as well.

4.5 Erza Mordvinian

The nonpast transitive paradigm of Erza Mordvinian (Uralic) is shown below:[4] (Abondolo 1982)

24)

			object					
			singular			plural		
			1	2	3	1	2	3
SU	sg	1		-d-an TNS-2	-s-a TNS-1		-d-ad-yz TNS-2-PL	-s-y-n TNS-PL-1
		2	-s-am-ak -TNS-1-2		-s-ak TNS-2	-s-am-iz TNS-1-PL		-s-y-t TNS-PL-2
		3	-s-am-am -TNS-1-1	-d-anz-at TNS-3?-2	-s-y TNS-DFL	-s-am-iz TNS-1-PL	-d-ad-yz TNS-2-PL	-s-y-nze TNS-PL-3
	pl	1		-d-ad-yz TNS-2-PL	-s-y-nek TNS-PL-1		-d-ad-y-z TNS-2-PL	-s-y-nek TNS-PL-1
		2	-s-am-iz TNS-1-PL		-s-y-˜k TNS-PL-2	-s-am-iz TNS-1-PL		-s-y-˜k TNS-PL-2
		3	-s-am-iz TNS-1-PL	-d-ad-yz TNS-2-PL	-s-yz-ø TNS-PL-3	-s-am-iz TNS-1-PL	-d-ad-yz TNS-2-PL	-s-yz-ø TNS-PL-3

Mordvinian appears to be split-□ language like Georgian. This is most apparent if the shaded block of forms is put aside temporarily. Outside of these cells the agreement

[4] NB: I have parsed the plural differently from Abondolo (1982) who separates – i-z/-y-z. I do not see what the basis for this is, and a simpler set of glosses is possible under my alternative. Even if I am wrong on this count, it should not really affect the substance of my claims here.

pattern is very regular: π-AGR is controlled by the DO unless it is 3rd person, in which case it is controlled by the SU (DO-controlled person markers are highlighted). #-AGR is controlled either by a plural SU, or if the SU is singular then by a plural DO. The fact that the SU is the preferred controller (and therefore the highest accessible NP in the domain of #-probe) can be seen by the fact that the morphology is sensitive to the contrast between singular and plural DOs just in case the SU is singular. Thus, the top two quadrants in (24) differ from one another, but the bottom two are identical to one another.

The set of exponents for person and number are given below, with a distinction made between 1st and 2nd cycle vocabulary person markers. Presumably, since #-AGR is high, second-cycle effects for number only occur when both SU and DO are singular; so the sole consequence of second-cycle of #-AGR is default #-agreement (ø in Mordvinian).

25)		*First cycle*	*Second cycle*		
	1	-am	-a	-n	-nek
	2	-ad	-ak	-t	-˜k
	3	N/A	-y	-nze	-ø
	pl	-yz/-iz/-y		-ø	

The alternations in the form of the plural are phonological: y~i after nasals and it would seem that z>ø before an obstruent. The second-cycle alternations, on the other hand, are suppletive. There are three series of second-cycle person affixes, and the conditioning context that describes their distribution can be stated in terms of number. The first series {-a, -ak, -y} occurs when both arguments are singular; the second series {-n, -t, -nze} is triggered by a plural DO; and the third series {-nek, -˜k, ø} is triggered by a plural SU. This is precisely what a split-□ analysis predicts should be possible. First of all, the fact that sensitivity to number is only possible for 2nd-cycle affixes and not for 1st-cycle affixes follows from the □-configuration, which for a split-□ language as follows (again, I abstract away from the second-cycle for #, presumably there could be one if all arguments were singular):

26)	a.	[T	[$T_{\#}$	[*v*	[v_{π}	V]]]	*no second-cycle*
	b.	[$T_{\#}$	[T	[*v*	[v_{π}	V]]]	*second-cycle for #-Agree*
	c.	[T	[$T_{\#}$	[v_{π}	[*v*	V]]]	*second-cycle for π-Agree*
	d.	[$T_{\#}$	[T	[v_{π}	[*v*	V]]]	*second-cycle for π and #-Agree*

Configurations (a) and (b) are not relevant to us at present because the alternations we are looking at only occur when there is a second-cycle for π. The configurations in (d) arises when both arguments are singular (the shaded, irregular quadrant of (24)). (d) is the configuration in which the first series {-a, -ak, -y} is inserted. Since the trigger $T_{\#}$ here is not structurally adjacent to the target v_{π}, I assume that the suffixes of this series are basic, and are not actually subject to any restriction on their context (except for the restriction that prohibits them from being inserted in 1st cycle environments, see previous discussion of this for Georgian in section 4.1). This leaves configuration (c), which of the four is the only one that we would predict (independently of the actual Mordvinian data) to license a suppletive relation between π and #. This is precisely the configuration that gives rise to the second and third series of suffixes. The second series occurs when the controller of $T_{\#}$

is the DO (top-right quadrant of (24)), the third series occurs when the controller of is the SU (bottom quadrants of (24)).[5]

A second suppletive relation exists in the tense morphology of (24) . The usual tense marker is the suffix *–s*, however when the DO is 2nd person a special form *–d* occurs. A second-person DO always controls π-AGR on the first cycle, so the □ configurations in queston are:

27) a. [T [$T_{\#}$ [*v* [v_{π} V]]] *no second-cycle*

b. [$T_{\#}$ [T [*v* [v_{π} V]]] *second-cycle for #-Agree*

T is not structurally adjacent to π-AGR in either of these configurations, however, this is not necessarily problematic. ABH show that structural adjacency can be established in the morphological component by deletion of an intervening head. In (27) the deletion of the intervening head seems plausible, since it is a projection of *v* with no features for vocabulary insertion. Let us suppose that this deletion is a reflex of vocabulary insertion.:[6] Adopting the thesis that vocabulary insertion begins at the root of a complex syntactic word, and proceeds upwards (Halle and Marantz 1993), the derivation (as of v_{π}) would be:

[5] The fact that the context for the suppletion rule differentiates between the two controllers is formally (though not functionally) unexpected. If Agree only copies □-features and not Case or □ or any other such identifier of grammatical function, then the simplest assumption is that the suppletive rule triggered by $T_{\#}$ should not differentiate between the grammatical function of the #-AGR controller. In other words, there should be just one special form in the context of #. There are several possible explanations for this kind of phenomenon, which is probably relatively widespread. One possibility is that the trigger is also sensitive to some other property of the context. Another possibility is that singular NPs in Mordvininan are not in fact totally underspecified, they are interveners and as such trigger 2nd-cycle effects. The DO becomes an accessible controller when a 3rd person SU intervener displaces. Note that this 2nd-cycle would have to occur prior to Projection, otherwise the SU would again intervene.

[6] We are in fact obliged to assume this. If such deletion could occur prior to vocabulary insertion, then adjacency would be established between $T_{\#}$ and v_{π} in (26), and we would lose the generalization that suppletion is only predicted when there is a second-cycle of π-agreement.

28) a. T v $\boxed{v_\pi}$ V *VI :* *2 □ -da*

b. T $\boxed{v}$ v_π V *VI :* *v deletes if vacuous*

c. $\boxed{\mathrm{T}}$ v_π V *VI :* *T □ -d | __2nd person*

(1st cycle)[7]

Note that we make another clear prediction here, which is that the establishment of adjacency by deletion can only license an inwardly sensitive suppletive relation (where the target is higher than the trigger). In the first cases we looked at, where higher trigger (#) caused suppletion of a lower target (π), this mechanism is not available. Deletion of the vacuous *v* head would only occur after VI on the π-AGR had already taken place.

