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 $\phi$-feature structures in a cyclic syntax. Cyclicity and locality derive a preference for agreement control by the internal argument. Articulation of the probe decides when the agreement controller 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 by Agree.

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

1 Introduction

The operation Agree of recent minimalist syntax establishes a syntactic dependency correlating the morphosyntactic features of one terminal with another. The most direct evidence for the dependency defined by this operation is morphological covariance of two elements, of which verb agreement is a core example. Familiar examples of verb agreement, such as the well-studied Icelandic verb-subject agreement discussed with respect to Agree in Chomsky (2000) and related work, have fostered rich research that supports a syntactic treatment deriving conditions on it from the properties of narrow-syntactic dependencies, such as locality.

There exist also complex agreement systems for which the morphological expression of agreement appears to have a more uneasy correspondence to the syntax. Here we examine one class of such agreement patterns, which can be characterized as having a single core agreement slot, for the control of which multiple arguments compete. The outcome is sensitive to the values of person features on both the candidate controllers, leading to the characterization of such systems as sensitive to person-hierarchies (PH).1 We will argue that the basic patterns arise as agreement displacement (Hale's (2001) eccentric agreement), whereby perfectly general mechanics of the syntactic derivation, namely constraints on Agree, result in an apparently non-

---

* We are grateful for comments from David Adger, Jonathan Bobaljik, Elizabeth Cowper, Daniel Currie Hall, Daniel Harbour, Alana Johns, Mélanie Jouitteau, Diane Massam, Eric Matthieu, Jochen Trommer, and audiences at McGill University, the University of Connecticut, the Workshop on Agreement at the Universidade Nova de Lisboa, and GLOW 2004 at Aristotle University of Thessaloniki. Extensive written comments from Jonathan Bobaljik and four anonymous Linguistic Inquiry reviewers have much benefited the article, and encouraging words by the editor kept us working at it. Responsibility for shortcomings rests with us. This research was supported by SSHRC fellowship #756-2003-0107 and a Tomlinson Postdoctoral Fellowship to Susana Béjar, and SSHRC fellowships #752-2000-1545 and #756-2004-0389 to Milan Rezac.

canonical agreement pattern. The fundamental principles that enter into the account are the following:

\[
\begin{align*}
(1) \quad & \text{a. Intervener-based locality (Rizzi 1990), relativized to features (Chomsky 1995): Agree for a feature [F] is only sensitive to other goals with [F].} \\
& \text{b. A fine-grained approach to cyclicity, where every syntactic operation defines a cycle and thus a potential feeding-bleeding relationship (Rezac 2003).} \\
& \text{c. A fine-grained approach to } \phi \text{-features (specifically person or } \pi \text{-features), and especially } \phi \text{-probes, associating with each person value a different feature structure and thus a different locality class (Béjar 2003).}
\end{align*}
\]

These mechanics will generate two natural classes of derivations for transitive clauses: one where the internal argument IA controls agreement, corresponding to so-called inverse contexts; and another where the external argument EA does, or direct contexts. Unlike in languages without person hierarchies, IA agreement emerges as the primary agreement relation, and EA agreement arises as agreement displacement, a pervasive empirical pattern in the paradigms we discuss. The inverse contexts thus characterized as a class of computations coincide with a set of EA-IA combinations known to disrupt core agreement patterns, for example by introducing extra agreement or special morphology. Our mechanics provide the basis for explaining the character of these disruptions: convergence requires an extra probe to Agree with the EA and so license it, which is reflected as added agreement or special case marking.

The paper is structured as follows: Section 2 introduces to the phenomenon of agreement displacement and the preference for IA as controller, illustrated with Basque, Georgian, Karok, and Erza Mordvinian. Section 3 presents the mechanics of our account of agreement displacement. We develop a proposal that spells out the interaction of person feature structures with Agree in such a way as to allow parameterization of sensitivity to person. We argue that cyclicity derives the contrasting behavior of IA and EA as potential controllers, with Agree of a person probe on \( v \) preferring the IA but allowing Agree with EA for any unvalued features. The interaction of Agree and cyclicity defines contexts where EA fails to Agree, the inverse contexts; we argue that a (Case) licensing problem independently known from the Person Case Constraint arises there and must be repaired for such derivations to converge. We illustrate the basic system with Nishnaabemwin. In section 4, we develop the details of two repair strategies found in inverse contexts: extra agreement morphology in Mohawk Nishnaabemwin, Basque, and special special IA case morphology in Kashmiri. We develop a unified treatment where both strategies result from a derivationally driven addition of a probe.

2 Person hierarchies and EA-IA interaction

Our point of departure is a class of languages in which there is a (core) agreement system whose controller cannot be characterized in terms of its grammatical function. Instead, it appears to alternate in transitives between the EA and the IA. This is illustrated in (2) for ergative displacement in Basque (Laka 1993), where underlining indicates the relevant agreement slot and its controller. The notation \( x \rightarrow y = x \) should be understood to mean that in a clause where the person (\( \pi \)) specifications of the EA are \( x \) and the IA \( y \), agreement tracks \( x \). In (2), the
underlined agreement slot marker tracks the *person* features of the EA, but elsewhere the IA (we return to Basque in section 4.2). 2

\[
\begin{align*}
\text{a. } & \text{ikusi } z\text{-in-t-u-da-n } I \rightarrow 2 = 2 & \text{b. } & \text{ikusi } n\text{-ind-u-en } 3 \rightarrow I = 1 \\
\text{seen } & \text{2-X-PL-have-1-PAST} & \text{seen } & \text{1-X-have-PAST} \\
\text{I saw you.'} & & \text{He saw me.'} \\
\text{c. } & \text{ikusi } n\text{-ind-u-zu-n } 2 \rightarrow I = 1 & \text{d. } & \text{ikusi } n\text{-u-en } I \rightarrow 3 = 1 \\
\text{seen } & \text{1-X-have-2-PAST} & \text{seen } & \text{1-have-PAST} \\
\text{You saw me.'} & \text{I saw him.'} & (\text{Basque})
\end{align*}
\]

