Draft May 14 2004 - Comments welcome

Cyclic Agree

Susana Béjar and Milan Rezac

Abstract: We propose that agreement displacement phenomena sensitive to person hierarchies arise from the mechanics of Agree operating on articulated φ-feature structures in a cyclic syntax. Cyclicity and locality derive preference for agreement control by the internal argument. Articulation of the probe derives when the agreement control cyclically displaces to the external argument, and differences in cross-linguistic sensitivity to person hierarchies. The system characterizes two classes of derivations corresponding to direct and inverse contexts empirically, and predicts the existence and nature of repair strategies in inverse contexts. The properties of agreement displacement thus reduce to properties of syntactic dependency formation, Agree.

Keywords: Agree, person hierarchies, agreement displacement, features, cyclicity

1 Introduction

This papers explores a set of complex agreement systems characterized by sensitivity to *person-hierarchies* (PH); see among others among others Silverstein 1976, Zwicky 1977, Allen and Frantz 1978, Delancy 1981, Jelinek and Demers 1983, Gerdts 1988, Macaulay 1992, Jelinek 1993, Dixon 1994, Rice and Saxon 1994, Nash 1995, Fabri 1996, Lakämper and Wunderlich 1998, Hale 2001, Nichols 2001. We will argue that the basic patterns arise as *agreement displacement* (Hale's (2001) *eccentric agreement*), whereby perfectly general mechanics of the syntactic derivation, particularly the featural-dependency operation Agree, result in an apparently non-canonical agreement pattern. The fundamental principles that enter into the account are the following:[1]

(1) a. Intervener-based locality (Rizzi 1990), relativized to features (Chomsky 1995): Agree for a feature [F] is only sensitive to other goals with [F].
b. A fine-grained approach to cyclicity, where every Merge/Agree operation defines a potential cycle and thus a potential feeding-bleeding relationship (Rezac 2003).
c. A fine-grained approach to φ-features, and specifically person (π) features, associating with each person value a different feature structure and thus a different locality class (Béjar 2003).

For comments on the present (May 2004) version, we would like to thank the audience of GLOW 2004.

[1] Various types of syntactic accounts can be found for the kinds of PH phenomena we deal with (see note 4), within the same broad framework adopted here (thus setting aside for example Relational Grammar or Optimality Theoretic approaches). One class differs from ours in that it assumes person hierarchies as a primitive, e.g. Jelinek and Demers 1983, Gerdts 1988. Another class differs from us in that it deals with PH-phenomena in the morphology, Albizu 1997, and partly Wiltschko 2003. Closest to ours are proposals that derive PH effects from mechanics of Case and agreement in the derivation, such as Laka 1993, Hale 2001, Nichols 2001. A somewhat different family of approaches may be termed cartographic: different person values are either mapped to or licensed in different positions of the clause-structure, with ideally general principles deciding subject/object competition for a particular position; see Johns 1993, Rice and Saxon 1994 for the licensing approach. A recent example of the mapping approach is Jelinek (1993) which exploits Diesing's (1992) Mapping Hypothesis according to which specific and non-specific arguments must map outside and within the VP respectively. Jelinek differentiates groups of arguments within the person hierarchy (e.g. $1^{st}/2^{nd}$ from 3^{rd}) by giving them different specificity values (see Wiltschko 2003 for a perceptive critique). We do not attempt comparisons between these approaches and ours.

Our point of departure is a class of languages in which there is a (core) agreement system whose *controller* cannot be characterized in such terms as grammatical function or subject/object. Instead, the controller appears to alternate in transitives between the *external argument* EA and the *internal argument* IA. This is illustrated in (2) for *ergative displacement* in Basque (Laka 1993), where underlining indicates the relevant agreement slot in the example, and the controller in the translation. In (2)d, the underlined agreement slot marker tracks the *person* (π) features of the (EA), but elsewhere the IA:[2]

(2)

a. Basque $1\rightarrow2 = 2$
ikusi <u>z</u>-in-t-u-da-n
seen 2-X-PL-have-1-PAST
'I saw <u>you</u>.'

b. Basque $3\rightarrow1 = 1$
ikusi <u>n</u>-ind-u-en
seen 1-X-have-PAST
'He saw <u>me</u>.'

c. Basque $2\rightarrow1 = 1$
ikusi <u>n</u>-ind-u-zu-n
seen 1-X-have-2-PAST
'You saw <u>me</u>.'

d. Basque $1\rightarrow3 = 1$
ikusi <u>n</u>-u-en
seen 1-have-PAST
'<u>I</u> saw him.'

The choice between EA and IA as controller is clearly partly conditioned by their π-specification, since all other variables remain constant. We might characterizes this as a PH effect where the controller is given by some ranking of EA, IA on the basis of their π-specification, such as 1^{st} person > 2^{nd} person > 3^{rd} person, where > means "outranks."

However, this would be an inadequate characterization of the data. Any person hierarchy will underdetermine the choice of controller in (2). This is because although a $1^{st}/2^{nd}$ person argument will always beat out a 3^{rd} person argument, as in (2)b and (2)d, the choice between two $1^{st}/2^{nd}$ person arguments cannot be resolved by PH: in (2)a, 2^{nd} person beats 1^{st} person, while in (2)c, 1^{st} person beats 2^{nd} person.

We propose that the correct characterization is as follows: it is always the IA that controls agreement, *if it can do so*. The latter condition holds if the IA is $1^{st}/2^{nd}$ person, in which case IA agreement *bleeds* potential EA control of the agreement slot. If the IA is 3^{rd} person, it simply cannot control agreement; it is just in this context that a $1^{st}/2^{nd}$ person EA is the controller, so that failure of IA agreement *feeds* EA agreement. This is the phenomenon we identify as *PH-driven agreement displacement*, where displacement refers to apparently non-canonical control of a typically IA-controlled agreement slot by the EA. This phenomenon is what this paper is about. Two determining conditions enter into it: a principle that prefers IA to EA, which we will reduce to cyclicity, and sensitivity of Agree to person value of the controller, which we will reduce to the structure of π-features. We refer henceforth to the sum of these proposals as the theory of cyclic Agree.

The IA > EA preference has been observed in Bobaljik and Wurmbrand 1999 for Itelmen (Chukoto-Kamchatkan). Analyses of this pattern as PH-driven agreement displacement can be found for Basque in Rezac 2003, and for Georgian (Caucasian, Harris 1981, Hewitt 1995, Nash

[2] We use the following glosses: 1, 2, 3 person; SG singular, PL plural; INV inverse; OBV obviative; AP anti-passive; N nominative, A absolutive/accusative, D dative, GEN genitive, ABL ablative, POSS possessive; TNS tense; ASP aspect; FUT future, INF infinitive, PAST past. X glosses material not relevant to our point, or in example (2) material which cannot at that point be analyzed; we give *ind* here the same analysis as the theme suffix in Algonquian in section 3.2.

1995), Karok (Hokan, Bright 1957), and Erza Mordvinian (Uralic, Abondolo 1982) in Béjar 2003. These are illustrated in Table 1.

Table 1: PH-driven agreement displacement

Basque

IA controller		EA controller	
z-STEM	I VERB you	z-STEM	You VERB him
2-STEM		2-STEM	
n-STEM	You VERB me		
1-STEM			
n-STEM	He VERBs me	n-STEM	I VERB him
1-STEM		1-STEM	

Georgian

IA controller		EA controller	
g-STEM	I VERB you	∅-STEM	You VERB him
2-STEM		2-STEM	
m-STEM	You VERB me		
1-STEM			
m-STEM-s	He VERBs me	v-STEM	I VERB him
1-STEM-X		1-STEM	

Karok

IA controller		EA controller	
ki-STEM-ap	I VERB you.PL	ʔi-STEM	You VERB them
2.PL-STEM-X		2-STEM	
kin-STEM	You VERB us		
1.PL-STEM			
kin-STEM	He VERBs us	ni-STEM	I VERB them
1.PL-STEM		1-STEM	

Erza Mordvinian

IA controller		EA controller	
STEM-d-ad-yz	I VERB you.PL	STEM-s-y-t	You VERB them
STEM-T-2-PL		STEM-T-PL-2	
STEM-s-am-iz	You VERB us		
STEM-T-1-PL			
STEM-s-am-iz	He VERBs us	STEM-s-y-n	I VERB them
STEM-T-1-PL		STEM-T-PL-1	

This convergence, across languages, suggests an account grounded in principles of UG. We argue for a syntactic account, because we will show that the pattern arises through (i) feature-relativized conditions on Agree, giving rise to PH-effects, and (ii) conditions of search-space given by cyclic construction of the phrase-marker, giving rise to the IA > EA preference.[3] The basic idea that a class of PH-effects arises via the Case/agreement mechanism is due to Nichols (2001), who shows that oblique arguments, which do not interact with core Case/agreement systems, are invisible to the class of PH effects she considers. To the extent that Case and agreement are syntactic as in Chomsky 1995, 2000, Nichols's argument strongly supports a syntactic, Agree-based account. Thus, one of the main intended contributions of the present paper is a syntactic, derivational model of PH-driven agreement displacement.[4]

We will interpret the core pattern, where IA bleeds EA agreement, to mean that the relevant AGR head with a π-probe has only IA in its search-space at first, placing it on the *v* head:

(3) $[_{vP}$ EA $[v$+AGR $[_{VP}$ V IA]]]

The very existence of the bleeding pattern tells us that we are dealing with a single π-probe, which correlates with the fact that in all cases we are dealing with a single agreement slot for the core pattern. The two questions that arise immediately are: (i) why does the IA fail to control agreement if it has a certain person specification; (ii) how does this entail that it is then possible for the π-probe on *v* to Agree with the EA. We answer these questions in the first half of this paper (section 2). Throughout, we situate our approach in the framework of Chomsky 2000, where the conditions on Agree are given as follows (p. 122):

(4) 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[:]
 a. Matching is feature identity.
 b. D(P) is the sister of P.
 c. Locality reduces to "closest c-command"
 Thus, D(P) is the c-command domain of P, and a matching feature G is *closest to* P if there is no G' in D(P) matching P such that G is in D(G').

[3] Our discussion limits itself to basic transitive and intransitive clauses, referring to EA and IA, for convenience only. Whenever the argument of a lower clause falls into the scope of Agree as in ECM, causatives (see Nichols 2001:523 for this point), or cross-clausal agreement constructions (see Branigan and MacKenzie 2001, Bruening 2001:chapter 5 on direct/inverse voice in Algonquian), it behaves exactly as the IA that we limit ourselves to for PH-driven agreement displacement. For ditransitives, languages which treat the applied object as the primary object, IA, like Nishnaabemwin, require no special discussion; while those which do not, like Basque, are subject to the Person Case Constraint that we briefly touch on in 3.1, and the applied object effectively counts for match as the IA, though it cannot control agreement.

[4] This does not mean that there are no other PH-like effects, such as the use of passive for certain argument combinations, which arise elsewhere. Two kinds of examples come to mind. One is the use of a form like the passive when derivation of the transitive would produce a form that has no morphological spell-out in a language. In this case a derivation that is fine in the syntax simply crashes at the interface, and using another numeration is the only way of achieving a similar meaning. Here fall uses of the inverse or passive voice to avoid idiosyncratic person/number argument combinations like 2.SG/PL→1.PL in languages like Upper Chehalis (Wiltschko 2003), which may extend to languages like Southern Tiwa (Allen and Frantz 1978, 1983) where the pattern is more systematic (3→1/2) but is still apparently due to the lack of the relevant agreement morphology for the transitive active form. The other classes of examples are such restrictions like the fact that 1[st]/2[nd] person arguments cannot occur as agent of the passive in Southern Tiwa, which involve restrictions unrelated to the Case/agreement system, and about which we have nothing to say.