Finally, what can be said about the irregularities in shaded forms of (24)? Here, I would argue, the syntax of agreement has nothing to say about the actual form of these irregularities, however it might reveal something about their distribution The □-configuration for these forms is the following:

29) [$T_{(\#)}$ [T [v [v_π V]]] *second-cycle for #-Agree*

Note that # on T here is default (both SU and DO are singular) and there is no evidence that default #-agreement receives any spell-out in Mordvinian. This means there is no agreement morphology corresponding to the first three X^0s on the left-edge of the syntactic word. It could simply be that a morphological well-formedness condition in the

[7] Recall the 2nd cycle 2nd persons do not trigger suppletion of T. This is not problematic for us, since the feature structure of 1st and 2nd cycle controllers is distinct, and we have assumed on for independent reasons that the vocabulary insertion operation is sensitive to this difference.

language disfavours the realization of T^0 *sans* agreement morphology (or alternatively, the string of vacuous heads in (29) may simply be too long or too prominent).[8] This would motivate the spurious 'extra' agreement markers that appear in all of the forms except 1SU>2DO. The latter may well be a syncretic 1/2 form that arises for independent reasons. I will have more to say on such forms in chapter 5.

[8] cf. Noyer's (1992) Q slots.

Chapter 5 Consequences, Questions and Conclusions

5.0 Introduction

This chapter is the most speculative of the thesis. Here I raise questions that explore the consequences of the core proposals for Case theory. The discussion of Case is in two parts. The first part (section 5.1) walks through certain predictions made by my core proposals in tandem with minimalist Case theory. The predictions are typologically problematic and raise questions about the nature of the relation between Case and Agree. The second part (section 5.2) explores a puzzle posed by defective interveners, which are unable to value because they have been assigned Case, although they are not necessarily in violation of the entailment condition on value. The chapter ends with a brief section (5.3) of concluding remarks about the thesis as a whole.

5.1 Agreement and Case

Recall from Chapter 1 that Chomsky 2000, 2001 takes Case assignment to be a reflex of Agree, along the lines of Schütze 1997 (also George and Kornfilt 1981) who argued for the unification of Case and agreement as a single operation (his *Accord*).[1] More specifically, Case is taken to be a reflex of value; the unification holds over controllers, not interveners. NPs that match without valuing fail to be Case-marked by the AGR target; conversely, NPs with Case may match but may not value (hence defective intervention effects). The Icelandic paradigm in (1)-(4) from Schütze (1997:102 (1)-(4))

[1] Chomsky's version of this thesis is slightly weaker than Schütze's. Chomsky allows agreement without Case assignment (and vice versa) if either the target or the goal is φ-defective.

illustrates this correlation between agreement and nominative Case assignment. (1) and (2) show that whichever argument gets nominative case also controls agreement, regardless of its grammatical function. (3) and (4) show that agreement is not obligatory in finite clauses, so if no nominative case is assigned, default agreement is acceptable. And (2) and (4) show that non-nominative NPs cannot control agreement.

1)	a.		Strákarnir	voru	ekki	svangir.
			the-boys(NOM)	were(3.PL)	not	hungry(NOM.PL.MASC)
	b.	*	Strákarnir	voru	ekki	svangt.
			the-boys(NOM)	were(3.PL)	not	hungry(SG.NEUT)

2)	a.	Henni	leiddust/?*leiddist	þeir.	
		she(DAT)	was.bored.by.(3.PL/?*3.SG)	they(NOM)	Taraldsen 1995:307

3)	a.		Strákunum	var	ekki	kalt.
			the.boys(DAT)	was(3.SG)	not	cold(SG.NEUT)
	b.	*	Strákunum	voru	ekki	köldum.
			the.boys(DAT)	were(3.PL)	not	cold(DAT.PL.MASC)

Sigurðsson 1993:38

4)	a.	Okkur	vantaði	vinnu.	
		us(ACC)	lacked(3.SG)	a.job(ACC)	
	b.	Mig	vantar	skó	
		me(ACC)	lacks(3.SG)	shoes(ACC)	Einarsson 1945:105

c. Ferðunum seinkaði

the.journey(DAT) was.delayed(3.SG) Sigurðsson 1993:4

This view of Case assignment, in conjunction with the theory of agreement proposed in this thesis, makes a clear set of predictions. Putting aside CSA effects momentarily, I predict that a low-□ language should have an ergative case system (5) (see also Levin and Massam 1985, Bobaljik 1993), a double-□ system should be accusative (6), and split-□ should be mixed, with the high AGR patterning in an accusative pattern and the low AGR patterning in an ergative pattern. A low-□ language should have an ergative case type in the basic sense that the object of a transitive should be licensed by the same AGR-head as the subject of an intransitive, as in (5). Likewise, a double-□ language should be accusative in the basic sense that the subject of a transitive should be licensed by the same AGR-head as the subject of an intransitive, as in (6).

5) low-□ = Ergative Pattern: O and S assigned Case by same AGR head

	key:	≠ *no Match*			□ *Match, no Value*		□ *Match, Value*
a. transitive:							
	A	…	*v*	…	O		Agree/Case with O
			□	□	□		
b. intransitive (unaccusative):							
	…	*v*	…	S			Agree/Case with S
		□	□	□			
c. intransitive (unergative):							
v	…	S	…	*v*	…	ø	Agree/Case with S
□	□	□		□	≠	ø	(second-cycle)[2]

[2] So-called split-S ergative systems might also be derivable along these lines. In such systems , unergative intransitives (agent-like subjects) take the same case marking as agents of transitives, while unaccusative intransitives (patient-like subjects) take the same markings the object of transitives. This pattern would follow from the system if it could be shown that there are morphological second-cycle effects in these languages, so that unaccusative and unergative S are marked differently because the former reflects a first cycle of Agree and the latter a second cycle. The plausibility of such an approach is unclear. There is considerable disagreement in the literature as to whether split-S subjects can be characterized structurally at all (see, for example, the discussion in Baker 1996: 5.4.1)

6) Double-φ = Accusative Pattern: A and S assigned Case by same AGR head

key:	≠ *no Match*			□ *Match, no Value*		□ *Match, Value*	
a. transitive:							
T	…	A	…	*v*	…	O	Agree/Case with A
□	□	□		□	□	□	
b. intransitive (unaccusative):[3]							
T	…	ø	…	v_{def}	…	S	Agree/Case with S
□			□			□	
c. intransitive (unergative):							
v	…	S	…	v_{def}	…	ø	Agree/Case with S
□	□	□					

It may well prove fruitful to think of varieties of Case patterns in these terms (the approach is reminiscent of Levin and Massam 1985, also Bobaljik 1993, Chomsky 1993). The problem is that Schütze's/Chomsky's unification of Case and agreement predicts that the Case patterns predicted above should be isomorphic with analogous agreement patterns. So, I further predict (a) that interveners of the class introduced in this thesis (entailment condition failures) should not receive Case because they are unable to value; (b) that domains for Case assignment should be dynamic in the way AGR domains have

[3] Here and in (c) the defectiveness of *v* implements Burzio's generalization (Chomsky 2000, 2001); an intransitive *v* cannot assign Case (for whatever reason), so an NP in the AGR-domain of *v* cannot be rendered inactive by an AGR-relation in this domain. This enables the NP to enter into an AGR-relation with a higher AGR-position. There is an important asymmetry between low-φ and double-φ languages in this respect. Intransitive *v* must be defective in the double-φ scenario, but not in the low-φ scenario. Intuitively, it would seem that low-φ *v* simply isn't 'expendable' since it is the only AGR head in the clause. But even assuming this characterization of the asymmetry is on the right track, what it tells us about the conditions of uninterpretable φ-distribution is unclear to me.

been proposed to be (Rezac 2003); and (c) that, in general, languages with context-sensitive agreement should also display context-sensitivity in Case assignment.[4] By analogy with agreement, we would thus expect it to be typologically common to find such things as Caseless 3.sg NPs in object (1st cycle) positions, as well as transitive subjects that are Case marked only if their object cannot be (these would be languages where, in Dixon's (1994) terms, A and O are aligned to the exclusion of S when the DO is 3rd person. I know of no such cases.