Laka (1993) shows that neither Case marking (EA ergative, IA absolutive) nor anaphora binding (EA binding IA) patterns are affected by ergative displacement, as (3) indicates (syntactically, Basque is nominative-accusative). Of the more complicated EA-IA oscillations discussed below, work on Algonquian confirms Basque: they may but do not need to correlate with a syntactic effect such as would be produced by movement (Dahlstrom 1986, Rhodes 1994; section 5).

\[
\begin{align*}
\text{(3) } & \text{Ni-k neure buru-a ikusten n-u-en. } \\
& \text{1-E my.own head-the.A seeing 1-have-PAST} \\
\text{I saw myself.' (Basque, Laka 1993:54)}
\end{align*}
\]

The choice between EA and IA as controller is clearly sensitive to their $\pi$-specification, since all other variables remain constant. We might characterize agreement displacement as a PH effect where the controller is given by some ranking of EA, IA on the basis of their $\pi$-specification, such as $1^{\text{st}} > 2^{\text{nd}} > 3^{\text{rd}}$ person, where $\succ$ means " outranks." This would indeed be an adequate characterization of a language like Algonquian or Mohawk, where the uniquely highest of EA, IA on such a scale is the agreement controller. However, it is inadequate for Basque-type languages. Any person hierarchy will underdetermine the choice of controller in (2). This is because although a $1^{\text{st}}/2^{\text{nd}}$ person argument will always beat out a $3^{\text{rd}}$ person argument, as in (2)b and (2)d, the choice between two $1^{\text{st}}/2^{\text{nd}}$ person arguments cannot be resolved by PH: in (2)a, $2^{\text{nd}}$ person beats $1^{\text{st}}$ person, while in (2)c, $1^{\text{st}}$ person beats $2^{\text{nd}}$ person. Thus in Basque, IA agreement *bleeds* potential EA control of the agreement slot unless the IA is $3^{\text{rd}}$ person, in which case the 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.

Taking our cue from the latter phenomenon, we will propose that despite their differences, both Basque-type and Algonquian-type hierarchies arise from fundamentally identical cyclic derivations: a unique probe seeks to agree with the IA and the EA in that order. Differences in the character of the PH-effects in these languages emerge because languages have probes with distinct feature structures, giving rise to different valuation potentials, so that in Basque a probe distinguishes only $1^{\text{st}}/2^{\text{nd}}$ from $3^{\text{rd}}$ person while in Algonquian $1^{\text{st}}$, $2^{\text{nd}}$, and $3^{\text{rd}}$ person are fully differentiated. Agreement displacement from IA to EA occurs for the same reason in both

2 Glosses are: 1, 2, 3 person; SG singular, PL plural; INV inverse; OBV obviative; N nominative, E ergative, A absolutive/accusative, D dative, GEN genitive, ABL ablative; T tense, FUT future, PAST past, INF infinitive, DFLT default; X irrelevant/unclear.
languages. The IA values the probe as much as it can, and the EA ends up controlling only if it can add to the value contributed by the IA. Two determining conditions enter into our account: a principle that prefers IA to EA, which we reduce to cyclicity, and a parametrizable sensitivity of Agree to different person values of the controller, which we will reduce to the structure of π-features. The interaction between EA and IA control emerges from serial locally-determinable valuation of the probe from by closest DP at each step, with no direct EA/IA interaction. 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 2002 for Itelmen (Chukotko-Kamchatkan). Analyses of this pattern as PH-driven agreement displacement can be found for Basque in Rezac 2003, for Georgian (Caucasian, Harris 1981, Nash 1995), Karok (Hokan, Bright 1957), and Erza Mordvinian (Uralic, Abondolo 1982) in Béjar 2003, and for Itelmen in Rezac 2006.3 Table 1 illustrates the first four.

Table 1: PH-driven agreement displacement

<table>
<thead>
<tr>
<th>IA controller</th>
<th>Basque</th>
<th>Georgian</th>
<th>Karok</th>
<th>Erza</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>z-in-t-STEM-zte-da-n</td>
<td>g-STEM-t</td>
<td>ki-STEM-ap</td>
<td>STEM-d-ad-yż</td>
<td>I V you.PL</td>
<td></td>
</tr>
<tr>
<td>g-in-t-STEM-zu-n</td>
<td>gv-STEM</td>
<td>kin-STEM</td>
<td>STEM-s-am-izu</td>
<td>You V us</td>
<td></td>
</tr>
<tr>
<td>g-in-t-STEM-n</td>
<td>gv-STEM-s</td>
<td>kin-STEM</td>
<td>STEM-s-am-izu</td>
<td>He Vs us</td>
<td></td>
</tr>
<tr>
<td>1.PL-X-PL-STEM-T</td>
<td>1.PL-STEM-1</td>
<td>1.PL-STEM</td>
<td>STEM-T-1-PL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EA controller</th>
<th>Basque</th>
<th>Georgian</th>
<th>Karok</th>
<th>Erza</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>z-en-it-STEM-n</td>
<td>∅-STEM</td>
<td>'i-STEM</td>
<td>STEM-s-y-nk</td>
<td>You V them</td>
<td></td>
</tr>
<tr>
<td>2-X-PL-STEM-T</td>
<td>2-STEM</td>
<td>2-STEM</td>
<td>STEM-T-PL-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>g-en-it-STEM-n</td>
<td>v-STEM-t</td>
<td>nu-STEM</td>
<td>STEM-s-y-ńek</td>
<td>We V them</td>
<td></td>
</tr>
<tr>
<td>1.PL-X-PL-STEM-T</td>
<td>1-STEM-PL</td>
<td>1.PL-STEM</td>
<td>STEM-T-PL-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 on the search-space given by cyclic construction of the phrase-marker, giving rise to the IA > EA preference. The basic idea that a class of PH-effects arises via the Case/agreement mechanism is due to Nichols