In the second half of this paper (section 3), we demonstrate that the derivational mechanics that account for PH-driven agreement displacement is also instrumental in accounting for another class of PH-sensitive agreement phenomena: disruptions of the core pattern in so-called *inverse* contexts where the EA is more highly specified than the IA for person on some scale. We will show that the theory of cyclic Agree characterizes inverse configurations as a class of computations, and provides the basis for accounting for their special properties.

2 Theory of cyclic Agree and PH-driven agreement displacement

2.1 Articulated probes and feature-relativized locality

We refer to the pattern where first the IA is evaluated with respect to Agree and then the EA as a *cyclic expansion* pattern: in section 2.2 we argue that it follows from cyclic construction of the phrase-marker, which leads to the IA becoming the first potential match for Agree by a probe on *v*, and EA the second. In this section, we address what it means for the IA to fail to control a π-probe on the first cycle, allowing subsequent Agree with the EA. We articulate π-features in general into a set of hierarchically organized features each of which can Agree independently, and each of which therefore defines a separate locality class. The IA will fail to Agree for a particular feature F of such an articulated π-probe simply when the IA lacks F; F on EA can then be the goal of Agree. Therefore, control by IA, (5)a, and by-passing of IA for control by EA, (5)b, display the same logic as classical feature-relativized locality for two arguments in (6), where NP_1 is a goal only if it bears F, (6)a, and is bypassed otherwise, (6)b.

(5) Cyclic expansion

a. ________ NP_1 Agrees

NP_2 H NP_1

F+ F- F+

b. ________ NP_2 Agrees

________ NP_1 Bypassed

NP_2 H NP_1

F+ F-

(6) Standard locality pattern

a. ________ NP_1 Agrees

H NP_1 NP_2

F- F+ F+

b. ____________ NP_2 Agrees

________ NP_1 Bypassed

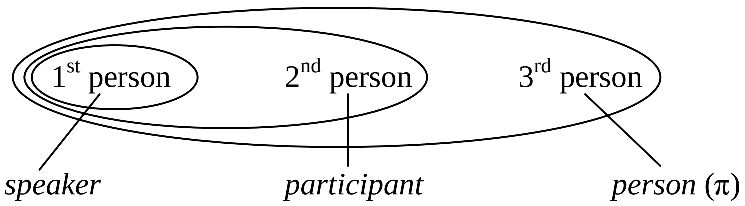
H NP_1 NP_2

F- F+

Both cyclic expansion as well as the standard locality patterns arise as a consequence of feature relativized locality, which is encoded in (4)a as the condition on matching: a probe for a feature F only sees the closest goal with a feature F in its search-space. The criteria for halting a search can thus be manipulated simply by manipulating assumptions about features. We take the data in Table 1 to establish that π-Agree must be sensitive to a fine grain of person specifications, so that π-Agree of a probe looking for a $1^{st}/2^{nd}$ person argument can ignore a 3^{rd} person NP, simply because the NP lacks the relevant feature(s). This suggests a system of features that lends itself to underspecification, so that the minimal contrasts within a subcategory like person can be captured in terms of the presence or absence of features.

One such system is developed by Harley and Ritter (2002) for morphological φ-features, which we extend to the φ-features visible to Agree, both interpretable and uninterpretable, following Béjar (2000, 2003). The φ-bundle is organized into subsets that reflect both natural classes and semantic entailment relations, as graphed in Figure 1 for a three-way person system. Here, all persons include some shared feature, our π. 1^{st} and 2^{nd} persons are, in addition to this, specified as discourse *participants* and thus grouped into a natural class to the exclusion of 3^{rd} persons. Finally, 1^{st} and 2^{nd} persons are themselves differentiated from one another by the further specification of 1^{st} persons with a *speaker* feature.

Figure 1: Entailment relations among person (π) features



This gives the entailments in (7); for example, being specified as [speaker] entails being specified as [participant] and as [π]. These entailments pattern translate to degrees of privative feature specification through a heuristic of logical underspecification, where π-values are differentiated only by the presence vs. absence of features, as in (8). This requires specifying default interpretations for underspecified representations: for example, [π] is common to all persons, but a bare [π] feature is interpreted as 3^{rd} person.

(7) Entailment: SPEAKER > PARTICIPANT > π

(8)	3^{rd} person	2^{nd} person	1^{st} person
	π	π	π
		PART	PART
			SPKR

We adopt these feature specifications, but for convenience we employ a shorthand from hereon in: we write [π] as [3], [participant] as [2], and [speaker] as [1], and refer to each of [3], [2], and [1] as a *segment*, meaning feature in a hierarchical feature structure given by entailment. The representations corresponding to (8) in this abbreviated system are given in (9). The notation is convenient because the most commonly assumed default interpretable of underspecified feature sets can be transparently read off the representation by inspection of the bottom-most segment in the feature bundle. It is crucial, however, that these segments not be read as person categories; *1* and *2* in (9) do not refer to the categories of 1^{st} and 2^{nd} person, but rather [speaker] and [participant], respectively. It is only the feature structure as a whole that corresponds to a traditional category like 1^{st} or 2^{nd} person.

(9)	3^{rd} person	2^{nd} person	1^{st} person
	3	3	3
		2	2
			1

The system permits limited variation in the selection of features. Of relevance in sections 2.2 and 3.2 will be that some languages differentiate 1st and 2nd persons by specifying the latter as [addressee] rather than specifying the former as [speaker], and by contrast interpreting a bare [participant] as 1st person. The specifications and default interpretations in such a system, translated into our shorthand, are given in (10):

(10)	3rd person	1st person	2nd person
	3	3	3
		1	1
			2

Now consider the consequences of modeling Agree in light of this feature-theoretic approach to π-specification. Matching can now be relativized for various persons along the entailment patterns of π-structures, using specifications of a probe: the more highly *articulated* a probe is, the more highly specified must an NP be to count as a match for all of a probe's features. (11)-(13) show this for the three possible articulations of the probe in the system just outlined: a *flat* probe that is just [π] ([3] in our notation) in (11), a medium-articulated probe in (12), and a fully articulated probe in (13). For each probe, an NP as (or more) highly specified as the probe will be a match for every feature of the probe (signified by a dash). However, an NP less specified than a probe will only match a subset of the probe's features, leaving an *active residue*, set in bold in (12) and (13). This active residue can, by feature-relativized locality, Agree with another NP in the search-space of the probe; it is this active residue that will lead to agreement displacement.

(11)	a.	*v*		NP	b.	*v*		NP	c.	*v*		NP
		3	—	3		3	—	3		3	—	3
								2				2
												1

(12)	a.	*v*		NP	b.	*v*		NP	c.	*v*		NP
		3	—	3		3	—	3		3	—	3
		2				2	—	2		2	—	2
												1

(13)	a.	*v*		NP	b.	*v*		NP	c.	*v*		NP
		3	—	3		3	—	3		3	—	3
		2				2	—	2		2	—	2
		1				**1**				1	—	1

(12) schematizes exactly what we are looking for to account for PH-driven agreement displacement in languages like those Table 1: a system where a 1st/2nd person IA will fully match a probe, but a 3rd person argument will leave the probe with an active residue, the segment 2, which may Agree with another argument.

The amount of specification of a probe is subject to variation, giving cross-linguistic differences in PH-sensitivity (Bejar 2003). For example, the fact that language families like Germanic and Romance do not show any PH-sensitivity in their agreement systems is modeled by assuming that they have the flat probe of (11); any NP will fully match a probe. On the other hand, languages like Dakota, Nishnaabemwin, Mohawk, and Kashmiri will be seen to have the fully articulated probe of (13), so that there is agreement displacement also between two 1st/2nd person arguments. Languages therefore vary parametrically in their choice of a *characteristic probe* for π, which determines their PH-sensitivity.

Three details must be discussed about the application of Agree to such feature structures. First, for feature structures as a whole we assume that some match must be found. If no segment in a feature structure is checked by Agree, then the derivation fails, following Bošković's (1997:140f) Inverse Case Filter (see also Lasnik (1999:chapters 4, 6), Chomsky (2000:125-7)). This is formalized as follows:

(14) Match Requirement: for a probe feature [F], some [F'] entailed by [F] must match.

Second, if an individual segment in an otherwise licensed feature structure finds no match by the end of the derivation, it is licensed by default and effectively ignored, and we write this by strike-through; this will be the fate of the active residue e.g. (13)a if not later valued from the EA.

Third, our use of feature structure in characteristic probes must be kept distinct from any use of feature structure as a PF instruction expressing valuation as a consequence of Agree. The feature structure of characteristic probes delimits conditions on deactivation of the probe, but it does *not* delimit the set of values to which agreement can be sensitive. Clearly, languages with a flat probe like Icelandic or Spanish are not restricted to 3rd person agreement; valuation of their probes can for e.g. distinguish 1st and 2nd persons as well. Likewise, a 3-2 probe language like Basque is not restricted to expressing just 3rd person and 2nd person agreement. Intuitively, this contrast between feature structure of the probe and feature structure of the *spell out* of the probe can be captured by construing valuation as some form of feature copying between target and controller. There are various ways to model this; we adopt the following assumptions:

(15) Agree

a. Each feature that seeks to Agree is distinguished by being *active* upon insertion into the derivation.
b. Agree for a feature [F] upon matching with a goal [F'] constitutes in copying [F'] to [F] and any feature that entails [F']; this copying constitutes valuing.
c. An active feature that is locally related to a non-active feature (as a consequence of (15)b) is no longer active.

Thus, what technically happens in (12)c is that the 3-segment of the probe, active, matches the 3-segment of the IA, the entire feature structure of the IA is copied to the probe, which deactivates both the 3 and 2 segments; while in (12)a, copying of the 3-segment of the goal to the probe upon match by the 3-segment of the probe leaves still active the 2-segment of the probe. For simplicity, we will throughout dispense with indicating copying (valuation), because what is important to us is the deactivation or lack thereof of individual π-features. This can be discussed more simply by indicating only match relationships between active features as in (12) and by speaking of checking between individual features/segments, with the understanding that the dash indicates in fact copying (valuation) and consequent deactivation, and that copying copies as much of the π-structures as there are on the goal, not just those which are active on a probe.

(14) and (15) together indicate the role played in Agree by organizing features into structures characterized by entailment (subset relations). In a sense, the entire structure behaves as a unit, in that it is the entire structure that is characterized by the Match Requirement, and that matching by any feature in the structure copies the entire structure of the goal to the probe, with attendant deactivation. However, individual features in the structure are capable of match on their own; that is, a feature F like [2] in (12)a is crucially capable of match once [3] has been deactivated.

To summarize, we adopt an approach to π-features that allows us to distinguish individual π-values by representing them as subsets of features of a single feature structure. Under the feature-relativized locality hypothesis implemented through using feature-identity as the criterion for matching, this means that the PH-sensitivity of agreement displacement can be modeled by (i) the fact that match of a proper subset of a the features of a probe by a goal leaves an active residue able to match another goal, and (ii) different cross-linguistic PH-sensitivities follow from different articulations of the probe. This now provides for correct interaction between π-probes and IAs; in the next section, we will derive the IA > EA preference of the cyclic expansion pattern.

2.2 Cyclicity and Agree

The pattern of PH-driven agreement displacement is preference for an IA controller, which is superceded by an EA controller if the IA does not suffice to check all segments of the characteristic probe of a language. The explanation of the IA-to-EA displacement lies in the derivational mechanics of a strongly cyclic construction of the phrase-marker (Rezac 2003), combined with a locating of the relevant π-probe on an AGR head, *v*, between EA and IA (Béjar 2000, 2003). We will now show how the EA falls into the search-space of a probe on *v* as *v* projects, if it retains an active probe.