In many languages, only one of Case or agreement is morphologically marked and this prediction cannot be tested. However, there exist many other languages where Case marking and agreement coexist and it is clear that counter-examples to the predictions abound, making the hypothesis that Case and agreement are reflexes of a unique syntactic relation difficult to maintain. Schütze demonstrated persuasively that his theory made the correct predictions for a fair range of languages with both morphological Case and agreement, however his sample was broad only in the sense that a variety of language families were analyzed. In terms of the typology developed in Chapter 3 of this thesis, his sample was in fact quite narrow, as all of the languages were of the non-CSA variety, which I showed to be a narrow subset of the typology. When a broader (in my sense) range of languages is considered, the necessary conclusion seems to be that the

---

[4] The existence of □-sensitive Case splits is well known, but they are not the ones predicted here (see Dixon 1994, Silverstein 1985). So, for example, one reported empirical generalization is that 1/2 pronouns are less likely to be ergative than 3 pronouns. For my system, this is tantamount to saying that in clauses with 1/2 subjects, □ is less likely to be low than in languages with 3 subjects. This is not a prediction made by my theory (though it would not be incompatible with my theory, given auxiliary assumptions). What my proposals actually predict is that clauses with 1/2 DOs are more likely to show ergative pattern than those with 3 DOs. The extensions of these two predictions merely overlap (for e.g. both predict that a 3>1/2 clause is ergative). They differ in that for me a 1/2>1/2 clause should be as ergative as a 3>1/2 clause.

relation between Case assignment and Agree is not straightforward.[5] The anticipated isomorphy between CSA effects in agreement and Case patterns simply does not arise.

Schütze himself notes that split-Case systems pose serious problems for the unification hypothesis. This is dramatically illustrated by Georgian which has a 3-way Split Case system: NOM-ACC, ERG-ABS, DAT-NOM and in which agreement and Case do not covary at all (except in the latter DAT-NOM case, which I showed in Chapter 3 to result from an argument structure permutation in that the base position of dative subjects is different from that of non-dative subjects). There are two pivots for Case splits in Georgian: (a) the tense/aspect/mood (TAM) system, and (b) verb class. I will focus here just on the former, but I assume that the proposals to be discussed in the next section are generalizable to verb class pivots as well). The Accusative/Ergative split is illustrated below. (7) and (8) show that nominative arguments cannot control agreement unless the □-configuration warrants it, in which case they must do so (target-controller pairs are highlighted by underlining; where emphasis is on the failure of an argument to agree, that argument is set in a box). In (7) the object is assigned nominative Case by T^0 even though it cannot control number agreement (the AGR-relation proper to T^0). It cannot be the case that valuing is optional, because in (8) the nominative argument *must* value; hence the ungrammaticality of (8) where agreement is dropped. (7) and (9) show that non-nominative arguments can control agreement (and must do so if the □ configuration warrants it).

---

[5] Note that if Case and Agree both covary with the distribution of uinterpretable features, then it should be expected that in some cases they would line up in the way described by Schütze.(cf. Marantz 1991).

7) a. (me) cign-i da-v-c'er-e

I(ERG) book-NOM PREV-1-write-INDIC.AOR

'I wrote the/a book.' cf. Hewitt 1995:117

b. (me) cign-eb-i da-v-c'er-e

I(ERG) book-PL-NOM PREV-1-write-INDIC.AOR

'I wrote the books.'

c .* (me) cign-eb-i da-v-c'er-en/t

I(ERG) book-PL-NOM PREV-1-write-INDIC.AOR.PL/PL

'I wrote the books.'

8) a. (me) tseril-s da-v-c'er-e

I(NOM) letter-DAT PREV-1-write-INDIC.AOR

'I am writing the/a letter.'

b. * (me) tseril-s da-ø-c'er-e

I(NOM) letter-DAT PREV-1-write-INDIC.AOR

'I am writing the/a letter.'

9) a. kmar-i (me) pul-s m-a-jl-ev-s

husband-NOM me(DAT) money-DAT 1-APPL-give-THEME-TNS

'(My) husband gives me (the) money.'

b. * kmar-i (me) pul-s ø-a-jl-ev-s

husband-NOM me(DAT) money-DAT Ø-APPL-give-THEME-TNS

'(My) husband gives me (the) money.' (cf. Hewitt 1995:121)

If Case and agreement are both expressions of a unique relation, then we expect to be able to pinpoint that relation precisely. Given that the locus of the relation is, by hypothesis, Agree, then the natural expectation is that Case should correlate with either match or value. We have just seen that Case assignment and value need not pattern together. A weaker hypothesis (in the sense that it allows a wider set of possibilities) is that Case is a reflex of match. (For this to work, an auxiliary assumption is necessary to the effect that match with an inactive intervener fails to assign Case, presumably due to properties of the intervener itself.) However even this weaker version of Case/Agreement unification fails to account for the Georgian problem, because in Georgian we have evidence for Case assignment without matching (10), as well as matching that does not lead to Case assignment (11).[6] In (10) the 2.sg SU (a) gets NOM Case despite the fact that it fails to match the #-probe on T. The same can be said for the 1.sg SU in (b). In (11) the SU *does* enter into match — with π-AGR on *v* — but it does not receive the expected ACC Case, it gets NOM instead. This cannot be due to a subsequent AGR relation with T (multiple case-checking scenario) as the outcome is the same regardless of whether or not SU matches the high #-probe (which it does in (b) but not (a)). If there *is* indeed multiple Case assignment happening in (11), it must be due to mechanisms that are independent of Agree. Either there is more than one mechanism for structural Case assignment; or some as yet unknown third factor is responsible for Case in both matching and non-matching

[6] The same problem arises for any language with split-φ and total underspecification of at least one interpretable φ feature. Note further that an even more general problem arises for all single-φ languages where there is a one-to-many ratio of AGR elements, not all of which need enter into Agree in my theory, but all of which (may) need Case.

environments.[7] Various takes are possible on what the latter conclusion might entail. One possibility is that Case cannot be adequately formalized as a reflex of feature checking at all. Another possibility is that it *is* driven by feature-checking, but the locus is not □, or at least not in the way □ are construed here.[8]

10) Case without match

key:	≠ *no Match*		□ *Match, no Value*		□ *Match, Value*	
a. (šen) m-xedav (me)			'You(NOM) see me(ACC)			
T	…	you(NOM)	…	*v*	…	me(ACC)
		$\pi_{[\text{PART}]}$		$\pi_{[\text{PART}]}$	□	$\pi_{[\text{PART}]}$
#	≠					
b. (me) g-xedav (šen)			'I(NOM) see you(ACC)[9]			
T	…	I(NOM)	…	*v*	…	you(ACC)
		$\pi_{[\text{PART}]}$		$\pi_{[\text{PART}]}$	□	$\pi_{[\text{PART}]}$
#	≠					

[7] We cannot claim that either of the Case assignment/match mismatches in (10) and (11) arise due to default Case or inherent case assignment, because they behave exactly as predicted given the TAM pivot. If NOM was default or inherent Case we would expect it to be insensitive to structural Case patterns in the clause.