3 Karok and Erza paradigms have gaps in the agreement displacement pattern in certain cells, where agreement morphology lacks systematicity altogether: X>2.SG in Karok, SG>1/2.SG in Erza. Following Béjar (2003), we abstract away from these; they belong to feature combinations known to typologists for susceptibility to skewed forms (cf. Heath 1998, Beck 2003), and make the systematicity of the rest of the paradigm all the more striking. Functionalist discussions relate these anomalies to politeness/formality coding.
(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. A correct prediction is that whenever the argument of a lower clause falls into the scope of Agree, it behaves exactly as the IA of our discussion, as in ECM, causatives (Nichols 2001:523), or cross-clausal agreement constructions (Branigan and MacKenzie 2001, Bruening 2001:chapter 5). We thus limit our examples to the IA of basic transitives. If Case and agreement are syntactic as in Chomsky 1995, 2000, Nichols's argument strongly supports a syntactic account. Rhodes's (1994) demonstration that in some Ojibwa varieties the PH oscillation between EA and IA as agreement controller correlates with it being the pivot of certain syntactic diagnostics confirms the syntactic reality of PH effects (Rezac, in prep.). One of the main intended contributions of the present paper is a syntactic, derivational model of PH-driven agreement displacement.

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

\[ [\pi \text{EA} \ [v+\text{AGR} \ [v\pi \text{V} \text{IA}]]] \]

By postulating a unique low locus of agreement (on v) we depart from the received practice of associating the core AGR in a clause with a higher head, usually T (Béjar 2000). This is strongly motivated by the data. The very existence of the bleeding pattern tells us that we are dealing with a single π-probe oscillating between two controllers, which correlates with the fact that in all cases we are dealing with a single agreement slot for the core pattern. The IA-preference manifested in the bleeding pattern tells us that this single π-probe must be low. No bleeding pattern would be expected if the clausal architecture included one low AGR for the IA and another high AGR for the EA, since the availability of an AGR for the EA would then be independent of whether or not the IA had agreed. A further argument for (4) will come precisely from situations where the core pattern of a single probe is disturbed by the addition of a second agreement slot dedicated to the EA. In the systems under discussion, this occurs only in contexts where the bleeding of EA-Agree by IA-Agree with the core probe on v would leave the EA without φ-Agree entirely, leaving it without Case licensing (viewed as licensing of person features). Our account will predict this distribution of extra EA-controlled agreement, as an added probe to repair the failure of EA licensing. Positing throughout a dedicated high EA-controlled probe would predict neither this distribution of added agreement, nor its limitation to EA control: one would expect either two independent agreement slots or extra agreement added in contexts where the IA fails to be Case-licensed by φ-Agree.

We limit our discussion to person features. The system also predicts similar phenomena for other φ-features, of which Béjar (2003) explores number. Among person hierarchies, our proposal is for one class of PH phenomena only. For illustration of PH-effects that do not fall into the scope of our approach, we refer the reader to two recent analyses of certain transitive forms repaired by non-agreement, Bobaljik and Branigan (2006) for Chukchi (Chukotko-Kamchatkan) and Wiltshcko (2003) for Thompson River Salish. In both cases, derivation of the transitive would yield a form that has no morphological spell-out because of idiosyncratic morphological gaps in agreement, such as 3.SG→1.SG but not 3.PL→1.SG/PL. One option then is to resort to another independently available numeration to achieve a similar meaning, such as the passive or non-agreeing strong pronouns (Wiltshcko 2003). Another is to repair the derivation at the interface by removing the offending agreement (Bobaljik and Branigan 2006). If the analyses are correct, these phenomena make reference to conditions and operations within the
post-syntactic component that are governed by different properties than those of narrow syntax, for example bans on arbitrary feature combinations on the same node or under linear adjacency (see Noyer 1997, Bobaljik 2000, Embick and Noyer 2001, Embick and Marantz 2006). For the PH pattern we characterize as agreement displacement, a morphological analysis is of course possible (Bobaljik (2000) for Itelmen, Halle and Marantz (1993) for Potawatomi). But in contrast to the special character of the phenomena in the analyses just discussed, those analyzed here conform under our mechanism to conditions and outcomes that are determined by the mechanics of narrow syntax itself. This predicts that there may be narrow-syntactic consequences beyond agreement, as seems correct (section 5). Our proposal thus draws a sharp line between syntactic and morphological PH phenomena (Béjar 2003, Rezac 2006).