In a strongly cyclic interpretation of the derivation, each operation stands in a potential ordering relationship with other operations. Suppose that the ordering of operations triggered by features on a single *locus* of the derivation (see Frampton and Gutmann 1999, Chomsky 2000:132, Collins 2002:46 on this concept) such as *v* is given by the Earliness Principle, which requires a feature to probe as early as possible (Pesetsky and Torrego 2001). Suppose further that we take seriously the idea that upon Merge, the label of the selector projects and that labels are non-distinct from lexical items (Chomsky 2000:133f.), modulo the effects of Agree (ibid., p. 126). Then we have two consequences: (i) any probe on *v* will first seek match in the object first Merged with *v*, the VP, because of the Earliness Principle (Rezac 2003); (ii) upon subsequent Merge of the EA and further projection of *v*, [Spec, *v*] falls into the domain of any remaining probe on *v* according to (4)b because *v*' is the sister of the new projection of *v* under Bare Phrase Structure (Rezac 2002). We walk through this below for a transitive construction, where we annotate the projections of *v* as v_n for convenience; each is identical to v_{n-1} modulo checking/valuation:

(16) Derivation of a transitive *v*P

Step 0: VP constructed as {V, {V, IA}}; *v* becomes locus

Step 1: Merge(*v*, VP) ⇒ {v_I, {*v*, {V, {V, IA}}}}

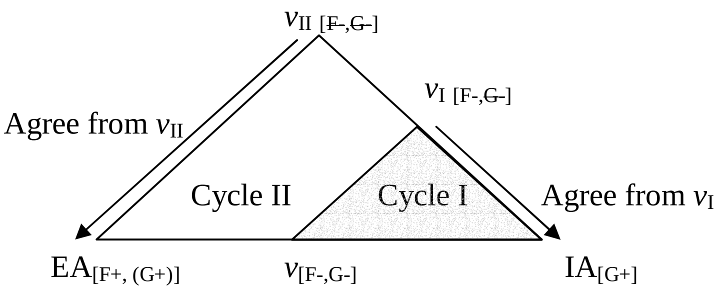
Step 2: Agree(v_I, IA)

Step 3: Merge(*v*P, EA) ⇒ {v_{II}, {EA, {v_I, {*v* {V, {V, IA}}}}}}

Step 4: Agree($v_{II,}$ EA), if there is still a probe on v_{II}

It is thus the cyclic architecture of the derivation, and locating a π-probe on *v*, that is responsible for PH-driven agreement displacement: we term this cyclic expansion of search-space. Thus schematically for two features (segments) F, G on the AGR head *v*:[5]

Figure 2: Cyclic search-space expansion



Consider how this works for Nishnaabemwin (Algonquian, Valentine 2001), which is a language with a fully articulated probe, and where as noted in 2.1 π-features have the structure [π [participant [addressee]]] so that 2nd person is the most specified, in our notation [3 [1 [2]]]. This means that if the IA is 3rd person, a 1st/2nd person EA will control agreement because the segments 1 and 2 are not affected by Agree with the IA, and will project unchecked to v_{II}; similarly, if the IA is 1st person, the segment 2 will project. The examples showing this fully articulated agreement displacement are in (17); the core agreement slot is underlined, while the slot glossed INV is treated in section 3.2. The notation $x \rightarrow y = x$ should be understood to mean that in a clause where the π-specifications of the EA are *x* and the IA *y*, agreement tracks *x*. The corresponding derivations are in Table 2, where a probe valued on a lower projection of *v* and thus deactivated on a higher one is in brackets, and strike-through indicates a segment that never matches in the derivation.

(17) a. Nishnaabemwin 1→2 = 2
g-waabm-in
2-see-1.INV
'I see <u>you</u>.'

b. Nishnaabemwin 2→1 = 2
g-waabm-i
2-see-DFLT.1
'<u>You</u> see me.'

c. Nishnaabemwin 3→1 = 1
<u>n</u>-waabm-ig
1-see-3.INV
'He sees <u>me</u>.'

d. Nishnaabemwin 3→2 = 2
g-waabm-ig
2-see-3.INV
'He sees <u>you</u>.'

---

[5] It should be clear that what is crucial is the derivational mechanics and the position of the π-probe between EA and IA, not the actual identification of the AGR head with *v*, the introducer of the EA. If the AGR head is on a head H c-commanded by *v*, syntactic head-movement can bring H to *v*, and along with H any still-active features of H. These features then probe from H+*v*, if head-adjunction is so defined that the entire head adjunction structure counts as the locus/pivot of the derivation; this captures some of the effects of Baker's (1988:64) Government Transparency Corollary (cf. Roberts 1991, Chomsky 1995).

Table 2: Core agreement in Nishnaabemwin

EA→IA	1				2			
1	--				v_{II}	EA	v_I	IA
					(3)	3	3	— 3
					(1)	1	1	— 1
					(2)		2	— 2
2	v_{II}	EA	v_I	IA	--			
	(3)	3	3	— 3				
	(1)	1	1	— 1				
	2	— 2	2					
3	v_{II}	EA	v	IA	v_{II}	EA	v	IA
	(3)	3	3	— 3	(3)	3	3	— 3
	(1)		1	— 1	(1)		1	— 1
	~~2~~		2		(2)		2	— 2

The same logic accounts for the PH-driven agreement displacement discussed in section 1 for languages like Basque, ex. (2); there we are dealing with a 3-2 probe, so that a 3rd person IA will not check the 2-segment, while both a 2nd and 1st person IA will. In a language with a flat probe like English, the IA will always count as a match for the sole segment of the probe, and there will never remain any active residue on v_{II} for the EA.

An important consequence of the cyclic expansion mechanics for agreement displacement is the prediction of *second cycle effects* (cf. Béjar 2003:79). Agree with the IA takes place on a different cycle than Agree with the EA, and Merge of the EA intervenes between the two cycles. This difference in derivational mechanics is reflected in the morphology of agreement in languages like Dakota (a 3-2-1 probe) and Karok (a 3-2 probe). In such languages, a π-probe value F on *v* is systematically spelled out using one morpheme if it has been valued on the first cycle (IA ≥ EA), and a different morpheme if it has been valued on a second cycle (EA > IA). A second cycle effect is different from the morphological contrast that might arise in a language with separate AGR heads to track the IA and the EA, because second cycle morphology correlates precisely with PH-driven agreement displacement in a language; it occurs only with those EAs that are more specified than IAs with respect to the characteristic probe of a language.

Consider Dakota (Siouan, Buechel 1939, Schwartz 1979, Miner 1980). Dakota has are two classes of intransitives verbs, which we identify as unergative and unaccusative.[6] π-agreement with the subject is spelled out differently in the two classes of verbs. With unaccusative verbs the subject is the IA and thus always a first-cycle controller, with unergative the EA and thus always a second-cycle controller; we thus identify the morphology involved as first (I) vs. second (II) cycle morphology spelling out the values of a single π-probe on *v*:

[6] In Dakota terminology agentive and non-agentive; this classification does not necessarily reduce to unergative and unaccusative (cf. Baker 1996:212f. for Mohawk). It suffices for us that the agentive class base-generate its subject higher than the *v* probe, and the non-agentive class lower, as in Travis 2003. The position of the former does not have to correspond to the position of the transitive EA introduced by its special head *v* if the core π-probe is not placed on *v* but another AGR head between EA and IA; see note 5.

Table 3: Dakota agreement morphology

Person	First-cycle (I)	Second cycle (II)
1	ma	wa
2	ni	ya
3	∅	∅

(18) a. unaccusative
ma-zí
1.I-pale
'I am pale.'

b. unergative
wa-wáChi
1.II-dance
'I dance.'

If we are right that the contrast between the two series of Dakota affixes reflects a contrast between 1st cycle and 2nd cycle morphology, then we expect the same contrast to occur in the transitive paradigm in the appropriate environments. This is in fact what we find. In the transitive paradigm of Dakota, which is a fully articulated probe language (presented in more detail in 2.3), the second cycle morphology turns up exactly when EA > IA, and the first cycle morphology appears elsewhere:

(19) a. $3 \rightarrow 1$ = 1.I
ma-khíze
1.I-attacked
'He attacked me.'

b. $1 \rightarrow 3$ = 1.II
wa-khíze
1.II-attacked
'I attacked him.'

Thus the distribution of the unaccusative/unergative series in the transitive paradigms correlates with PH-driven agreement displacement from IA to EA. The environments are unified by referring to the derivational mechanics. Second cycle morphology shows up when the π-probe is valued on the second cycle, which is from the second projection of *v*, v_{II}, c-commanding the EA; first cycle morphology when it is valued on the first cycle, from v_I (here ↔ symbolizes Agree).

(20) $[v_{II}$ EA $[v_I$ $[$V IA$]]]$
$\pi\text{-} \leftrightarrow \pi\text{+}$ $\pi\text{-}$ $\leftrightarrow$ $\pi\text{+}$

This morphological sensitivity to stages of syntactic derivation can be given a natural expression in a realizational theory of morphology like Distributed Morphology (Halle and Marantz 1993). With post-syntactic vocabulary insertion, the second-cycle effect can be reduced to contextual allomorphy so long as we allow each projection of *v* to be a potential vocabulary insertion site (a move that is, in turn, natural given Bare Phrase Structure; see Bejar 2003 for further evidence supporting this approach). Adjacent projections of *v* are in a sufficiently local relation with one another to trigger alternations that can be captured by vocabulary insertion rules of the form in (21):

(21) A set of features Γ of v_n is spelled out as the phonological string σ in the environment of $v_{n\pm1}$

Thus, in the example of Dakota 1st person agreement (19),

(22) Vocabulary insertion rules for Dakota 1st person agreement

a. First-cycle vocabulary item

ma $\leftrightarrow [1]_{v\mathrm{I}}$

b. Second-cycle vocabulary item

wa $\leftrightarrow [1]_{v\mathrm{II}} // \_\_ v_{\mathrm{I}}$

As with all cases of grammatically or lexically conditioned allomorphy, the occurrence of the alternation cannot be said to be predicted. Whether or not a language manifests this morphological contrast is a language-specific and (at least synchronically) idiosyncratic property. We expect that some languages will do so fully, like Dakota, while others will not at all, or only partially (see Nishnaabemwin, in section 3). It suffices to establish that cyclic expansion creates a plausible conditioning environment for such allomorphy. In this case, the plausibility of the environments rests on the locality between target and trigger of the alternation (Lieber 1980, Sproat 1985, Bobaljik 1995), which in turn follows automatically from the intrinsic locality between adjacent projections. Given the plausibility of this environment, we might expect other variants of this kind of allomorphy to occur. For example, vocabulary insertion on v_{II} could be sensitive not only to the existence of the adjacent projection v_{I}, but also to particular features on v_{I}; or v_{II} could be realized by morphology that distinguishes it v_{I} from independently of Agree. We will see examples of both kinds in section 3.

In this section we have shown how the search-space of any probe on *v* expands from including first the IA and then, if it remains active, the EA, through independent minimal assumptions about the derivation. The pieces of our account are now in place: an articulated structure of person features, which permits PH effects, and the mechanics of cyclic displacement, which captures the IA > EA preference, and predicts the existence of second cycle effects.

2.3 Dakota

In this section we go more carefully through a partial Dakota paradigm, to show how the cyclic Agree mechanism developed above derives the basic pattern of PH-driven agreement displacement, as well as second cycle effects. Dakota is a language with a fully articulated (3-2-1) probe. The illustration is typical in the sense that the agreement-displacement pattern exemplified from here shows up throughout in the 3-2-1 languages (Dakota, Nishnaabemwin, and Mohawk), but of the examples explored here only Dakota shows pervasive second cycle effects in the morphology.

The Dakota singular agreement paradigm is given in Table 4. The underlined morphemes are the core agreement slot that is so far fully predicted by the system, using the morphemes given in Table 3 above.