[8] If we assume some independent feature □ to be responsible for Case, then the split-Case pattern is quite straightforwardly (albeit mechanically) derivable with the added assumption that the pivot of the split determines whether □ is high (accusative pattern) or low (ergative pattern). However, certain auxiliary assumptions must be introduced. For example, as noted previously, low AGR in accusative configurations conforms to Burzio's generalization, whereas low AGR in ergative configurations does not.

[9] Note that there is no morphological Case contrast on Georgian 1st and 2nd person pronouns. One might wonder whether we are justified in assuming abstract Case assignment here. Since overt pronoun use is emphatic in Georgian, we might suppose that they are licensed independently by focus (see 5.4). However, in clauses with both nominal and pronominal arguments, the morphologically visible Case on the nominal(s) is precisely what it would be if all Cases in the clause were being assigned, so I assume that 1st and 2nd person pronouns do get assigned Case (cf. Levin and Massam 1985 on Case assignment to CPs in Niuean).

11) Match without Case assignment

key:	≠ *no Match*			□ *Match, no Value*			□ *Match, Value*
a. (me) v-xedav (ma-s)				'I(NOM) see him-ACC			
T	…	*v*	…	I(NOM) …	*v*	…	him-ACC
		π	□	$\pi_{[\text{PART}]}$	π~~$_{[\text{PART}]}$~~	□	π
#	≠						
b. (čven) v-xedav-t (ma-s)				'We(NOM) see him-ACC			
T	…	*v*	…	we(NOM) …	*v*	…	him-ACC
		π	□	$\pi_{[\text{PART}]}$	π~~$_{[\text{PART}]}$~~	□	π
#	□	#					

The only generalization that seems to hold up for this data is that the match relation between AGR and a DO results in Case assignment independently of whether or not value also succeeds. The same is not true for match between the same AGR head and the SU. Thus, match between *v*-AGR and the DO results in *v*-Case (ACC) being assigned to the DO in both (10) and (11), whereas match with the SU in (11) does not. One possibility is that a first match is privileged over subsequent matches. The relevant notion of 'first match' must be first match for the probe, as opposed to first match for the goal. In (11) the target-controller pair <π, SU> is the result of first match for the goal, but second match for the probe, and accusative Case is not assigned. This is supported by the ergative Case pattern of (5). Here the prediction is that *v*-Case (absolutive) will be assigned to the specifier if (and only if) there is no NP in the domain of the 1st cycle. In other words, *v*-Case is only assigned to the S argument if the target-controller pair <*v*, S> is the result of first match for the *v* probe. Note that first match need not occur on the first cycle of Agree. In the case of (5), first match occurs on the second cycle. On the basis of this subset of attested patterns, I offer the following tentative generalization:

12) Case assignment correlates with first match.

The generalization is certainly compatible with canonical Case assignment configurations, but it remains to be seen whether it holds up to closer and broader empirical scrutiny. A serious consideration of these options is beyond the scope of the present work.

5.2 Unification of intervention effects

In chapter 2 I proposed a new class of interveners that match but fail to value, not because they are inactive (in the sense of Chomsky 2000, 2001) but rather because they are *inadequate* in their φ-specification and cannot meet the entailment condition on value defined in Chapter 2. This class of intervention effect is natural within the system, since such interveners simply fail to meet the conditions necessary for Agree. Inactive intervener, on the other hand, pose a puzzle: why should case assignment strip an NP of its ability to Agree? This is the question that is addressed in this section. I am not questioning the existence of defective intervention effects; I am asking what they tell us about the computation.

It would be helpful to have a hypothesis about what an inactive intervener actually looks like, in particular with respect to its φ-structure. The crucial characteristic of an intervener (of any variety) is that it must have sufficient φ-structure to satisfy match, but insufficient (or perhaps inappropriate) φ-structure for value. The canonical inactive intervener is an NP with θ-related/inherent Case. Bejar and Rezac (2003) propose that NPs with inherent Case are arguments of prepositional or applicative AGR heads, and therefore the asymmetry between structural and inherent Case reduces to whether NP enters into an AGR-relation at the clause level (where AGR is v^0 or T^0) or within the nominal complex, (where AGR is P^0 or $Appl^0$) (see also Bejar and Rezac in prep).

The languages we have seen in the previous chapters have 'fluid' TCP sets (as opposed to the rigid TCP sets of non-CSA languages) which demonstrates that the deactivation of an NP cannot be all-or-nothing since the same NP can enter into multiple AGR relations with different targets, so long as these are for different φ features. For example, in Georgian/Mordvinian/Dakota the DO can control low π-AGR and high #-AGR. Thus, the deactivation of a goal must be relative to individual φ-features, it cannot be a property of the NP as a whole. (See Rezac 2003 for a related proposal based for partial agreement in Basque.)

Our unanswered questions about Case make it difficult to formulate a precise theory of how deactivation of NP φ-features occurs. For present purposes I will characterize the contrast between active and inactive φ in terms of deletion. This gives us a parallel with our mechanism for deactivation of uninterpretable AGR features. It is unconventional to posit deletion of *interpretable* features, however this is not as drastic as it might seem, given that deletion has been defined here essentially as a diacritic for deactivation/inertness, and is to differentiated from actual erasure (as in Chomsky 1995, 2000, 2001). Thus, deletion makes features inert, but it does not actually strip them away (and it seems reasonable to adopt the position that no erasure should actually be necessary at any point in the derivation if a feature is interpretable).

I will take this rather simplistic approach and represent φ-deactivation by deletion. Thus, given an AGR head and an NP goal in a local configuration, there are four possible outcomes of a φ-probe.

13)

	AGR	NP	
a.	φ	φ	No value, no Case Assignment
b.	~~φ~~	φ	Value, but no Case Assignment
c.	φ	~~φ~~	Case Assignment, but no value
d.	~~φ~~	~~φ~~	Value and Case assignment

(b) and (c) are disallowed by Schütze, but allowed by Chomsky. I will remain agnostic on this question and abstract away from the question of what conditions *permit* interpretable φ-deletion on NP and whether/how they relate to Case. The relevant question here is, how does deletion relate to intervention.