There exist various other approaches to PH phenomena within the same broad framework adopted here, although we cannot attempt a comparison here. One class differs from ours in that it assumes person hierarchies as a primitive, e.g. Jelinek and Demers 1983. Closer to ours are proposals that derive PH effects from Case and agreement, such as Laka 1993, Hale 2001, Nichols 2001. Another family of approaches may be termed cartographic: different person values map to different positions of the clause-structure: see Johns 1993, Rice and Saxon 1994, Nash 1997 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. 1st/2nd from 3rd) by giving them different specificity values.

Given (4), two questions arise immediately: (i) why does the IA fail to control agreement if it has a certain person specification; (ii) how does this allow the same π-probe to Agree with the EA. We turn to these questions in the next section.

3 Theory of cyclic Agree and PH-driven agreement displacement

3.1 Articulated probes, feature-relativized locality, and person licensing

We refer to the pattern where first the IA is evaluated with respect to Agree and then the EA as cyclic expansion. In section 3.2 we argue that it follows from cyclic construction of the phrase-marker, which makes the IA the first potential match for Agree by a probe on v, and EA the second (Rezac 2003). 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 situate our approach in the framework of Chomsky 2000, where the conditions on Agree are given as follows:

(5) 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" (Chomsky 2000: 122)
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 \( D(G') \) (for clarification, see note 7, and Collins (2002:57-9), Fitzpatrick (2002), Rezac (2004:24ff.)).

We articulate \( \pi \)-features 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 \([\pi F]\) (designating an uninterpretable / unvalued occurrence of \([F]\)) of such an articulated \( \pi \)-probe simply when the IA lacks matching \([F]\) (interpretable / valued); \([F]\) on EA can then be the goal of Agree. Therefore, control by IA, Figure 1 (A), and bypassing of IA for control by EA, (B), display the same logic as classical feature-relativized locality for two arguments in (C-D), where \( DP_1 \) is a goal only if it bears \([F]\), (C), and is bypassed otherwise, (D).

Figure 1: Locality patterns

Cyclic expansion

(A) \[ DP_2 \rightarrow DP_1 \text{ Agrees} \]
\[ [F] \quad [\pi F] \quad [F] \]

(B) \[ DP_2 \rightarrow DP_1 \text{ Bypassed} \]
\[ [F] \quad [\pi F] \quad [F] \]

Standard locality pattern

(C) \[ \rightarrow DP_1 \text{ Agrees} \]
\[ H \quad [F] \quad DP_1 \quad DP_2 \]

(D) \[ \rightarrow DP_1 \text{ Bypassed} \]
\[ H \quad [\pi F] \quad [F] \quad DP_1 \quad DP_2 \]

Both cyclic expansion as well as the standard locality patterns arise as a consequence of feature relativized locality, which is encoded in (5)a as the condition on matching: a probe for a feature \([\pi F]\) 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 \( \pi \)-Agree must be sensitive to a fine grain of person specifications, so that \( \pi \)-Agree of a probe looking for a 1\textsuperscript{st}/2\textsuperscript{nd} person argument can be undervalued by 3\textsuperscript{rd} person DP, simply because the DP lacks the features to fully value it. 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.\(^4\)

One such system is developed by Harley and Ritter (2002) for morphological \( \phi \)-features, which we extend to the \( \phi \)-features visible to Agree, both interpretable and uninterpretable, following Béjar (2000, 2003). The \( \phi \)-bundle is organized into subsets that reflect both natural classes and semantic entailment relations, as in Figure 2 for person. Here, all persons include some shared feature, our \( \pi \). 1\textsuperscript{st} and 2\textsuperscript{nd} persons are additionally specified as discourse participants and thus grouped into a natural class to the exclusion of 3\textsuperscript{rd} persons. Finally, 1\textsuperscript{st} and

\(^4\) In principle, the same result could be obtained in a fully specified feature system with bivalent values (e.g. Halle 1997’s \([+/-\text{participant}, +/-\text{author}]\), if the ability to Match or Agree were made contingent on having \([+]\) or \([-]\) values for \( F \), rather than on presence or absence of \( F \) (cf. Nevins 2007). However this predicts a much broader typology of agreement systems than we can motivate empirically; see further Béjar (2003).
2nd persons are themselves differentiated from one another by the further specification of 1st persons with a feature distinguishing it as speaker.

Figure 2: Entailment (subset) relations among person (π) features

![Diagram showing entailment relations among person features: 1st person → 2nd person → 3rd person]

This yields the entailments in (6), given that a set containing a feature (structure) $F$ entails a feature (structure) $F'$ iff $F'$ is a subset (including identity) of the least set containing $F$. For example, being specified as [speaker] entails being specified as [participant] and as $\pi$.

5 These entailments translate to degrees of privative feature specification through a heuristic logical underspecification, where $\pi$-values are differentiated only by the presence vs. absence of features, as in Table 2 (A). This requires specifying default interpretations for underspecified representations: for example, $\pi$ is common to all persons, but a bare $\pi$ feature is interpreted as 3rd person.