Table 4: Dakota singular paradigm

EA→IA	1	2	3
1	—	c^hĩ-kte ‡ 1/2-killed 'I killed you.'	wa-kte 1.II-killed 'I killed him.'
2	ma-yá-kte ‡ 1.I-2.II-killed 'You killed me.'	—	ya-kte 2.II-killed 'You killed him.'
3	ma-kte 1.I-killed 'He killed me.'	ni-kte 2.I-killed 'He killed you.'	kte killed 'He killed him.'

Table 5 summarizes the cyclic Agree derivations for this paradigm. Agree for an individual feature (segment) is represented by —. First and second cycle Agree, which correspond to Agree taking place from the first (v_I) vs. second (v_{II}) projection of v, are represented by — to the right of the probe (first cycle) and to the left of the probe (second cycle), which is both convenient and clear.

Table 5: Cyclic Agree for the Dakota singular paradigm

EA→IA	1	2	3
1	--	EA AGR IA 3 3 — 3 2 2 — 2 1 — 1	EA AGR IA 3 3 — 3 2 — 2 1 — 1
2	EA AGR IA 3 3 — 3 2 2 — 2 1 — 1	--	EA AGR IA 3 3 — 3 2 — 2 ~~1~~
3	EA AGR IA 3 3 — 3 2 — 2 1 — 1	EA AGR IA 3 3 — 3 2 — 2 ~~1~~	EA AGR IA 3 3 — 3 ~~2~~ ~~1~~

The shaded cells in the paradigm are those where there is only on Agree step, with the IA; the probe has no segments left that can Agree with the EA. The remaining filled cells are those where the 3-2-1 characteristic probe of Dakota does have an active segment left after Agree with the IA, and this segment Agrees with the EA on its second cycle. Cyclic Agree thus directly characterizes two classes of derivations, which turn out to be natural classes because they are found independently in the empirical data:

(23) Direct/inverse contexts

a. *Inverse context* (shaded): when IA checks the characteristic probe of a language as fully as possible, so that the EA cannot itself Agree with it at all. For Dakota, (1)/2/3→1; (2)/3→2;

3→3; for Basque, 1/2→1/2, 3→1/2/3; for English, all contexts. There is no Agree of the core π-probe of *v* with the EA in inverse contexts.

b. *Direct context* (unshaded): when EA is more specified than the IA, so that after the characteristic probe of a language has Agreed as fully as possible with the IA, it then Agrees for its unchecked segments with the EA. For Dakota, 1→2/3, 2→3; for Basque, 1/2→3; for English, not possible. In direct contexts, the core π-probe of *v* Agrees with both the IA and EA both, for different segments.

The characterization of direct and inverse contexts will be crucial in section 3.

Continuing the exposition of the paradigm, two cells in Table 4, signaled by ‡, do not fully meet our expectations: the 1→2 and the 2→1 cells. For the 1→2 form, we would expect **wa-kte* using the 1.II morpheme *wa*. What we see instead is a special morpheme, *c^h^ĩ*, which by contrast to other forms in the paradigm tracks both 1st and 2nd person; it is a portmanteau for the 1→2 combination. The use of a special morpheme precisely in 1→2 contexts (and not e.g. 2→1, 1→3), is not unique to Dakota; the same phenomenon shows up in exactly the same environment in Mohawk and in Algonquian (already pointed out in 3.2).

Interestingly, our system permits a precise characterization of the relevant environment: it is the only direct context where the two arguments are 1st/2nd person. We take the recurrent existence of portmanteau morphology precisely here as further evidence that there is indeed a single π-probe that Agrees with both the IA and the EA on two cycles. The spell-out of such a derivation can make reference to both steps of Agree because they occur from different projections of v, v_{I} vs. v_{II} respectively:

(24) Insertion rules of Dakota 1st person morphemes

a. First cycle: ma ↔ $[1]_v$

b. Second cycle: wa ↔ $[1]_v$ // __v

c. Portmanteau: c^hĩ ↔ $[1]_v$ // __$[2]_v$

Thus the system naturally limits portmanteau morphology to direct contexts where the probe enters into separate steps of Agree for both EA and IA. The reason why only 1st/2nd person arguments are subject to portmanteau rules like (24)c is intuitively obvious due to their more marked status, though it remains to be determined whether UG actually disallows insertion rules referring specifically to the context __$[3]_v$ distinct from __v. The cross-linguistic robustness of portmanteau morphology in this environment is especially suggestive given the fact noted earlier that contextual allomorphy of this kind is almost by definition unforced. We leave to future research the matter of why allomorphy seems to in fact be forced here.

The second anomalous form, the 2→1, is quite different. The core agreement morphology that it expresses, *ma* 1.I, is as predicted. However, there is also an extra agreement morpheme, which tracks the EA. This turns out to be a general feature characterizing in many languages exactly inverse contexts. In the next section, we will show that this is a systematic consequence of the need of person features to be licensed.

3 The Person Licensing Condition and inverse contexts

3.1 Introduction

The insertion of extra agreement morphology only in the inverse contexts turns out to be a subset of a more general pattern characterizing highly diverse languages traditionally described as sensitive to person hierarchies. Inverse contexts are typically morphologically distinguished from direct contexts by special morphology; we may speak of disruption to the core paradigm. In Dakota we see an extra agreement slot; this is true of Mohawk and Algonquian as well. Kashmiri instead puts the object into an apparent oblique Case and does not Agree with it. We will speak of both of these phenomena as repair strategies for reasons that will become clear.

Both strategies are precisely coextensive with the inverse, and they are thus *derivational* strategies: their availability depends on the person specifications of EA and IA in a particular derivation. In this they contrast with *periphrasis,* such as the use of an independently available passive, which are not limited to inverse contexts. It is therefore very significant that the mechanics of PH-driven agreement displacement developed in section 2 independently characterizes two classes of computations by whether or not the core π-probe Agrees with the EA; inverse derivations are those where it does not. It is these derivations that are characterized by repair strategies.

The nature of what happens in the inverse is suggested by the nature of the repair strategies. The extra agreement slot of Dakota tracks the EA, which would otherwise never Agree in inverse contexts. We take this to reflect the introduction of an extra π-probe. The same result will turn out to be true of the inverse-only oblique Case on IA in Kashmiri in 3.3: it allows non-agreement by the core π-probe with the IA, so it can Agree with the EA. Therefore, we propose that the special character of inverse contexts, by which they are subject to these strategies, arises from the fact that the EA never Agrees with the core probe. We take the derivational disruption of non-agreement contexts to suggest a requirement that π-features be licensed by Agree, the *Person Licensing Condition* (PLC); the repair strategies satisfy the PLC.

(25) Person Licensing Condition: a person feature must be licensed by Agree of some feature F' that it (reflexively) entails.

The PLC is proposed in Béjar and Rezac 2003 for the Person Case Constraint (Bonet 1991, Boeckx 2000, Ormazabal and Romero 2003), where there is a single π-probe c-commanding two NPs as in (26). The closer NP_1 is "quirky", which means it is visible to and moveable by the π-probe, but at the same time oblique so that it cannot value it; after it is moved by the probe, there is no π-probe left on AGR for NP_2, though there may be a number probe (Chomsky 2000:127-128, Anagnostopoulou 2003:chapter 5, Béjar and Rezac 2003). This leads to a crash of such derivations if the farther NP_2 does not have another sufficiently local probe to Agree with it and license its person feature. We posit that inherent Case and focus both involve shells around a DP that contain a local φ-probe for this purpose (PP for inherent Case; cf. Cardinaletti and Starke's (1999:179-192) γ' and Travis and Lamontagne's (1992) K, both of whom compare it to a complementizer); these are not present for clitics or *pro*. Consequently, the Person Case Constraint blocks an NP with a person feature where the closest AGR head is separated from it by another NP.[7]

[7] We cannot further discuss the PCC here for reasons of space. In the languages we are looking at such as Mohawk, 3rd person animates out-rank 3rd person inanimates; the contrast seems to be that the former are specified for [3] and the latter simply absent. This is consistent with Ormazabal and Romero's (2003) argument that in languages like Spanish, grammatical animacy is the relevant notion of "person" in the PCC, so that NP_2 cannot be animate 3rd person. However, other languages draw the line slightly differently; Anagnostopoulou's (2003:199f.) finds that

(26) PCC: In [$_\alpha$ AGR … NP_1-oblique … NP_2 …], where α includes no other NP or AGR, NP_2 cannot have a marked person feature ($1^{st}/2^{nd}$, with variation for 3^{rd}).

(27) a. Je le$_1$ /*te$_1$ lui$_2$ ai présenté t_1 t_2. PCC for *v*
I 3.SG-A /*2.SG-A 3.SG.D have introduced
I introduced him/*you to her. (French)

b. Okkur *þóttuð/%þótti þið fyndin. PCC for T
us-D though-2.PL/3.SG you.PL.N amusing
We found you amusing. (Icelandic; Hrafnbjargarsson 2001:18)

The PLC bears obvious similar to the *Case Filter* as a requirement on NPs, distinct from the Inverse Case Filter as a requirement on probes. Indeed, given our approach to inherent Case as including a DP/PP-internal π-probe, the two seem to have the same empirical strategy; the PLC is simply a way of spelling out what our approach needs the Case Filter to specifically require.

In the next two subsections, we examine in detail the two derivational strategies used for avoiding PLC violations in inverse contexts. The important result of cyclic Agree at this point is that it directly characterizes the problematic derivation set, inverse contexts, as those where a single (articulated) π-probe does not reach the EA. The PLC identifies what goes wrong in inverse contexts, why they cannot converge unless there is a repair strategy that provides the EA with a π-probe, and the nature of the repair strategies.

3.2 Added probe

The following partial paradigm of the singular from Mohawk (Iroquoian, Lounsbury 1953, Beatty 1974, Postal 1979:chapter 3, Baker 1996:chapter 5) exemplifies the added probe phenomenon even more clearly than Dakota. In Dakota most 3^{rd} person morphology is ∅, so we only see the extra agreement slot overtly in 2→1 combinations. This is not the case in Mohawk where 3^{rd} person agreement is overt, so it can be observed transparently that the distribution of the extra agreement slot (here underlined, beside the core agreement) is coextensive with the inverse.[8]

animate NP_2 clitics are fine in Greek active ditransitives provided there is φ-Agree between *v* and NP_2 (diagnosed by cliticization of the intervening dative). Since dative-nominative constructions demonstrate that in PCC contexts the number/gender probe can by-pass the trace of NP_1 after it has been dislocated by the π-probe of AGR and Agree with NP_2 (Anagnostopoulou 2003:chapter 5), it must be this probe that is licensing an animate 3^{rd} person in these languages, which is consistent with another interpretation of the structure of 3^{rd} persons.

[8] In 3→3 combinations in Mohawk and Nishnaabemwin, we pick the forms where in the different 3^{rd} person specifications of these languages, dependent on gender and obviation respectively, EA ≤ IA (obviative→proximate in Nishnaabemwin, masculine→neuter in Mohawk); there are also direct 3→3 direct forms when EA > IA, and which show up as expected: only the core agreement slot valued to 3.

Table 6: Singular paradigm for Mohawk (added probe underlined)

EA→IA	1	2	3
1	—	ku-*see* ‡ 1/2 'I see you.'	k-*see* 1 'I see him.'
2	(h)s-k-*see* 2-1 'You see me.'	—	hs-*see* 2 'You see him.'
3	wa-k-*see* 3.INV-1 'He sees me.'	(h)s-(w)a-*see* 2-3.INV 'He sees you.'	hra-wa-*see* > hra-o-*see* 3.M-DFLT 'He sees him.'