We could simply stipulate that deleted φ is visible but unable to value, however this does not actually follow from the conditions on either operation. Fortunately, the empirical fact is that interveners *intervene*, so the question of whether deleted φ are visible need not be resolved on theory internal grounds; the answer must be that they *are* visible. Given this, the next question is why, if they are visible, can they not value (assuming entailment conditions are met). I propose that the answer to this resides in conditions on pied-piping (in the sense of Chomsky 2000, 2001). Following Rezac 2003, we have been assuming that the projection of a head is tantamount to copy movement, and the copied projection does not include valued/deleted subcategories of the head. In chapter 4 I presented independent morphological evidence in support of this approach. In light of this, I propose the following generalization:

14) Deleted features cannot be copied/moved

Now recall that the value operation was defined in chapter 2 as structure building, followed by deletion. If structure building entails actual feature movement of the interpretable □ of the goal then it would follow from this that inactive (i.e. deleted) interpretable □ cannot value because deleted features cannot be pied-piped. Thus we have a unification of the partial projection of uninterpretable □ and the inability to value inactive interpretable □. Furthermore, if valuing is feature-movement, then deletion of the interpretable □ on NP can be seen as deletion of the downstairs copy of the moved element (bringing it in line with other known cases of deletion of interpretable elements).

We can therefore hold that the sole condition on valuing that need be stated is the entailment condition given in Chapter 2. Independent conditions on feature movement are responsible for the inability of inactive □ to control agreement. This analysis of inactive □ provides a unified account for both the deactivation of structurally case-marked NPs as well as the inertness of inherently case-marked NPs. In the latter, the interpretable □ of the NP have already entered into an AGR relation with P/Appl and deleted. match with PP is presumably (by locality) detecting the valued/deleted uninterpretable features of P (and not those of the embedded NP itself). Nothing in the framework prohibits an Agree relation between two sets of uninterpretable features so long as the entailment conditions are met, but because valued features are deleted, they are unable to value a higher AGR. A schematization of these conditions with a 2nd person NP are shown below. In the (a) NP matches , values and deletes the π-AGR on P. Deleted π matches a higher π AGR, but cannot value because feature movement/pied-piping of deleted features is not allowed. In (b) the inactive NP has previously entered into an AGR relation that resulted in its interpretable π features valuing (i.e moving to) the AGR head, leaving a copy which deletes. If it is subsequently in a local configuration with another AGR with uninterpretable π, it will match, but it cannot value for the same reason as (a).

15)	a. Inherently case-marked intervener	b. Structurally case-marked intervener
	AGR [P NP] π □ ~~π~~ □ ~~π~~ | ~~PART~~	AGR $NP_{INACTIVE\ \pi}$ π □ ~~π~~ | ~~PART~~

5.3 Concluding Remarks

The proposals in this thesis have aimed to uncover the common architecture of agreement systems across languages. Amongst the goals of generative linguistics is to discover general conditions and principles that, by their interactions, yield the broadest possible empirical coverage. The hypotheses put forward in this thesis were arrived at with this aim in mind. Cross-linguistic differences between agreement systems were argued to reduce to the outcome of conditions on match and value, with the locus of variation being the specification and distribution of formal □ features. Derivational consequences of these outcomes are mediated by general locality principles, and by the cyclic character of syntactic computation.

The findings of this thesis support Chomsky's 2000, 2001 decomposition of Agree into two separate operations: match and value. In combination with the theory of φ-features presented here, it was shown to be possible to account for two kinds of anti-superiority effects (locality paradoxes) that are responsible for much of the surface complexity in agreement systems. Agreement patterns that have previously been considered to reflect mismatches between syntax and morphology were shown instead to reflect mismatches between probe, match and value. In canonical configurations, all three operations 'converge' in the sense that the probe is deactivated by a single cycle of Agree, and match and value are simultaneously satisfied by the highest NP in the AGR-domain of that cycle. However it is possible for match and/or value to fail, either creating an anti-intervention patter or leading to a second-cycle of Agree with an expanded AGR-domain and altered locality conditions.

The importance of feature theory is made salient in this thesis. First, intrinsic entailment relations between features are shown to be crucial to the determination of the outcomes of match and value. Second, the interaction between feature underspecification and locality was demonstrated to have important derivational consequences. Third, the selection/construction of formal feature inventories, was shown to be crucial to determining the core typological properties of a language; this was shown to be especially true with respect to the selection/distribution of uninterpretable features. New questions were articulated for future research concerning correlations between morphological and formal features, interpretable and uninterpretable features. Whether or not there are innate conditions on the distributional properties of uninterpretable features is one of the salient questions raised.

Finally, □-syntax was shown to have important consequences for the morphological component, creating conditioning contexts for suppletive allomorphy and other special forms, as well as offering new ways to view conditions on morphological well-formedness. It was shown that the vocabulary insertion operation must be sensitive to the distinction between inherently valued feature specifications, and feature specifications derived by Agree (second-cycle effects). Overall implications for the morphology-syntax interface are potentially substantial. On the one hand, a syntactic basis for erstwhile complex morphological patterns was established. On the other hand, new avenues for exploring systematicity within the morphological component itself were opened up. The syntactic redistribution of features (by Agree), and the introduction of new vocabulary insertion sites (by projection) offer new ways of exploring asymmetries in the distribution of special vocabulary choices.

References

Abondolo, Daniel. 1982. Verb paradigm in Erza Mordvinian. Folia Slavica 5: 11-24.

Adger, David, Susana Bejar & Daniel Harbour. 2001. Allomorphy: Adjacency and Agree. Paper presented at GLOW 24, Braga, Portugal.

Anagnostopoulou, Elena. 2003. The syntax of ditransitives: Evidence from clitics. The Hague: Mouton de Gruyter.

Anderson, Stephen. 1992. A-Morphous Morphology. New York: Cambridge University Press.

Aronson, Howard I. 1989. Georgian: A reading grammar. Columbus, OH: Slavica.

Arregi, Karlos. 1999. Person and number inflection in Basque. In V. Lin et al. (eds.), Papers in Morphology and Syntax, Cycle Two, MIT Working Papers in Linguistics 34, Cambridge, MA.

Avery, Peter, & Keren Rice. 1989. Segment structure and coronal underspecification. Phonology 6.179-200.

Bach, Emmon. 1964. Subcategories in transformational grammars. In H. Lunt (ed.), Proceedings of the Ninth International Congress of Linguists. The Hague: Mouton & Co.

Barrie, Michael. 2002. Pronominal agreement on Iroquoian nouns and verbs. Paper presented to the IVth meeting of the Niagara Linguistic Society.

Baker, Mark C. 1996. The polysynthesis parameter. Oxford: Oxford University Press.

Baker, Mark & Kenneth Hale. Relativized Minimality and Pronoun Incorporation. Linguistic Inquiry 21.2: 289-297.

Bejar, Susana. 1999. Structural markedness and minimalist checking theory. In S.F. Bird et al. (eds.), WCCFL 18: Proceedings of the 18th West Coast Conference on Formal Linguistics. Cascadilla Press. 27-42.

Bejar, Susana. 2000a. Structural markedness in formal features: Deriving interpretability. Revue québecoise de linguistique 28.47-72.

Bejar, Susana. 2000b. Locality, cyclicity and markedness in Georgian verbal morphology. Paper presented at GLOW 23, Vitoria, Spain.

Bejar, Susana. 2000c. Locality, cyclicity and markedness in Georgian verbal morphology. Generals paper, University of Toronto.

Bejar, Susana & Daniel Currie Hall. 1999. Marking markedness: the underlying order of diagonal syncretisms. In R. Daly and A. Riehl, eds., Proceedings of the 1999 Eastern States Conference on Linguistics. Ithaca, N.Y.: Cornell Linguistic Circle Publications.