(6) Entailment: $[\text{speaker}] \rightarrow [\text{participant}] \rightarrow [\pi]$

Table 2: Person Specifications

<table>
<thead>
<tr>
<th>A: Person Specifications</th>
<th>B: Shorthand 1&gt;2&gt;3</th>
<th>C: Shorthand 2&gt;1&gt;3</th>
</tr>
</thead>
<tbody>
<tr>
<td>3rd person</td>
<td>2nd person</td>
<td>1st person</td>
</tr>
<tr>
<td>$[\pi]$</td>
<td>$[\pi]$</td>
<td>$[\pi]$</td>
</tr>
<tr>
<td>[participant]</td>
<td>[participant]</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td>[2]</td>
<td>[1]</td>
</tr>
<tr>
<td></td>
<td>[1]</td>
<td>[1]</td>
</tr>
</tbody>
</table>

We adopt these feature specifications, but for convenience we employ a shorthand from hereon in: we write $[\pi]$ 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. The representations corresponding to Table 2 (A) in this abbreviated system are given in Table 2 (B). The notation is convenient because the interpretation can be transparently read by inspecting the bottom-most segment in the feature bundle. It is important, however, that these segments not be read as person categories; [1] for example in Table 2 (C) does not refer to the category of 1st person, but rather to [speaker]. It is only the feature structure as a whole that corresponds to a traditional category like 1st person.

The system assumes limited variation in the selection of features (see Harley & Ritter 2002). Of relevance below 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. This is shown in Table 2 (C).

---

5 The entailment relation between feature segments is integral to our formalization of the operations Match and Value, as shall be seen shortly. This excludes feature systems that do not encode intrinsic entailment relations, like Anderson's (1992) [+-me, +-you].
In light of this feature-theoretic approach to π specification, matching requirements can be relativized to specific π-structures by manipulating the specifications of a probe: the more highly articulated a probe is, the more highly specified must a DP be to match all of a probe's features (cf. Béjar 2003). (7)-(9) show this for the three possible articulations of the probe: a flat probe that is just \([uπ]\) ([u3] in our notation) in (7), a medium-articulated probe in (8), and a fully articulated probe in (9). For each probe, a DP as (or more) highly specified as the probe will be a match for every feature of the probe (signified by a dash). However, a DP less specified than a probe will only match a subset of the probe's features, leaving an active residue, set in bold in (8) and (9). This active residue can, by feature-relativized locality, Agree with another DP in the search-space of the probe; it is this active residue that will lead to agreement displacement.

(7)  a. \(v\) DP    b. \(v\) DP    c. \(v\) DP
    \([u3]— [3]\)    \([u3]— [3]\)    \([u3]— [3]\)
    \([2]\)    \([2]\)    \([2]\)
    \([1]\)    \([1]\)    \([1]\)

(8)  a. \(v\) DP    b. \(v\) DP    c. \(v\) DP
    \([u3]— [3]\)    \([u3]— [3]\)    \([u3]— [3]\)
    \([u2]\)    \([u2]— [2]\)    \([u2]— [2]\)
    \([1]\)    \([1]\)    \([1]\)

(9)  a. \(v\) DP    b. \(v\) DP    c. \(v\) DP
    \([u3]— [3]\)    \([u3]— [3]\)    \([u3]— [3]\)
    \([u2]\)    \([u2]— [2]\)    \([u2]— [2]\)
    \([u1]\)    \([u1]\)    \([1]\)

(8) schematizes exactly what we are looking for to account for PH-driven agreement displacement in languages like those in Table 1: a system like Basque or Georgian 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 \([u2]\), which may Agree with another argument.

The specification of a probe is subject to variation, giving cross-linguistic differences in PH-sensitivity (Béjar 2003). A language with no PH-sensitivity in its agreement system is modeled by assuming the flat probe of (7); any DP will fully match a probe. A convincing example requires a language with genuine object agreement rather than the spell-out of a pronoun, as in the following Swahili (Bantu) example, where the object marker \(ki\) agrees with \(chochote\) for the noun class 7, and not with the subject 'I' (\(pro\)) which controls the 1st person singular subject marker (cf. Bresnan and Mchombo 1987: 777-8, Morimoto 2002). Béjar (2003: 91ff.) posits a flat probe for Abkhaz and Choctaw, with overt non-interacting subject and object agreement, and for Germanic and Romance.

(10)  \(si-ja-ki\-ona\)               \(chochote\,
    1.SG-NEG-7-see anything
    I haven't seen anything. (Swahili, Morimoto 2002, Wald 1979; gloss adapted)

On the other hand, languages like Nishnaabemwin, Mohawk, and Kashmiri will be seen to have the fully articulated probe of (9), 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 \( \pi \), which determines their PH-sensitivity. The surface reflex of this variation lies precisely in the patterns of person sensitivity in cyclic displacement, presumably serving as input to acquisition of the structure of the probes.

Three details must be discussed about the application of Agree to such feature structures. First, for uninterpretable features we assume that some match must be found. This is part and parcel of Full Interpretation, which requires uninterpretable features to have been deleted by LF, under the standard assumption that deletion is not free and requires Agree of an uninterpretable feature (bundle) with a corresponding interpretable one. What counts as correspondence requires spelling out. In our system, features are organized into structures whose internal properties determine both feature classes (e.g. person) and values (e.g. speaker). We take the deletion-licensing requirement to be the Match Requirement (11), which allows correspondence between two non-identical feature structures if the interpretable one is identical to a subset of the uninterpretable one (cf. Chomsky (2000:124) where identity of feature, not value, matters).

(11) Match Requirement: for a probe segment \([uF]\), a subset \([uF']\) of \([uF]\) must match.

If an uninterpretable feature structure satisfies the Match Requirement, its deletion at the interface is licensed. Unmatched segments within such a structure pose no problem. This will be the fate of the active residue e.g. (9)a, if not later valued.