The role of the added probe in PLC licensing the EA becomes clear in considering the actual derivations. It is exactly in inverse contexts where person features of the EA would not enter into Agree were it not for the added probe. This is shown in Table 7 (for a 3-2-1 language like Mohawk), where the added probe is in bold; only the segments of the added probe that Agree are actually shown (for reasons to become clear in 3.4):

Table 7: Cyclic Agree and added probe

EA→IA	1	2	3
1	—	EA *v* IA 3 3 — 3 2 2 — 2 1 — 1	EA *v* IA 3 3 — 3 2 — 2 1 — 1
2	EA *v* IA 3 — **3** 3 — 3 2 — **2** 2 — 2 1 — 1	—	EA *v* IA 3 3 — 3 2 — 2 ~~1~~
3	EA *v* IA 3 — **3** 3 — 3 2 — 2 1 — 1	EA *v* IA 3 — **3** 3 — 3 2 — 2 ~~1~~	EA *v* IA 3 — **3** 3 — 3 ~~2~~ ~~1~~

The added probe is thus responsible for PLC licensing. Its character as a derivational option is evident from its limitation to inverse contexts, which depend for their characterization on the EA ≤ IA relationship with respect to a characteristic probe. The latter point can be illustrated with Basque, which is a 3-2 probe language, unlike the 3-2-1 probe of Dakota and Mohawk.[9] In Basque PH-driven agreement displacement illustrated in (2), there is extra agreement morphology beside the underlined agreement slot in (2)a and (2)c that has not been discussed. This morphology tracks the EA whenever the core agreement slot does not, that is outside of

[9] In languages like Georgian, it is not clear that any added probe morphology shows up, where our account would predict there to be one. However, nothing forces overt spell-out of an added probe.

direct contexts with respect to a 3-2 probe, which are 1/2→3. The independent effect of the characteristic probe, which is used to capture when IA fails to control agreement, here determines when the added probe show up: it shows up in Basque in 1→2 (ex. (2)a) because that is an inverse contexts with respect to a 3-2 probe (and IA controls core agreement), but not in Mohawk because it is a direct context with respect to a 3-2-1 probe (and EA controls core agreement). Therefore, we have a very strong relationship between the cyclic Agree account for core agreement, which is used to determine when there is IA > EA agreement displacement and when there is not, and the distribution of the added probe, which shows up exactly when there is no displacement: in inverse contexts.[10]

Empirically, the added probe does not interfere in Agree between the core probe and the IA, so it seems to be added at the point in the cyclic derivation after Agree between the core probe and the IA has taken place. An important cue in narrowing down the nature of the added probe comes from its spell-out in Dakota. There the added probe makes use of the same morphology as a second cycle core probe (the 2→1 combination in Table 4, repeated here with the added probe underlined):

(28) ma-<u>yá</u>-kte
1.I-2.II-killed
'You killed me.'

We have identified second cycle probes as those that probe the EA has been added by cyclic expansion from the projection of *v* above the EA. This captured the restriction of second cycle morphology for core probe spell-out to direct contexts; only in those cases will the core probe Agree with the EA on a cycle after it has Agreed with the IA. We will suppose, therefore, that Dakota provides evidence that an added probe is added on the projection of *v* above the EA, v_{II}. Given the Dakota morpheme insertion rules, a particular π-value on v_{II} will be spelled out using the same morphology whether it reflects a second-cycle core probe projected from v_{II} or an added probe when the core probe has been fully checked on v_{I}.

Because we model second-cycle morphology as contextual allomorphy, we do no expect added probes to be uniformly realized across languages; rather we expect free and even arbitrary variation. While spell-out of the added probe in Dakota replicates that of second cycle probes, implemented as in (24), in Mohawk the added probe makes use of the same morpheme as the core probe for 2nd person agreement (with no 2nd cycle distinction), and a distinct morpheme *wa*- for 3rd person agreement.[11]

---

[10] The distribution of the added probe with respect to ergative displacement in Basque is discussed in Rezac 2003, where it is analyzed as the π-probe of T. This illustrates an alternative way of construing the added probe: it could be taken as a separate AGR head above the EA, and limited to inverse contexts by being allowed to Agree with the EA only when the core probe on *v* has not Agreed with the EA, as a general consequence of deactivation under Agree (see op. cit. for discussion).

However, there are problems with the "Active Goal Hypothesis" on which this would need to crucially rely (Carstens 2001); more gravely, in direct contexts the high probe would never Agree with any goal at all, violating the Match Requirement of (14). It would also be predicted to be an accident that the second cycle core probe and an added probe receive the same spell-out as in Dakota. Below we pursue a unification of the added probe and spurious oblique Case strategy (3.4), which requires conceiving of both as reflex of the core probe. The alternative of this footnote is worth mentioning; fundamentally, the cyclic Agree mechanism has as large a role to play in accounting for it as in the hypothesis pursued in the text.

[11] Mohawk *wa* also shows up in unaccusative-type intransitives, whose inflection is identical to 3→X combinations. It is not the spell-out of a position unlike *aa* in Nishnaabemwin, because of its mobile placement with respect to the

Algonquian provides us with an even more striking illustration of the range of possible variation with respect to the spell out of an added probe:[12, 13]

Table 8: Singular paradigm for Nishnaabemwin

EA→IA	2	1	3
2	—	g-*see*-i ‡ 2-*see*-DFLT.1 'You see me.'	g-*see*-aa 2-*see*-DFLT 'You see him.'
1	g-*see*-in 2-*see*-1.INV 'I see you.'	—	n-*see*-aa 1-*see*-DFLT 'I see him.'
3	g-*see*-ig 2-*see*-3.INV 'He sees you.'	n-*see*-ig 1-*see*-3.INV 'He sees me.'	w-*see*-igw-n 3-*see*-3.INV-OBV 'That sees this.'

The Nishnaabemwin verb has multiple agreement slots. The core person agreement is realized on the (unique) prefix. In direct environments this prefix cross-references person features of the EA (second cycle agreement), whereas in inverse environments it cross-references the IA. This morphology does not show second-cycle effects, for e.g. 1st person agreement is realized as *n-* whether it be from 1st or or 2nd cycle Agree.

Our real interest here is the suffix position closest to the root, which is known in the Algonquian literature as the *theme suffix* (underlined in Table 8). In our analysis, this position is the morphological realization of the second projection of *v*, and it is in this position that the added probe in Algonquian is spelled out. Spell out of this projection is obligatory for transitive verbs in Algonquian, independently of Agree. We suggest that the prefixal agreement and the theme suffix are fundamentally different kinds of morphemes. Instructions to PF for the prefix have a 'mobile' source, originating either on v_I or v_{II} depending on whether the probe was

core probe spell-out and its incompatibility with direct context second cycle, both contrasting with *aa*. So it seems to spell out specifically an added probe. This would indicate that what we are here calling unaccusatives actually have a default EA in Mohawk, perhaps as in a derived causative analysis of unaccusatives (Davis and Demirdache 2000); cf. Baker 1996:201f. for evidence that at least some Mohawk unaccusatives are derived from transitives, comparing *se/si* unaccusatives in Romance.

[12] Certain details of the Nishnaabemwin paradigm are deliberately not addressed here, but we think their omission is tangential to our argument. First of all, the language in fact as a 4-way system of person contrasts, not just a 3-way system: in addition to there being 1st and 2nd person categories, there is also a 1st person inclusive category. Nothing in the analysis changes as a result of this, but an exhaustive analysis of the Algonquian person system requires a more complete decomposition of the person feature geometry (see Harley and Ritter 2002 for specific proposals). That the 1st person inclusive has properties of both 1st and 2nd person is clear. It cannot co-occur with another 1st or 2nd person argument (just as 1→1 and 2→2 cannot co-occur). Furthermore, the 1st person has the peculiar property (well-known in the Algonquian literature) of behaving like a 2nd person with respect to person agreement, but like a 1st person with respect to number agreement (like many languages, the combination of [1.PL] in Algonquian has special properties). This asymmetry is beyond the scope of our treatment, and not clearly relevant. For our purposes, what is relevant is that the 21 category behaves as we predict a category with 2nd person features should do: it fully deactivates a 3-1-2 probe.

[13] We note that the 1→2 forms (marked by ‡) show portmanteau morphology distinct from regular 1 (or 2) morpheme in exactly the same location as Dakota as discussed in 2.3, confirming the systematic nature of the phenomenon.

deactivated on the first or second cycle. Thus, the prefix does not have a structural locus as the spell out of a syntactic head as such. We take it to simply be the spell out of a valued probe.[14]

The theme suffix on the other hand seems to be fundamentally the spell out of a head (for us v_{II}), in that its spell out can be tied to just one syntactic position. This distinction between spelling out φ-features and spelling out a head is one that is independently motivated on empirical grounds. In an extensive typological study of verbal morphology, Julien (2002) finds that whereas one can observe striking cross-linguistic generalizations correlating the syntactic positions of functional heads in the verbal complex with the positions in which they are spelled out, no such generalizations are possible about agreement morphology. This strongly suggests that the spell out of agreement is typically not discharging (in Noyer's (1992) sense) a head, but merely discharging the φ-features on a head (compare Distributed Morphology's *fission* operation).

Despite being fundamentally different from one another, spell out of the v_{II} head interacts with spell out of the core probe in crucial ways. In direct contexts this head is the locus of a probe by virtue of second-cycle agreement. In inverse contexts this head is the locus of a probe by virtue of the added probe. Direct and inverse derivations are given in (29) and (30). Shaded parts of (29) and (30) show where core agreement features are discharged. Depending on whether or not the core agreement morphology discharges the probe on v_{II}, φ-features on this head may or may not be visible when it comes to spell out of the head itself. (Clearly this approach commits us to saying that the core probe must be discharged earlier than other properties of the head; the validity of this assumption is beyond our present scope.) For each derivation, the remnant of v_{II} after vocabulary insertion of the core agreement morphology is indicated in the unshaded box. With the mechanics in place, generalizations as to the form of the theme suffix start to emerge. The default *–aa* suffix is realized whenever core agreement morphology discharges the phi-features on v_{II} leaving a bare head to be spelled out. The exception to this is (29)c, but this we have already discussed; it falls into the class of special portmanteau morphology in 2→1 direct contexts and, like the other cases we have seen of this, we take it to be reducible to contextual allomorphy (a special form in the context of marked agreement on v_I). In the inverse contexts (30) the core probe is discharged on v_I, which means that the added probe on v_{II} is not discharged independently of the head itself, and so it is able to affect vocabulary insertion at v_{II}, as it in fact does: v_{II} is spelled out as *–igw* when it has an undischarged 3 probe (30)a-(30)b, and *–in* when it has an undischarged 3-1 probe (30)c. Thus the theme suffix in inverse contexts directly spells out the added probe.

(29) Direct contexts

a.	$\boxed{v_{II}}$	EA	v_I		IA	1→3=1
	(3)	3	3	—	3	Theme -*aa*
	1 —	1	1			
	~~2~~		2			
b.	$\boxed{v_{II}}$	EA	v_I		IA	2→3=2
	(3)	3	3	—	3	Theme -*aa*
	1 —	1	1			
	2 —	2	2			

[14] A priori one might expect partial agreement on v_I in direct environments to receive some morphological spell out, but empirically it clearly does not. This suggests that vocabulary insertion is selective in that it only discharges checked/valued features of the probe *on the projection where the probe is deactivated*, be that v_I or v_{II}.

c.

v_{II}	EA	v_I	IA	2→1=2
(3)	3	3 —	3	Theme -*i*
(1)	1	1 —	1	
2 — 2		2		

(30) Inverse contexts (bold-type indicates added probe)

a.

v_{II}	EA	v_I	IA	3→1=1
3 —	3	3 —	3	Theme -*igw*
		1 —	1	
		~~2~~		

b.

v_{II}	EA	v_I	IA	3→2=2
3 —	3	3 —	3	Theme -*igw*
		1 —	1	
		2 —	2	

c.

v_{II}	EA	v_I	IA	1→2=2
3 —	3	3 —	3	Theme -*in*
1 —	1	1 —	1	
		2 —	2	

In summary, we have seen three examples of languages with added probes. The morphology of the phenomenon varies considerably but is in all cases bounded by derivational properties imposed by cyclic Agree. In all three languages the added probe is a derivational strategy that appears only in inverse contexts, which cyclic Agree characterizes, and in them it avoids a violation of the PLC by Agreeing with the EA. It is added cyclically after the Agree cycle between the core probe and the IA, and on the same projection, v_{II}, from which second cycle of the core probe takes place in direct contexts. We return further to the inverse context restriction in 3.4; first, we examine the other PLC repair strategy.