Bejar, Susana & Milan Rezac. 2003a. Person licensing and the derivation of PCC effects. Selected papers from LSRL 32. Benjamins.

Bejar, Susana & Milan Rezac. 2003b. Cyclic Agree. Paper presented at the Lisbon Workshop on Agreement, Lisbon, Portugal.

Bejar, Susana & Milan Rezac. In prep. Cyclic Agree. Ms. University of Toronto.

Benveniste, E. 1966. Problèmes de linguistique générale. Paris: Gallimard.

Bittner, Maria. 1994. Case, Scope, and Binding. Dordrecht: Kluwer Academic Publishers.

Bittner, Maria & Ken Hale. 1996. The structural determination of case and agreement. Linguistic Inquiry 27.1: 1-68

Bloomfield, Leonard. 1962. The Menomini language. Yale University Press.

Bobaljik, Jonathan David. 1992. Nominally absolutive is not absolutely nominative. Proceedings of WCCFL XI. Stanford: CSLI.

Bobaljik, Jonathan David. 1993. Ergativity and ergative unergatives. In Colin Phillips, ed.MIT Working Papers in Linguistics 19: Papers on case & agreement II.

Bobaljik, Jonathan David. 2000. The Ins and Outs of Contextual Allomorphy. In K.K. Grohmann and C. Struijke (eds.), University of Maryland Working Papers in Linguistics, volume 10, 35-71.

Bobaljik, Jonathan David & Phil Branigan. 2003. Eccentric agreement and multiple case checking. Ms. McGill University & Memorial University of Newfoundland. [To appear in proceedings of the University of Toronto Ergativity Workshop.]

Bobaljik, Jonathan David & Susi Wurmbrand. 1997. Preliminary notes on Agreement in Itelmen. PF: Papers at the Interface. MITWPL 30. Cambridge, Massachusetts.

Bobaljik, Jonathan David & Susi Wurmbrand. 2002. Notes on Agreement in Itelmen. Linguistic Discovery 1.1.

Bonet, Eulàlia. 1991. Morphology after syntax: Pronominal clitics in Romance. Doctoral dissertation, MIT.

Branigan, Phil & Marguerite Mackenzie. 2001. How much Syntax can you fit into a Word? Late lnsertion and verbal agreement in Innu-aimûn. In S. Gessner, S. Oh & K. Shiobara (eds.), Proceedings of WSCLA 5: The Workshop on Structure and Constituency in Languages of the Americas . UBC Working Papers in Linguistics 5.

Bright, William. 1957. The Karok Language. Berkeley and Los Angeles: University of California Press.

Brittain, Julie. 2001. The Morphosyntax of the Algonquian Conjunct Verb: A Minimalist Approach. New York: Garland.

Buechel,Eugene. 1939. A Grammar of Lakota: The Language of the Teton Sioux Indians. Saint Francis, South Dakota: Saint Francis Mission.

Burzio, Luigi. 1991. The morphological basis of anaphora. Journal of Linguistics 27, 81-105.

Burzio, Lugi. 1992. on the morphology of reflexives and impersonals. In C. Laeufer and T.A. Morgan, eds. Theoretical analyses in oromance linguistics. Current issues in linguistic theory 74, 399-414. Amsterdam: John Benjamins.

Campana, Mark. 1992. A movement theory of ergativity. PhD Dissertation. McGill University, Montreal.

Carmack, Steven. 1997. Blocking in Georgian verb morphology. Language 73.2, 314-338.

Carstairs, Andrew. 1984. Constraints on allomorphy in inflexion. Bloomington, Indiana: Indiana University Linguistics Club.

Chomsky, Noam. 1951. Morphophonemics of Modern Hebrew. Unpublished Master's thesis, University of Pennsylvania.

Chomsky, Noam. 1956. Three models for the description of language. I.R.E. Transactions on Information Theory, vol. IT-2, pp.113-124. Reprinted, with corrections, in R.D. Luce, R. Bush and E. Galanter (eds.), Readings in Mathematical Psychology II. New York: Wiley.

Chomsky, Noam. 1965. Aspects of the theory of syntax. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1981. Lectures on government and binding: The Pisa lectures. The Hague: Mouton.

Chomsky, Noam. 1982. Some concepts and consequences of the theory of government and binding. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1986a. Knowledge of language. New York: Praeger.

Chomsky, Noam. 1986b. Barriers. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1991. Some notes on economy of derivation and representation. In Principles and parameters in comparative syntax, 417-454, ed. R. Freidin. Cambridge, Mass.: MIT press.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Hale & Keyser 1993b.

Chomsky, Noam. 1994. Bare Phrase Structure. MIT occasional papers in linguistics #5. Cambridge, Mass.: MITWPL.

Chomsky, Noam. 1995. The minimalist program. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2000. Minimalist Inquiries: The Framework. In Roger Martin, David Michaels & Juan Uriagareka (eds.), Step by Step: Essays on minimalist syntax in honor of Howard Lasnik, 89-153. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In Michael Kenstowicz (ed.), Ken Hale: A life in language, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001. Beyond explanatory adequacy. MITOPL 20. Cambridge, Mass.: MITWPL.

Chung, Sandra. 1998. The design of agreement. Chicago: University of Chicago Press.

Clements, George N. 1985. The geometry of phonological features. Phonology Yearbook

2.225-252.

Corbett, Greville. 1988. Agreement: A partial specification based on Slavonic data. In Agreement in Natural Language: Approaches, Theories, Descriptions. Michael Barlow & Charles A. Ferguson, eds. PP 23-53. Stanford: CSLI.

Cowper, Elizabeth. 2001. Feature geometry and verbal inflection. Toronto Working Papers in Linguistics 17.

Cowper, Elizabeth. 2003a. The geometry of interpretable features: Infl in English and Spanish. Manuscript. University of Toronto.

Cowper, Elizabeth. 2003b. Formal properties of feature geometries. Manuscript. University of Toronto.

Cowper, Elizabeth & Daniel Hall. forthcoming. The syntactic manifestation of nominal feature geometry. Proceedings of the 2002 meeting of the Canadian Linguistics Association.

Davies, William. 1986. Choctaw Verb Agreement and Universal Grammar. Studies in Natural Language and Linguistic Theory. Dordrecht: D. Reidel Publishing Company.

Déchaine, Rose-Marie. 1999. What Agonquian morphology is really like: Hockett revisited. In Papers from the Workshop on Structure & Constituency in Native American Languages. MIT Occasional Papers in Linguistics 17, ed. L. Bar-el, R-M. Déchaine & C. Reinholtz, 25-72. Cambridge, MA: MITWPL.

Déchaine, Rose-Marie and Charlotte Reinholtz.1998. Direct/Inverse as Case-Licensing. Paper presented at the Workshop on Structure and Constituency in the Languages of the Americas, Saskatchewan Indian Federated College.

Dixon, R.M.W. 1994. Ergativity. Cambridge: Cambridge University Press.

Dresher, Elan. 2003. The contrastive hierarchy in phonology. In Daniel Currie Hall (ed.), Toronto Working Papers in Linguistics 20: Special issue on contrast in phonology. University of Toronto.

Dresher, Elan, Glyne Piggott & Keren Rice. 1994. Contrast in phonology: an overview. In Carrie Dyck (ed.), Toronto Working Papers in Linguistics 13: Proceedings of the MOT Phonology Workshop, February 1994. University of Toronto.