Second, our use of feature structure in characteristic probes must be kept distinct from the use of features as a PF instruction expressing valuation as a consequence of Agree. The characteristic probe delimits conditions on matching and deactivation of the probe, but not the values expressed by agreement. Clearly, languages with a flat probe like Icelandic or Spanish are not restricted to 3rd person agreement; valuation of their probes can distinguish 1st and 2nd persons as well. Likewise for a \([u-3-2]\) probe language like Basque. 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 feature copying to the target. There are various ways to model this; we adopt the following:

(12) Assumptions for Agree:

a. Each feature that seeks to Agree is active upon insertion into the derivation.
b. Agree for a feature \([uF]\) upon matching with a goal \([F']\) copies the feature structure containing \([F']\) (i.e. all features that entail \([F']\)) to \([F]\); this constitutes valuing.
c. An active feature that is locally related to a non-active feature (that is, standing in the configuration created by (12)b) is no longer active.

Thus, what technically happens in (8)c is that the \([u3]\) and \([u2]\) segments of the probe, active, match the \([3]\) and \([2]\) segments of the IA, the entire feature structure of the IA including the unmatched \([1]\) segment is copied to the probe, which deactivates both the \([u3]\) and \([u2]\) segments; while in (8)a, copying of the \([3]\) segment of the goal to the probe upon match by the \([u3]\) segment of the probe leaves still active the \([u2]\) segment of the probe. For simplicity, we dispense henceforth with indicating copying (valuation), because what is important to us is the deactivation or lack thereof of individual \(\pi\)-features. This can be discussed more simply by indicating only match relationships between active features as in (8) 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 transfers as much of the \( \pi \)-structure as there is on the goal, not just those segments which are active on a probe.

(11) and (12) together indicate the role played in Agree by organizing features into structures characterized by subset relations (entailment). 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 [\( u2 \)] in (8)a is crucially capable of match once [\( u3 \)] has been deactivated.\(^6\)

Third, the decomposition of person into combinations of syntactically independent units interacts with proposals that (certain) \( \phi \)-features must be licensed, for example by Case (the Case Filter) or in designated configurations. We adopt the following condition:

(13) Person Licensing Condition (PLC): a \( \pi \)-feature \( F \) must be licensed by Agree of some feature of which \( F \) is a subset.

The PLC is proposed by Béjar and Rezac (2003: 53) for the Person Case Constraint (Bonet 1991), where a single \( \pi \)-probe c-commands two DPs, as in (14) (Anagnostopoulou 2003, Béjar and Rezac 2003, Rezac in press, forthcoming; cf. Albizu 1997, Boeckx 2000, Chomsky 2000: 127-128, Ormazabal and Romero 2007, Adger and Harbour 2007). The closer DP1 is "quirky": visible to and moveable by the \( \pi \)-probe, but at the same time oblique so that it cannot value it. After it is moved, there is no \( \pi \)-probe left on AGR for DP2. This leads to a crash of such derivations if the farther DP2 does not have another sufficiently local probe to Agree with it and license its person feature. Béjar and Rezac (2003: 54) posit that inherent Case and focus, which both protect a DP from PCC effects, involve shells around a DP that contain a local \( \phi \)-probe for this purpose (cf. Cardinaletti and Starke's (1999) \( \gamma' \)); these are missing on clitics, pro, and on ni in (15).

(14) PCC: In [\( \alpha \) AGR … DP1-oblique … DP2 …], where \( \alpha \) includes no other person AGR, DP2 cannot have a marked person feature (1\text{st}/2\text{nd}, sometimes 3\text{rd} animate).

(15) Zu-kï etsaia-rij misil-akk / *nik saldu d-i-zkid-oj-zuk / *nik-(a)-i-oj-zui.


You have sold the missiles / *me to the enemy. (Basque, cf. Albizu 1997)

As the PCC indicates, a derivation may fail to satisfy the PLC. In the person hierarchy phenomena we discuss, this will occur whenever the IA fully controls the unique \( \pi \)-probe of \( v \) in a transitive structure, leaving person features of the EA without Agree. In such environments

\(^6\) The matter is unclear for one occurrence of a probe c-commanding two DPs, the closer DP1 less specified (e.g. [3]) than the farther DP2 (e.g. [3-1-2]). The [\( u-3-1-2 \)] probe might be expected to Agree first for [3] with DP1 and then for [1-2] with DP2. Assuming double object constructions fit this description, this does not seem to be the case: the Person Case Constraint (below) bans a 3\text{rd} person goal DP1 + 1\text{st}/2\text{nd} person theme DP2 in [\( u-3-1-2 \)] or [\( u-3-2-1 \)] probe languages (e.g. Baker 1996:194ff. for Mohawk), indicating to us that the non-[3] segments of DP2 were never matched. This is evidence that a segment (e.g. [\( u2 \)]) cannot match past a feature structure that intervenes for any segments it entails (e.g. [3]) (or alternatively, for the root segment entailed by all the others). Such a condition is consistent with the letter and spirit of our system. When reprojection of a probe brings another DP into the search-space, earlier-matched segments encounter no interveners.
there may surface repair strategies to Agree-license the EA through otherwise impossible agreement or case morphology. These are the subject of section 4, where we propose that they reflect the derivational addition of a $\pi$-probe to $v$. We adopt the proposal of Adger and Harbour (2007) that in contrast to an IA, where some 3rd persons have no person features in the sense relevant to the PLC, an EA always has one: in our system, a 3rd person EA is always at least [3].