3.3 R-Case

In Kashmiri inverse contexts are special because the IA gets a superficially inherent/oblique Case, distinct from the Case it gets in direct contexts, and furthermore it does not control the π-probe as cyclic Agree suggests. We call the special Case R-Case, R for *repair*. Like the added probe, R-Case assignment is a derivational strategy limited to inverse contexts, and as such it cannot be inherent. We will propose that R-Case is assigned by a probe as a reflex of Agree with the IA if the probe has a morphological property that allows it to not be valued; otherwise it assigns its normal Case. The property is limited to inverse contexts, to which we return in 3.4.

Kashmiri (Indo-European, Wali and Koul 1997) has a nominative-accusative system in the present, and ergative-nominative in the past. Table 9 gives the Cases of the core arguments, with the argument that controls core agreement underlined; the core agreement slot is also underlined in the examples. There are both unergatives whose sole overt argument behaves like the EA of transitives, and unaccusatives whose sole argument behaves like the IA of transitives.[15]

[15] Beside the core agreement slot tracking the underlined argument, there are two clitic doubling series, one of which tracks ergative subject and the unmarked object (that is, the non-agreeing core argument of the clause), which we gloss E/A, and one the agreeing nominative subject, which we gloss N; there is also an independent dative clitic series, glossed D.

Table 9: Kashmiri Case/agreement system

	EA	IA	Dative
Present/future	Unmarked, N-clitic	Unmarked, E/A-clitic	Dative, D-clitic
Past/perfective	Ergative, E/A-clitic	Unmarked, N-clitic	Dative, D-clitic

The ergative behaves fully as an inherent Case in Kashmiri (Woolford 1997, Nichols 2001): it is restricted to and obligatory on EAs in the past/perfective, and it does not interfere with assignment of the nominative (identified as the unmarked Case tracked by the N-clitic series) or with core agreement (tracking the nominative). The inherent dative has similarly fully inherent properties; it is assigned for example to the goal argument of ditransitives and it remains under passivization:

(31) Mohnas a:yi kəmi:z aslamni zəriyi dinɨ.
Mohan.M-D pass.F.SG shirt.F Aslam.M by give
'Mohan was given the shirt by Aslam.' (Wali and Koul 1997:154)

Table 9 omits the fact that there is a Case morphologically identical to the dative that is assigned to the IA in the present tense in all and only inverse contexts, as determined with respect to a 3-2-1 rather than 3-2 probe:

(32) a. bɨ chu-s-ath tsɨ parɨna:va:n. 1 > 2, *direct*
I.N be.M.SG-1.SG.N-2.SG.E/A you.N teaching
'I am teaching you.' (Wali and Koul 1997:155)
b. tsɨ chu-kh me parɨna:va:n. 2 > 1, *inverse*
you.N be.M.SG-2.SG.N me.D teaching
'You are teaching me.' (Wali and Koul 1997:155)

This is the R-Case. As can be seen, it is tied to person interaction of the EA→IA combination rather than to a theta role. Therefore, it disappears under passivization, when the inverse context is removed. (33) shows this, as well as that the R-Case dative is not affected by the presence of an inherent dative (33)a, and that the two datives differ in that only the former disappears under passivization (33)b (Wali and Koul 1997:208-9).

(33) a. su kari-y tse me hava:lɨ. 3 > 2, *inverse*
he.N do.FUT-2.SG.D you.D me.D handover
'He will hand you over to me.' (Wali and Koul 1997:208)
b. tsɨ yi-kh me hava:lɨ karnɨ təm'sɨndi dəs'
you.N come.FUT-2.SG.N me.D handover do.INF.ABL he.GEN by
'You will be handed over to me by him.' (Wali and Koul 1997:208)

Similarly, because the ergative is an inherent Case, the EA of the past tense is never accessible to φ-Agree, and never figures into establishing inverse contexts (Nichols 2001). Therefore, the IA always gets the nominative and agrees:

(34) a. tsɨ vuch-a-kh me
you.N see-2.SG-2.SG.N me.D
'You will see me.' (Wali and Koul 1997:156)
b. tse vuch-u-th-as bɨ
you.E saw-M.SG-2.SG.E-1.SG.N me.N
'You saw me.' (Wali and Koul 1997:156)

In both direct and inverse configurations, the core agreement slot of Kashmiri tracks the EA. Thus the pattern in Table 10 emerges, where the left-hand reflects the π-value of core agreement (always the EA), and the right-hand is the Case assigned to the IA: ∅ in direct contexts and when there is no other non-oblique, R-Case in inverse contexts, where direct/inverse are those characterized by a 3-2-1 probe.

Table 10: Kashmiri Person Hierarchy effect

EA→IA	1	2	3
1		1/∅	1/∅
2	2/D		2/∅
3	3/D	3/D	3/D

The Kashmiri pattern is important to us for three reasons. First, it employs the direct/inverse patterning determined by cyclic Agree. Second, it uses a novel strategy to avoid PLC violations, which is intuitively consistent with the PLC: somehow the IA is not Agreed with, reflected in R-Case assignment, and the core probe is available for the EA. Third, the pattern is so far not predicted by the theory of cyclic Agree, which predicts Agree with the IA in inverse contexts.

We propose that R-Case is assigned when a probe matches a goal, and is thus structural, but (due to a property of the probe) valuation does not result. Since identity of Case assigned is determined by the identity of the probe, R-Case can be different from the normal Case of a probe: it is assigned under Agree-without-value.

In the rest of this section, we consider the R-Case phenomenon, delimited as the assignment of an atypical Case to a DP by a probe [F-] when it "needs" for convergence to Agree with another DP, and so "needs" not be valued/deactivated by the first DP. We bring support for the phenomenon from 2-to-3 retreat in K'ekchi, and then from Bobaljik and Branigan's (forthcoming) analysis of Romance causativization, which posits the equivalent of R-Case. In 3.4 we return to the restriction to inverse contexts.

K'ekchi (Mayan, Berinstein 1985, 1990) is an ergative-absolutive VOS language. Only ergatives and absolutives trigger agreement, by the form of which they are distinguished as ergative vs. absolutive. There are no person hierarchy effects: ergative and absolutive agreement morphology separately track their respective arguments. Dative arguments are morphologically realized as possessors, using ergative agreement morphology, of so-called relational nouns, which may equally be viewed as inflected adpositions or Case markers, and are glossed as such; for other oblique arguments the possessed relational noun is construed as the argument of an independent adposition.

We propose that R-Case arises in K'ekchi in what Berinstein (1990) analyzes in Relational Grammar terms as 2-3 Retreat (cf. Davies and Sam Colop 1990, Hale 2001). In K'ekchi, an EA cannot undergo Ā-extraction to the preverbal position when tracked by regular ergative agreement; instead, it must be tracked by absolutive agreement. When this occurs there is no ergative agreement, and the IA becomes dative:

(35) a. T-at-x-lop laat/*acu-e li c'anti'.
TNS-2.A-3.E-bite you/*you-D the snake
'The snake will bite you.' (Berinstein 1990:10)
b. Li c'anti' ta-∅-lop-o-k acu-e.
the snake TNS-3.A-bite-AP-ASP 2.E-D
'That's the snake that will bite you.' (Berinstein 1990:10)

We analyze this as a requirement for Ā-extraction to proceed from a non-derived position (cf. Rizzi 1982:chapter 4, 2004). Ergative agreement, which blocks Ā-extraction of EA, must then be a property of T (Bobaljik 1993), and force externalization of the EA. When the external argument is to Ā-extract, the ergative φ-Probe of T must be absent, and the external argument remain in [Spec, *v*P]. There it is Agreed with by *v*'s φ-Probe, and thus coded by absolutive agreement. Since the φ-Probe of K'ekchi is flat (there is no agreement displacement), the only way for it to be available for Agree with the EA is if it does not agree with the IA. This approach, which is a reformulation of the basic insights of Berinstein's 2-3 Retreat analysis, accounts for the core properties of the construction: the EA is tracked by absolutive agreement; the IA is not tracked by absolutive agreement; EA extraction requires this structure.[16] The presently interesting fact is that the IA is further realized as a dative, a Case that is not normally available to IA's outside of this construction (35)a, but required here, (35)b.

Berinstein (1990:27-31) shows that dative IAs in 2-3 Retreat behave as if they had structural rather than inherent Case. In K'ekchi', the ergative and absolutive are almost uniquely isolated as potential Controllers into adjuncts, (36)a, against both various types of obliques and the dative goals of ditransitives (36)b. However, the dative IAs of 2-3 Retreat clauses can Control (36)c.[17]

[16] There are two further properties of the construction: the presence of antipassive morphology (glossed AP) on the verb; and the requirement that the external argument undergo Ā-extraction. As Davies and Sam Colop's (1990, esp. p. 540-1, 526 nt. 10) treatment of and comparison with the related Mayan language K'iche' shows, both properties are independent of the crucial constellation of facts: EA getting absolutive forces the IA to get a separate Case (modulo the possibility, available in K'iche' but not K'ekchi, of IA not agreeing if 3rd person, cf. Hale 2001). Ā-extraction and the antipassive morphology show up independently in other constructions where there is not both 2→3 retreat and EA tracked by absolutive agreement.

More investigation is needed to clarify the limitation of the 2→3 Retreat to Ā-extraction of the EA in K'ekchi, but a plausible answer is that it is ordinarily blocked because some property of T is not being satisfied. Ordinarily, the EA would raise to [Spec, TP] to satisfy the EPP under Agree. If 2-to-3 retreat applied without Ā-extraction, the EA will have Agreed with *v* by the time the φ-probe of T seeks it; that renders it unavailable to further Agree, and thus for raising to [Spec, TP]. The EPP is not satisfied. If however, the EA is Ā-extracted, we may suppose that it satisfies the EPP regardless of Agree with T, because such EPP-of-T satisfaction under Ā-movement is well known, whether because extraction proceeds through [Spec, TP] (Spanish, Zubizarretta 1998), or because if [Spec, CP] is filled the EPP of T is satisfied parasitically (see Alexiadou and Anagnostopoulou 1998:526, note 43).

[17] This is an independent issue. Since the Control involved is into adjuncts, ergative and absolutive are plausibly isolated as those that attain a high structural position, either through φ-agreement alone or by agreement-driven A-movement (see Cardinaletti 1997, Chomsky 1995:272f.); alternatively, it is the extra PP-like structure introduced by inherent Case that blocks such DPs from controlling.

(36) a. Ch-o-a-tenk'a chi cuanc sa' x-yal-al.
TNS-1.AP-2.E-help in.order exist-INF in its-truth-POSS
'May you help us in order to live in peace.' (Berinstein 1990:27)
b. *X-∅-in-q'ue acu-e_i chi PRO_i cua'ac.
TNS-3.A-1.E-give 2.E-D in.order eat.INF
Intended: 'I gave it to you for the purpose of eating.' (Berinstein 1990:28)
c. Lain x-in-takla-n acu-e chi c'anjelac.
I TNS-1.A-send-AP 2.E-D in.order work.INF
'I'm the one who sent you for the purpose of working.' (Berinstein 1990:28)

The behavior of the dative in 2-3 Retreat fits into the theory of R-dative we have proposed for Kashmiri: it enters into a cycle of Agree with the absolutive π-Probe, but it receives R-Case without affecting the probe, which remains available for the EA to ensure a convergent derivation. The IA therefore surfaces as a non-agreeing dative, but one which has entered into Agree with *v*.[18]

Bobaljik and Branigan (forthcoming) arrive at a mechanism very similar to R-Case from independent data, of which we focus on datives in French causatives. These instantiate a cross-linguistically very common pattern of causativization (Baker 1988:chapter 4): the causativization of an intransitive treats its sole argument as the direct object of the causative verb; causativization of a transitive treats is direct object as direct object and its subject as indirect argument of the causative verb. (37) shows this for French: the sole argument of *venir* 'come' in (37)a can only be accusative, as the direct object of *acheter* 'buy' in (37)b, while the subject of *acheter* must be dative (formed with *à* 'to'). Bobaljik and Branigan show that the dative here is structural.