Dyck, Carrie. 1992. Case assignment in Cayuga. Generals paper. University of Toronto.

Dupuis, Fernande, Monique Lemieux & Daniel Gosselin. 1991. The Consequences of Underspecification of the Features of AGR in the Identification of PRO. Language Variation and Change 3.3: 275-299.

Einarsson, Stefán. 1945. Icelandic: Grammar, texts, glossary. Baltimore: Johns Hopkins Press.

Embick, David. 1997. Voice and the interfaces of syntax. Doctoral dissertation, University of Pennsylvania.

Fehri, Abdelkader Fassi. 1992. Underspecification, Agreement, and Pronouns in Arabic. Revue quebecoise de linguistique 22.1: 117-144.

Forchheimer, 1953. The category of person in language. Berlin: Walter de Gruyter. cited in 2/19. consider taking out if inappropriate.

Gazdar, Gerald and Geoffrey Pullum. 1982. Generalized Phrase Structure Grammar: A theoretical synopsis. Bloomington, IN: Indiana University Linguistics Club.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum and Ivan Sag. 1985. Generalized Phrase Structure Grammar. Oxford: Basil Blackwell.

George, Leland M. and Jaklin Kornfilt. 1981. Finiteness and boundedness in Turkish. In Frank Heny, ed., Binding and filtering. Cambridge, MA: MIT Press.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), Universals of language, 73-113. Cambridge, MA: MIT Press.

Hale, Ken. 2002. Eccentric agreement. Kasu eta Komunztaduraren gainean [On Case and Agreemnent]. B. Fernández and P. Albizu. Vitoria-Gasteiz, Euskal Herriko Unibetsitatea: 15-48.

Hale, Ken and Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In K. Hale and J. Keyser (eds.), The view from Building 20, 53-109. Cambridge, Mass.: MIT Press.

Hale, K. and J. Keyser. 1997a. On the complex nature of simple predicates. In A. Alsina, J. bresnan & P. Sells, eds. Complex Predicates, 29-66. Stanford: CSLI.

Hale, K. and J. Keyser. 1997b. On the double object construction. Ms. MIT.

Hall, Daniel Currie (ed.). 2003. Toronto Working Papers in Linguistics 20: Special issue on contrast in phonology. University of Toronto.

Halle, Morris. 1990. An Approach to Morphology. Proceedings of the North East Linguistics Society 20, Vol. 1, 150-84. Amherst: GLSA, University of Massachusetts, Amherst.

Halle, Morris. (1992). Latvian Declension. Morphology Yearbook 1991, edited by Geert Booij and Jaap van der Marle, 33-47. Dordrecht: Kluwer.

Halle, Morris. (1995). The Russian Declension. Perspectives in Phonology, edited by Jennifer Cole and Charles Kisseberth, 321-353. Palo Alto: CSLI, Stanford.

Halle, Morris. (1997). Distributed morphology: Impoverishment and fission. MITWPL 30: Papers at the Interface, edited by Benjamin Bruening, Yoonjung Kang and Martha McGinnis, 425-449. Cambridge: MITWPL.

Halle, Morris and Alec Marantz. 1993. Distributed morphology andthe pieces of inflection. In The view from building 20: Essays in linguistics in honour of Sylvain Bromberger, ed. Ken Hale and Samuel J. Keyser. Cambridge, Mass.: MIT Press.

Hanson, Rebecca, Heidi Harley and Elizabeth Ritter. 2000. Underspecification and Universal Defaults for Person and Number Features. Proceedings of the 2000 Meeting of the Canadian Lingusitics Association. University of Toronto Working Papers in Linguistics.

Harley, Heidi. 1994. Hug a tree: deriving the morphosyntactic feature hierarchy. Papers on phonology and morphology, MITWPL 21, 289-230.

Harley, Heidi and Rolf Noyer. 1999. State-of-the-Article: Distributed Morphology. In GLOT 4.4, April 1999, pp. 3-9.

Harley, Heidi and Elizabeth Ritter. 2000. Structuring the bundle: a universal morphosyntactic feature geometry. Pronouns: Features and representation, ed. by Horst Simon and Heike Weise. 23-39. Amsterdam: John Benjamins.

Harley, Heidi and Elizabeth Ritter. 2002. A feature-geometric analysis of person and number. Language, 78.3.

Harrington, J.P. 1928. Vocabulary of the Kiowa Language. Bureau of American Ethnology Bulletin 84. Washington, D.C.: Government Printing Office.

Harris, Alice. 1981. Georgian syntax: A study in relational grammar. Cambridge: Cambridge University Press.

Harris, Alice. 1984. Inversion as a rule of grammar: Georgian evidence. In Studies in relational grammar 2, ed. D. Perlmutter and C. Rosen, 259-291. Chicago: University of Chicago Press.

Heap, David. 1999. Split subject pronoun paradigms: a feature geometry. Ms. University of Western Ontario, London, Ontario.

Hewitt, B.G. 1995. Georgian: A structural reference grammar. Philadelphia: John Benjamins.

Jakobson, Roman. 1936. Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus. Travaux du Cercle Linguistique de Prague VI 240–299. Reprinted in 1984 as 'Contribution to the general theory of case: General Meanings of the Russian cases' in L. Waugh and M. Halle (eds.), Russian and Slavic Grammar: Studies 1931-1981, 59–103. Berlin: Mouton

Jakobson, Roman. 1971. Shifters, verbal categories, and the Russian verb. Selected writings, II: Word and language, 130-47. The Hague: Mouton.

Johns, A. 1992. Deriving ergativity. Linguistic Inquiry 23, 57-87.

Johns, A. 1993. Symmetry in Labrador Inuttut. In J.D. Bobaljik and C. Phillips, eds. papers on Case & Agreement I, MITWPL 18, 43-57. Cambridge, Mass.

Johns, A. 1996. The occasional absence of anaphoric agreement in Labrador Inuttut. In Microparametric syntax and dialect variation, ed. James R. Black and Virginia Motapanyane, 121-143. Amsterdam: John Benjamins.

Kayne, Richard. 1989. Facets of Romance Past Participle Agreement. In P. Beninca' (ed.), Dialect Variation and the Theory of Grammar, 85-103. Dordrecht: Foris.

Koopman, Hilda. 1987. On the absence of Case Chains in Bambara. Ms. UCLA. [Published in 1992 NLLT 10, 555-94)]

Koopman, Hilda. 1996. The Spec Head configuration. In F. Lee and E. Garret (eds.), Syntax at Sunset. UCLA Working Papers in Syntax and Semantics, 1, 37-65.

Laka, Itziar. 1993. Unergatives that assign ergative, unaccusatives that assign accusative. In Papers on case and agreement I: MIT working papers in linguistis 18, 149-172.

Levin, Juliette and Diane Massam. 1985. Surface ergativity: Case/theta relations reexamined, pp 286-301. In Stephen Berman, Jae-Woong Choe and Joyce McDonough, eds. Proceedings of NELS 15.

Lieber, Rochelle. 1980. On the Organization of the Lexicon. PhD dissertation. MIT.

Marantz, A. 1989. Relations and configurations in Georgian, Ms, University of North Carolina, Cahpell Hill.

Marantz, Alec. 1991. Case and licensing. In Proceedings of ESCOL'91, 234-253. Ithaca, NY: DMLL Publications.