We intend the PLC to fall under the Case Filter, articulated to take into account structured $\phi$-feature bundles. To be fully identified with the Case Filter, the PLC must be generalized to those third person IAs that are fine as DP$_2$ in the PCC context (14). Since these Agree for number even there, we need simply assume that their $\phi$-content is a subset of and thus licensed by a number probe (cf. Anagnostopoulou (2003:274ff., 291-4)). In Chomsky 2000, 2001 the Case Filter is implemented by [uCase] on DPs that renders their $\phi$-set visible to Agree. Being uninterpretable, [uCase] must be deleted for Full Interpretation, and those classical Case Filter effects that do not fall out from the need of $\phi$-probes to delete do so from that of [uCase]. Assuming this implementation of the Case Filter, then the PLC, like the Match Requirement (11), determines when Agree licenses the deletion of [uCase] in a system with feature structures. Such a full statement of the Case Filter would look like PLC, with the reference to "person feature" replaced by "feature" and licensing referring to valuation of [uCase]. However, our discussion only assumes the PLC as formulated, and we do not address the asymmetries between $\phi$-probes and [uCase] that this view of the Case Filter inherits from Chomsky 2000.

To summarize, we adopt an approach to $\pi$-features that allows us to distinguish individual $\pi$-values by representing them as subsets of a single feature structure. Given feature-relativized locality 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 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 $\pi$-probes and IAs; in the next section, we will derive the IA > EA preference of the cyclic expansion pattern.

3.2 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 cyclic construction of the phrase-marker (Rezac 2003), combined with locating the relevant $\pi$-probe on an AGR head, $v$, between EA and IA (Béjar 2000ab, 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 (Frampton and Gutmann 1999, Chomsky 2000:132, Collins 2002:46) such as $v$ is given by the Earliness Principle, which requires a feature to probe as early as possible (cf. 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, modulo the effects of Agree (Chomsky 2000:133-4, 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; (ii) upon subsequent Merge of the EA and further projection
of \(v\), \([\text{Spec}, v]\) falls into the domain of any remaining probe on \(v\) according to (5)b because \(v'\) is the sister of the new projection of \(v\) under Bare Phrase Structure. We walk through this for a transitive construction in (16), annotating the projections of \(v\) as \(v_n\). Each \(v_n\) is identical to \(v_{n-1}\) modulo checking/valuation, but in a different configuration (that of a label to the item that projects it), and we annotate them with numbers for convenience only.

(16) Derivation of a transitive \(vP\)

Step 0: \(\text{VP constructed as } \{V, \{V, \text{IA}\}\}; \text{ }v \text{ becomes locus}\)
Step 1: \(\text{Merge}(v, \text{VP}) \Rightarrow \{v_1, \{V, \{V, \text{IA}\}\}\}\)
Step 2: \(\text{Agree}(v_1, \text{IA})\)
Step 3: \(\text{Merge}(vP, \text{EA}) \Rightarrow \{v_{1l}, \{\text{EA}, \{v_1, \{V, \{V, \text{IA}\}\}\}\}\}\)
Step 4: \(\text{Agree}(v_{1l}, \text{EA}), \text{if there is still a probe on } v_{1l}\)

It is thus the cyclic architecture of the derivation, and locating a \(\pi\)-probe on \(v\), that is responsible for PH-driven agreement displacement: we term this cyclic expansion of search-space. Figure 3 schematizes this for two features \([uF], [uG]\) on the AGR head \(v\). What is crucial is the derivational mechanics and the position of the \(\pi\)-probe between EA and IA, not the actual identification of the AGR head with \(v\), the introducer of the EA.\(^7\)

Figure 3: Cyclic search-space expansion

Consider how this works for Nishnaabemwin (Algonquian, Valentine 2001), a language that has a fully articulated probe with the structure \([\pi \text{ participant } \text{[addressee]}]\) so that 2\textsuperscript{nd} person is the most specified, as noted in 3.3. In our shorthand this is \([ur-3-1-2]\). If the IA is 3\textsuperscript{rd} person, a

\(^7\) The interaction of projection with search-space needs spelling out. Under Bare Phrase Structure as in Chomsky (1995:241-9, 2000:116, 133), but assuming labels in syntax, Merge of \(v\) with VP creates \(\{v, \{v, \text{VP}\}\}\). If \(x\) is the sister of \(y\) iff \(x, y \in z\), then the sister and search-space of the lowest position of \(v\) is VP, and of the first label both \(v\) and VP. These do not give distinct goal potentials. It is thus to these that we refer as \(v_1\) (we could notate them \(v_0, v_1\)). The next Merge, that of EA, yields \(\{v_2, \{\text{EA}, \{v_1, \{v_0, \text{VP}\}\}\}\}\). The new label \(v\), underlined, has as its sister and search-space a constituent that includes EA as its highest element; this is our \(v_{1l}\). Alternatives viewing the label of the object \(\alpha\) resulting from Merge(\(v\), VP) as the sister of EA, along with \(\alpha\) itself, give the same results for \(v_1, v_{1l}\). In this discussion, we assume the projection of labels in narrow syntax, with Chomsky (1995), Donati (2006), Hornstein (2005), Boeckx (to appear), but not Chomsky (2005), Collins (2002). A label is a copy of the projecting item, save the consequences of Agree (cf. Chomsky 2000: 133), for us specifically the copied interpretable features to the probe that deactivate it and license its deletion. The basic cyclic Agree proposal does not require labels (Rezac (2003)), but our specific mechanics here do (Rezac (2002), Béjar (2003)). However, reformulation is available by treating \(v_{1l}\) as \(v\) raised to some head above the EA like \(T\) (Béjar (2000ab), Rezac (2006)), from which any active features remaining on it at that point can probe. This captures some effects of Baker's (1988:64) Government Transparency Corollary (cf. Roberts 1991, Chomsky 1995).
1st/2nd person EA controls agreement because the probe segments \([u1], [u2]\) are not affected by Agree with the IA, and project unchecked to \(v_{II}\). Similarly, if the IA is 1st person, the segment \([u2]\) projects. 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 4.2. The corresponding derivations are in Table 3, where a probe valued and thus deactivated on a lower projection of \(v\) is in brackets, and strike-through indicates a segment that never matches in the derivation.