(37) a. Luc a fait venir **les/*aux étudiants**
Luc-N has made to.come the/*to.the students-A
'Luc has made the students work.'
b. Luc a fait t_1 acheter un livre **aux/*les étudiants$_1$**
Luc-N has made to.buy a book-A to.the/*the students
'Luc has made the students work.'

Bobaljik and Branigan propose a multiple-Case checking mechanics in configurations where a single head F^0 must license two DPs: the first DP gets a special Case $K_{F\text{-}R}$ from its relationship to F^0, our R-Case, which leaves F^0 free to assign its regular Case K_F to a farther DP (with obligatory evacuation of DP_1 from the path between F^0 and DP_2 because of locality, cf. Chomsky 2000:131):

(38) a. F^0 DP_1-$K_{F\text{-}R}$ DP_2-K_F b. F^0 DP-K_F

There is thus a convergence of evidence that an Agree relation between a φ-probe and a DP can, because of some property of the φ-probe, involve only match and not valuation, which

[18] That Agree cycle has perhaps caused movement to a high position, since Berinstein (1990:14) also shows that 2-3 Retreat datives unlike all other datives are necessarily immediately post-verbal.

allows Agree of that φ-probe with a different DP; assignment of a special Case by the probe to the IA occurs when this happens. Other properties of the language determine when this is required, but it seems to happen only as last-resort to allow a derivation to converge. In the following section we investigate the limitation of R-Case and the coextensive added probe strategy to inverse contexts, and argue that they are in fact generally available options that only appear to be restricted to inverse contexts because they simply crash the derivation in direct contexts.

3.4 Derivation of PLC-licensing strategies

Both the added probe and the R-Case strategies of avoiding a PLC violation by the EA are derivational strategies coextensive with inverse contexts, which suggests a unified distribution despite their superficial differences. This is supported if both are consequent on some property of the core probe; this is evident for R-Case because Case assignment to the IA is a function of Agree with the core probe, and may be supposed for the added probe because it is on a projection of the same head as that which hosts the core probe. We propose therefore that the two strategies reflect the same fundamental option, which provides much-needed boundary conditions on understanding the phenomena.

The boundary conditions on the mechanism are the following: (i) the EA must enter into Agree, either by (i-a) an added agreement slot distinct from that of the core probe (Dakota, Mohawk, Algonquian), or (i-b) by the core agreement slot of a language, which becomes anomalously valued form the EA but nevertheless licenses the IA (Kashmiri); (ii) under option (i-b), the IA must receive a Case different from the Case it would receive in direct contexts; (iii) the mechanism must occur only in inverse contexts. We assume local determinability at each point in a cyclic derivation, and for us Agree between the core probe on *v* and the IA is crucially on an earlier cycle than Agree with the EA. This severely constraints our analytic options, because the application of (ii), which occurs at the point where the core probe only has the IA in its search-space, must be so restricted that it succeeds only in inverse contexts (iii), without making reference to properties of the yet un-Merged EA.

We characterize the mechanism behind both strategies as the availability of a morphological property P to the core probe. The core probe may or may not have P; its presence is manifested as added probe or R-Case.[19]

(39) P: If the core probe on AGR (*v*) Agrees for a feature specification F, a new probe is added to AGR, whose articulation is identical to the articulation of F.

[19] We leave open the question whether the availability of P itself needs to be parametrized. All the 3-2-1 probe languages in our data have P available (as added probe or R-Case). Of the 3-2 languages, Basque most clearly does so, with the situation unclear for the EA in Georgian, Karok, and Erza Mordvinian. Under idealization, we might suppose P is generally available to non-flat probes, regardless of spell-out.

For flat probes, we largely exceed the limits of this work in addressing this question. Observe, however, that a flat probe language will never have the core probe Agree with the EA, so it always needs an added probe or a separate higher probe, because all contexts are inverse contexts. Given that, it is an interesting possibility that what is normally taken as the π-probe of T in languages like English is in fact due to P. In particular, we might take the accusative in English to be R-Case, and the core probe of *v* to always Agree with the EA after assigning R-Case to the IA, because derivations where it Agrees with the IA always violate the PLC from the EA. If so, P might be universally available to a probe.

> Parametrization of P: P is available always (Dakota) or only when the core probe matches without valuation (Kashmiri)

One might think of P as a blank probe adjoined to the core probe, which is only activated by Agree of the core probe. The activation specifies the articulation of the blank probe: Agree of a 3-2-1 probe with 1st person (3-2-1) adds a probe of a 3-2-1 articulation, with 2nd person (3-2) adds one with a 3-2 articulation, and with 3rd person (3) one with a 3 articulation. Addition of a probe by P to the core probe on a projection of *v*, say v_{I}, creates a modification of v_{I}, and therefore requires insertion of the added probe on the next higher projection, v_{II}, since once inserted into the structure a term may not be modified without projecting (Rezac 2002, cf. Chomsky 2000:126). An added probe (only in inverse) and a second cycle of the core probe (only in direct) thus have this similarity: they are both probes from v_{II} standing in some relation to an earlier probe from v_{I}. For an added probe the relation is that it is an Agree-determined doubling of the core probe, while a second cycle probe is the remaining unchecked segments of the core probe.

The parametrization of P is the locus of difference between added probe and R-Case language, turning on the idea that R-Case is Agree without value. Intuitively, the difference between Kashmiri and Dakota is the relative "weakness" of the core probe in Kashmiri: if P is present and a probe is added, the core probe must not have been valued, because such double probes are not tolerated.

Consider now how desideratum (iii), limitation to inverse contexts, can be derived without look-ahead from the *v*-IA Agree step to properties of the EA. For constructions without an EA, the answer is clear: if the core probe has P and a probe is added, the added probe never finds a match, violating the Match Requirement (14). The logic of what rules the probe out in direct contexts when there is an EA will be similar: although the added probe will Agree with the EA, it will violate a principle. The locus of the problem is at step 4 of the direct/inverse derivations in Table 11 for a probe with P. The derivations focus on added probe languages, so that Agree-without-value is not distinguished (in an R-Case language it occurs at step 1 instead of Agree). Added probes in Table 11 are boxed. The still-active segments of a probe at any point in the derivation are in bold, these being the ones at a particular locus that had not matched on a previous cycle.

Table 11: Derivation of added probe (and R-Case)

1→2 (direct)	2→1 (inverse)
Step 1: Agree with IA	
v_I IA 3 — 3 2 — 2 **1**	v_I IA 3 — 3 2 — 2 1 — 1
Step 2: Add a probe following (39)	
v_I IA $\boxed{\mathbf{3}}$ 3 — 3 $\boxed{\mathbf{2}}$ 2 — 2 **1**	v_I IA $\boxed{\mathbf{3}}$ 3 — 3 $\boxed{\mathbf{2}}$ 2 — 2 $\boxed{\mathbf{1}}$ 1 — 1
Step 3: Add EA and project v	
Step 4: Agree with EA from v_{II}	
v_{II} EA v_I IA 3 $\boxed{\mathbf{3}}$ — 3 3 3 — 3 2 $\boxed{\mathbf{2}}$ — 2 2 2 — 2 **1** —— 1 1	v_{II} EA v_I IA 3 $\boxed{\mathbf{3}}$ — 3 3 3 — 3 2 $\boxed{\mathbf{2}}$ — 2 2 2 — 2 1 $\boxed{\mathbf{1}}$ — 1 1 1 — 1

At step 4, the inverse context converges without further comment: the core (original) and added probe both Agree. However, when a probe is added in a direct context, there will always be a segment (feature) of the original probe that has not Agreed with the IA and corresponds to a segment on the EA; moreover, as consequence of (39), the added probe itself has no corresponding segment. This segment on our approach must enter into regular second cycle Agree with the EA, because there is no intervener and the EA falls into its search-space upon projection to v_{II}.

It is here that we localize what goes wrong. We propose that such Agree configurations are banned because individual segments of the π-structure of the EA enter into Agree with segments of different probes, and thus the π-structure as a whole enters into a multiple dependency for PLC-licensing (Case Filter). Such multiple dependencies seem generally banned, as movement of an NP from one Case position to another, or of a *wh*-word from one Q-checking (scope-taking) position to another:[20]

(40) *What did Azenor wonder t_1 Nolwenn saw t_1?

The unified treatment of the added probe and R-Case strategies proposed in this section goes in some ways beyond the data. R-Case suggests that the strategy is determined locally at the point the core probe Agrees with the IA, and therefore before the EA is added; the identical distribution of R-Case and added probe suggests the two mechanisms should be unified as fully as possible. We do not quite reduce one to the other; while in both a probe is added, languages

[20] In Chomsky 2000:122f., 128, the ban is implemented in terms of deactivation of a feature-set upon Agree (viewed as Case assignment in for φ-features). This is can be implemented in the present situation as follows: if the still-active 1-segment of the core probe Agrees with the EA from v_{II} and deactivates its π-set, the added probe has no match, leading to a violation of the Match Requirement (14) in exactly the same way as Lasnik's (1999) use of it to rule out **There seems to someone that S*.

that have R-Case do not have the core probe valued under Agree, unlike those without R-Case. This reflects the intuition that R-Case is assignment of Case under Agree-without-value, something evidently true for Kashmiri but not for added probe languages. Stepping back from the details, the theory of cyclic Agree is crucial in characterizing the class of inverse contexts where these strategies occur, and the Person Licensing Condition, perhaps a statement of the Case Filter, identifies what goes wrong there and the nature of the repair strategies.

4 Conclusion

In this paper we have displayed a pattern of PH-driven agreement displacement, and argued that it follows from simple, conceptually elegant assumptions about the mechanics of the derivation, (41)a, (41)b, combined with an independently motivated understanding of φ-features, (41)c:

(41)
a. Intervener-based locality (Rizzi 1990), relativized to features (Chomsky 1995).
b. A fine-grained approach to cyclicity, where every Merge/Agree operation defines a potential cycle, and where Merge can thus change the search-space for a probe after projection (Rezac 2002, 2003).
c. A fine-grained approach to φ-features, and specifically π-features, associating with each person value a different feature structure and thus a different locality class (Béjar 2003).

Once the core π-probe of a language is placed between EA and IA, PH-driven agreement displacement patterns emerge from this system:

(42)
a. Partial agreement sensitive to specifications of goal, which is prerequisite for Person Hierarchy phenomena, arises from interaction between the articulation of the characteristic probe of the language (giving PH-sensitivity, with a 3 probe having no sensitivity, 3-2 distinguishing $1^{st}/2^{nd}$ person from 3^{rd} person, and 3-2-1 all persons) with person specification of the goal.
b. The design of the agreement displacement pattern, going from IA (preferred) to EA: this follows from cyclic expansion, determined by a bottom-up derivational mechanics where the EA is added latter than the IA.
c. The same prioritization of the EA characterizes a natural class of computations where the EA does not Agree with the core probe, the inverse contexts, which map into an empirically distinguished class. The Person Licensing Condition identifies the special character of these derivations as those where the EA is not licensed as such, and the abstract nature of the repair strategies.