Massam, Diane. 1985. Case theory and the projection principle. Doctoral dissertation, MIT, Cambridge, Mass.

Massam, Diane. 2001a. Fully internal case: Surface ergativity can be profound. Paper presented at AFLA 8, MIT, April 27-29, 2001.

Massam, Diane. 2002. On predication and the status of subject in Niuean. In William Davies and Stanley Dubinsky, eds. Objects and other subjects. Kluwer.

McGinnis, Martha. 1995. Projection and position: Evidence from Georgian. In Proceedings of ConSole 4.

McGinnis, Martha. 1997a. Reflexive External Arguments and Lethal Ambiguity. Ms, MIT.

McGinnis, Martha. 1997b. Case and locality in l-syntax: Evidence from Georgian. In Heidi Halrey, ed. MIT Working Papers in LInguistics 32: Papers from the UPenn/MIT Roundtable on Argument Structure and Aspect.

McGinnis, Martha. 1998. Locality in A-movement. Doctoral dissertation, MIT.

McGinnis, Martha. 1999. Is there syntactic inversion in Ojibwa? In Leora Bar-el, Rose-Marie Déchaine, & Charlotte Reinholtz (eds.), Papers from the Workshop on Structure & Constituency in Native American Languages, 101-118. MIT Occasional Papers in Linguistics 17. MIT Working Papers in Linguistics.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. Language 67.3: 510-546.

Murasugi, Kumiko. 1992. Crossing and Nested Paths: NP Movement in Accusative and Ergative Languages. PhD Dissertation. Cambridge, Mass. MIT.

Nash-Haran, Léa. 1992. La catégorie AGR et l'accord en géorgien. Recherches Linguistiques 21, 65-79.

Nash, Léa. 1995. Portée argumentale et marquage casuel dans les langues SOV et dans les langues ergatives: L'exemple du géorgien. Doctoral dissertation, Université de Paris 8.

Nash, Léa. 1996. The Internal Ergative Subject Hypothesis. In Kiyomi Kusumoto, ed., Proceedings of NELS 26. GLSA, University of Massachusetts, Amherst.

Nevins, Andrew Ira. 2002. When 'we' disagrees in circumfixes. Talk presented at the 31st Annual Meeting of the Linguistic Association of the Southwest, Los Angeles and Pasadena, California.

Nichols, Lynn. 2001. The syntactic basis of referential hierarchy phenomena: Clues from languages with and without morphological case. Lingua 111. 515-537.

Noyer, Robert Rolf. 1992. Features, positions and affixes in autonomous morphological structure. PhD dissertation. MIT.

Perlmutter, David. 1971. Deep and Surface Structure Constraints in Syntax. New York: Rinehart & Winston.

Perlmutter, D. and R. Rhodes. 1988. Thematic-syntactic alignments in Ojibwa: evidence for subject-object reversal. Paper presented at the LSA.

Piggott, G. L. 1989. Argument structure and the morphology of the Ojibwa verb. In D. Gerdts and K. Michelson (Eds.), Theoretical Perspectives in Native American Languages. Albany, New York: SUNY Press.

Pollard, Carl & Ivan Sag. 1994. Head-Driven Phrase Structure Grammar. Chicago: CSLI.

Pollock, Jean-Yves. 1989. Verb-movement, Universal Grammar, and the structure of IP. Linguistic Inquiry 20, 356-424.

Reuland, Eric and Tanya Reinhart. 1995. Pronouns, Anaphors and Case. In H. Haider, S. Olsen and S. Vikner, eds. Studies in Comparative Germanic Syntax, 241-327. Dordrecht: Kluwer Academic Publishers.

Rhodes, R. 1994. Agency, inversion, and thematic alignment in Ojibwe. In Proceedings of BLS 20, 431-447.

Rezac, Milan. 2003. The fine structure of Cyclic Agree. Ms. University of Toronto. [to appear in Syntax].

Rezac, Milan. 2002. Cyclic domains and (mutiple) specifier constructions. Manuscript. University of Toronto.

Richards, Norvin. 1997. What moves where when in which language? Doctoral dissertation, MIT.

Richards, Norvin. 1999. Featural Cyclicity and the Ordering of Multiple Specifiers. in Working Minimalism, edited by Sam Epstein and Norbert Hornstein: MIT Press.

Rice, Keren & Leslie Saxon. 1994. The subject position in Athapaskan languages. In Heidi Harley & Colin Phillips (eds.), The morphology-syntax connection, MIT Working Papers in Linguistics 22, 173-195.

Ritter, Elizabeth. 1997. Agreement in the Arabic prefix conjugation: evidence for a non-linear approach to person, number and gender features. In Proceedings of the 1997 Annual Conference of the CLA, University of Calgary.

Rizzi, Luigi. 1990. Relativized Minimality. Cambridge, Mass.: MIT Press.

Roberts, Ian. 1991. Excorporation and Minimality. Linguistic Inquiry, 22, 209-218.

Rooryck, Johan. 1994a. On 0- and alpha- underspecification in syntax and phonology. In Heidi Harley & Colin Phillips (eds.), The morphology-syntax connection, MIT Working Papers in Linguistics 22, 197-216.

Rooryck, Johan. 1994b. On two types of underspecification: towards a feature theory shared by syntax and phonology. Probus 6, 207-233.

Sagey, Elizabeth. 1986. The representation of features and relations in non-linear phonology. Cambridge, MA: MIT dissertation.

Schacter, Paul. 1962. Review: R.B. Lees, Grammar of English nominalizations. International Journal of American Linguistics 28, 134-145.

Schütze, Carson. 1997. INFL in child and adult language: Agreement, case and licensing. Doctoral dissertation, MIT. Distributed by MIT Working papers in Linguistics.

Sigurðsson, Halldór Ármann. 1993. Agreement as head visible feature government. Studia Linguistica 47, 32-56.

Sigurðsson, Halldór Ármann. 1996. Icelandic Finite Verb Agreement. Working Papers in Scandinavian Syntax 57, 1-46.

Silverstein, Michael. 1985. Hierarchy of features and ergativity. In Pieter Muysken and henk van Riemsdijk (eds.), Features and Projections, 163-232. Dordrecht: Foris.

Sproat, Richard. 1985. On Deriving the Lexicon. PhD dissertation. MIT.

Starke, Michael. 2001. Move Dissolves into Merge: a Theory of Locality. PhD dissertation. University of Geneva.

Stockwell, Robert & Paul Schacter. 1962. Rules for a segment of English syntax. Ms. Los Angeles, University of California.

Stump, Gregory T. 2001. Inflectional Morphology: A Theory of Paradigm Structure. Cambridge: Cambridge University Press.

Taraldsen, Knut Tarald. 1995. On agreement and nominative objects in Icelandic. In Studies in comparative Germanic syntax, ed. H. Haider, et al, 307-327. Dordrecht: Kluwer.

Trommer, Jochen. 2002. Distributed Optimality. PhD Dissertation. University of Postdam.

Valentine, Randolph J. 2001. Nishnaabemwin Reference Grammar. Toronto: University of Toronto Press.

Watkins, L. with P. Mackenzie. 1984. A Grammar of Kiowa. Lincoln and London: University of Nebraska Press.

Wiklund, Anna-Lena. 2001. Dressing Up for Vocabulary Insertion: The Parasitic Supine. Natural Language & Linguistic Theory, 19.1: 199-228