These results are based on, and support, a syntactic treatment of a class of PH-effects: they result from the mechanics of the formation of featural syntactic dependencies in a cyclic derivation.[21]

[21] Quite different support for a syntactic treatment is given by Bruening (2001:chapter 2), who shows that the inverse in Passamaquoddy (Algonquian) allows the IA to take scope over the EA, which is not possible in the direct. Although we have not discussed this class of phenomena here, they are easy to model once the Agree-Merge interaction that results in movement (Chomsky 2000:101-2) is spelled out, as in Rezac 2004:chapter 3. In

References

Abondolo, Daniel. 1982. Verb paradigm in Erza Mordvinian. *Folia Slavica* 5:11-24.
Albizu, Pablo. 1997. Generalized Person-Case Constraint: A case for a syntax-driven inflectional morphology. In *Theoretical issues on the morphology-syntax interface*, ASJU XL, ed. Miryam Uribe-Etxebarria and Amaya Mendikoetxea 1-33. Donostia: Gipuzkoako Foru Aldundia/EHU.
Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing AGR: Word-order, V-movement, and EPP-checking. *Natural Language and Linguistic Theory* 16:491-539.
Allen, Barbara J., and Donald G. Frantz. 1978. Verb agreement in Southern Tiwa. In *Proceedings of the Fourth Meeting of the Berkeley Linguistic Society*, 11-17. University of California, Berkeley.
Allen, Barbara J. and Donald G. Frantz. 1983. Advancements and verb agreement in Southern Tiwa. In *Studies in Relational Grammar 1*, ed. D. M. Perlmutter, 303-316. Chicago: University of Chicago Press.
Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: Evidence from clitics*. The Hague: Mouton de Gruyter.
Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
Baker, Mark. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.
Beatty, John. 1974. *Mohawk Morphology*. Greenley, Col.: Museum of Anthropology, University of Northern Colorado.
Béjar, Susana, and Milan Rezac. 2003. Person licensing and the derivation of PCC effects. In *Romance linguistics: Theory and acquisition*, ed. Ana-Teresa Pérez-Leroux and Yves Roberge, 49-62. Amsterdam: John Benjamins.
Béjar, Susana. 2000. Structural markedness in formal features: Deriving interpretability. *Revue québecoise de linguistique* 28:47-72.
Béjar, Susana. 2003. Phi-syntax: A theory of agreement. Doctoral dissertation, University of Toronto, Toronto, Ont.
Berinstein, Ava. 1985. *Evidence for multiattachment in K'ekchi Mayan*. New York: Garland.
Berinstein, Ava. 1990. On distinguishing surface datives in K'ekchi. In *Studies in Relational Grammar 3*, ed. Paul M. Postal and Brian D. Joseph, 3-48. Chicago: University of Chicago Press.
Bobaljik, Jonathan David. 1993. Ergativity and ergative unergatives. In *MIT Working Papers in Linguistics 19*, 45-88. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
Bobaljik, Jonathan David. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral dissertation, MIT, Cambridge, Mass.

movement Merge base-generates a non-thematic object in the specifier of a head that has Agreed with a goal in its c-command domain, so the identity of the Merged object must be constrained by the result of Agree. In our system where *v* enters potentially into Agree with multiple goals, suppose it is the last which determines the identity of a possible goal for re-Merge: EA in direct contexts, IA in inverse contexts. Movement of *v* to T, which provides the EPP position for non-thematic Merge, will then allow only the EA to be re-Merged (moved) in direct contexts, and only the IA in inverse contexts, giving ambiguous EA >> IA (base-generated) and IA >> EA (moved) scope relations only in the inverse.

Bobaljik, Jonathan David, and Phil Branigan. Forthcoming. Eccentric agreement and multiple Case checking. In *Ergativity: Emerging issues*, ed. Alana Johns, Diane Massam., and Juvenal. Ndayiragije. Dordrecht: Kluwer.

Bobaljik, Jonathan David, and Susanne Wurmbrand. 1999. Notes on agreement in Itelmen. *Linguistic Discovery* 1 (on-line).

Boeckx, Cedric. 2000. Quirky agreement. *Studia Linguistica* 54:354-380.

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

Boškovic, Željko. 1997. *The syntax of nonfinite complementation*. Cambridge, Mass.: MIT Press.

Branigan, Phil, and Marguerite MacKenzie. 2001. Altruism, Ā-movement, and object agreement in Innu-aimûn. *Linguistic Inquiry* 33:385-407.

Bright, William. 1957. *The Karok Language*. Berkeley, Calif.: University of California Press.

Bruening, Benjamin. 2001. Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy. Doctoral dissertation, MIT, Cambridge, Mass.

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

Cardinaletti, Anna. 1997. Agreement and control in expletive constructions. *Linguistic Inquiry* 28:521-533.

Cardinaletti, Anna, and Michal Starke. 1999. The typology of structural deficiency. In *Clitics in the languages of Europe*, ed. Henk van Riemsdijk, 145-233. Berlin: Mouton de Gruyter.

Carstens, Vicky. 2001. Multiple agreement and Case deletion: Against (φ-)incompleteness. *Syntax* 4:147-163.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89-156. Cambridge, Mass.: MIT Press.

Collins, Christopher. 2002. Eliminating labels. In *Derivation and explanation in the minimalist program*, ed. Samuel David Epstein and T. Daniel Seely, 42-64. Oxford: Blackwell.

Davis, Henry, & Hamida Demirdache. 2000. On lexical verb meanings: Evidence from Salish. In *Events as grammatical objects*, ed. James Pustejovsky & Carol Tenny. Stanford, Calif.: CSLI.

Davies, William D., and Luis Enrique Sam-Colop. 1990. K'iche' and the structure of antipassive. *Language* 66:522-549.

Delancy, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57:626-657.

Diesing, Molly. 1992. *Indefinites*. Cambridge, Mass.: MIT Press.

Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Fabri, Raymond. 1996. The inverse morphology of Plains Cree (Algonquian). In *Yearbook of morphology 1995*, ed. Geert Booij and Jaap van Marle, 17-41. Dordrecht: Kluwer.

Frampton, John, and Sam Gutmann. 1999. Cyclic computation, a computationally efficient minimalist syntax. *Syntax* 2:1-27.

Gerdts, Donna. 1988. A nominal hierarchy in Halkomelem clausal organization. *Anthropological Linguistics* 30:20-36.

Hale, Kenneth. 2001. Eccentric agreement. In *On Case and agreement*, ed. Pablo Albizu and Beatriz Fernández, 15-48. Bilbo: Euskal Herriko Unibertsitatea.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger,* ed. Kenneth Hale and Samuel Jay Keyser, 111-176. Cambridge, Mass.: MIT Press.

Harley, Heidi, and Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78:482-526.

Harris, Alice. 1981. *Georgian syntax: A study in Relational Grammar*. Cambridge: Cambridge University Press.

Hewitt, B.G. 1995. *Georgian: A structural reference grammar*. Amsterdam: John Benjamins.

Hrafnbjargarson, Gunnar Hrafn. 2001. An optimality theory analysis of agreement in Icelandic DAT-NOM constructions. *Working Papers in Scandinavian Syntax* 68:15-47.

Jelinek, Eloise. 1993. Ergative "splits" and argument type. In *MIT working papers in linguistics 18*, 15-42. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Jelinek, Eloise, and Richard Demers. 1983. An agent hierarchy and voice in some Coast Salish languages. *International Journal of American Linguistics* 49:167-185.

Johns, Alana. 1993. Symmetry in Labrador Inuttut. In *MIT working papers in linguistics 18*, 43-57. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.

Julien, Marit. 2002. *Syntactic heads and word formation: A study of verbal inflection*. New York: Oxford University Press.

Laka, Itziar. 1993. The structure of inflection: A case study in X^0 syntax. In *Generative studies in Basque linguistics*, ed. José Ignacio Hualde and Jon Ortiz de Urbina, 21-70. Amsterdam: John Benjamins.

Lakämper, Renate, and Dieter Wunderlich. 1998. Person marking in Quechua – A constraint-based minimalist analysis. *Lingua* 105:113-148.

Lasnik, Howard. 1999. *Minimalist Analysis*. Oxford: Blackwell.

Lieber, Rochelle. 1980. The organization of the lexicon. Doctoral dissertation, MIT, Cambridge, Mass.

Lounsbury, Floyd G. *Oneida verb morphology*. New Haven, Conn.: Yale University Press.

Macaulay, Monica. 1992. Inverse marking in Karuk: The function of the suffix *–ap*. *International Journal of American Linguistics* 58:182-201.

Miner, Kenneth L. 1980. The order of Dakota person prefixes: The rest of the data. *Glossa* 14:35-42.

Nash, Léa. 1995. Portée argumental et marquage casuel dans les langues SOV et dans les langues ergatives: L'exemple du géorgien. Doctoral dissertation, Université de Paris 8.

Nichols, Lynn. 2001. The syntactic basis of referential hierarchy phenomena: Clues from languages with and without morphological case. *Lingua* 111:515-537.

Noyer, Rolf. 1997. *Features, positions and affixes in autonomous morphological structure*. New York: Garland Publishing.

Ormazabal, Javier, and Juan Romero. 2003. Agreement Restrictions. Ms.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 355-426. Cambridge, Mass: MIT Press.

Postal, Paul M. 1979. *Some syntactic rules in Mohawk*. New York: Garland.

Rezac, Milan. 2002. Cyclic domains and (multiple) specifier constructions. Ms., University of Toronto, Toronto, Ont., and University of the Basque Country, Vitoria-Gasteiz.

Rezac, Milan. 2003. The fine structure of cyclic Agree. *Syntax* 6:156-182.

Rezac, Milan. 2004. Elements of cyclic syntax: Merge and Agree. Doctoral dissertation draft, University of Toronto, Toronto, Ont.
Rice, Keren, and Leslie Saxon. 1994. The subject position in Athapaskan languages. In *MIT working papers in linguistics 22*, 173-195. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, Mass.: MIT Press.
Rizzi, Luigi. 2004. On the form of chains: Criterial positions and ECP effects. Ms., University of Siena.
Roberts, Ian. 1991. Excorporation and minimality. *Linguistic Inquiry* 22:209-218.
Schwartz, Linda J. 1979. The order of Dakota person prefixes: An evaluation of alternatives. *Glossa* 13:3-12.
Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical categories in Australian languages*, ed. R. M. W. Dixon, 112-171. Canberra: Australian Institute of Aboriginal Studies.
Sproat, Richard. 1985. On deriving the lexicon. Doctoral dissertation, MIT, Cambridge, Mass.
Travis, Lisa. 2003. Agents and causes in Malagasy and Tagalog. In *MIT Working Papers in Linguistics 44*, 355-370. Department of Linguistics and Philosophy, MIT, Cambridge, Mass.
Travis, Lisa, and Greg Lamontagne. 1992. The Case Filter and licensing of empty K. *Canadian Journal of Linguistics* 37:157-174.
Valentine, Randolph J. 2001. *Nishnaabemwin reference grammar*. Toronto: University of Toronto Press.
Wali, Kashi, and Omkar N. Koul. 1997. *Kashmiri: A cognitive-descriptive grammar*. New York: Routledge.
Wiltschko, Martina. 2003. Person hierarchies without a person hierarchy. Ms., University of British Columbia, Vancouver.
Woolford, Ellen. 1997. Four-way Case systems: Ergative, nominative, objective, and accusative. *Natural Language and Linguistic Theory* 15:181-227.
Zubizarretta, Maria-Luisa. 1998. *Prosody, focus, and word-order*. Cambridge, Mass.: MIT Press.
Zwicky, Arnold. 1977. Hierarchies of person. In *Proceedings of the Thirteenth Annual Meeting of the Chicago Linguistic Society*, 714-733. Chicago: Chicago Linguistic Society.

Susana Béjar
Department of Linguistics
1085 Dr. Penfield Avenue
Montreal, Quebec
H3A 1A7
sbejar@chass.utoronto.ca

Milan Rezac
milan_rezac@yahoo.ca