\[
\begin{align*}
\text{(17) a. } & \text{g-waabm-in} & 1 \rightarrow 2 = 2 & \text{b. } & \text{g-waabm-i} & 2 \rightarrow 1 = 2 \\
& \text{2-see-1.INV} & & & \text{2-see-DFLT.1} & \\
& 'I \text{ see you.'} & & & 'You \text{ see me.'} & \\
\text{c. n-waabm-ig} & 3 \rightarrow 1 = 1 & \text{d. } & \text{g-waabm-ig} & 3 \rightarrow 2 = 2 \\
& \text{1-see-3.INV} & & & \text{2-see-3.INV} & \\
& 'He \text{ sees me.'} & & & 'He \text{ sees you.'} & \text{(Nishnaabemwin)}
\end{align*}
\]

Table 3: Core agreement in Nishnaabemwin

<table>
<thead>
<tr>
<th>EA→IA</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>([u2]), ([u3])</td>
<td>([u3])</td>
<td>([u3])</td>
</tr>
<tr>
<td>([u1])</td>
<td>([u1])</td>
<td>([u1])</td>
</tr>
<tr>
<td>([u2])</td>
<td>([u2])</td>
<td>([u2])</td>
</tr>
</tbody>
</table>

The same logic accounts for the PH-driven agreement displacement discussed in section 2 for languages like Basque, ex. (2); there we are dealing with a \([u-3-2]\) probe, so that a 3rd person IA will not check the \([u2]\) segment, while both a 2nd and 1st person IA will. In a language with a flat probe like Swahili, the IA will always count as a match for the sole segment of the probe, and no active residue will ever remain on \(v_{II}\) for the EA.

As the derivations in Table 3 make clear, we end up with a single articulated probe whose individual segments can match with segments on different DPs through its reprojection. As each segment matches, valuation ensues via copying of the whole \(\pi\)-value of the matching DP, and all segments of the probe that have a corresponding segment in the copied value are deactivated (see (12)). Given the Nishnaabemwin \([u-3-1-2]\) probe with a \([3-1]\) DP in its scope, it does not matter whether \([u3]\) or \([u1]\) or both match; in both cases the \([3-1]\) value of the DP is copied and the \([u3], [u1]\) segments of the probe are deactivated. Projection of the category hosting the probe (\(v_1\) to \(v_{II}\)) can add a new DP to the search-space of the copy of the probe on the projection (\(v_{II}\)). From here any unmatched segments of the probe can match again.
The logic of entailment among segments, such that \([2]\) entails \([1]\) and \([1]\) entails \([3]\) in a \([3-1-2]\) structure, has as consequence that matching deactivates and values the probe in a continuous monotonic fashion. Given the "lowest" unmatched segment \([u\alpha]\) (the one entailed by all other unmatched segments), any segment entailed by \([u\alpha]\) will necessarily have been matched and deactivated, and any segment that entails \([u\alpha]\) will not have. It is not possible to end up with \([u1]\) matched and \([u2, u3]\) unmatched. Thus as a probe projects, any unmatched segments are in a contiguous bundle at its "top", and their match can only value the probe to a value that is a superset of (entails) the earlier value (say from \([3-1]\) to \([3-1-2]\)). This renders moot the question of what happens to the original value. It is always represented as a subset of the new value, and no questions about de-linking arise as would if (re)-valuation to a lower value were possible.

At the same time, the syntax retains the representation of the valuation of a probe on each phrase-structural locus until spell-out deletes it (Rezac 2002). Thus the cyclic expansion mechanics for agreement displacement predicts the possibility of second cycle effects (Béjar 2003:79), where such multiple representation of a probe is referred to by contextual allomorphy. Agree with the IA takes place on a different cycle or different projection of \(v\) than Agree with the EA, and this difference in derivational mechanics is reflected in the morphology of agreement in languages like Georgian and Karok (both with \([u-3-2]\) probes). In such languages, a particular value of a \(\pi\)-probe on \(v\) is systematically spelled out using one morpheme if it has been valued on the first cycle (\(\text{IA} \geq \text{EA}\)), and a different morpheme if it has been valued on a second cycle (\(\text{EA} > \text{IA}\)). A second cycle effect differs from the morphological contrast that might arise in a language with separate AGR heads for 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 (e.g. \(1 \rightarrow 3 = 1\) but not \(1 \rightarrow 2 = 2\) in Georgian), whereas a system that has for IA and EA a dedicated \(\pi\)-probe each (for example on \(v\) for IA, \(T\) for EA) predicts distinct morphology for EA regardless of the value of the IA (e.g. \(1 \rightarrow 3, 1 \rightarrow 2\)).

Georgian 1.SG yields an example of second cycle morphology (Béjar 2003: 127ff., 151ff.; 159-161 for Karok). When 1.SG IA controls the \(\pi\)-probe, which occurs regardless of the value of the EA, 1.SG is spelled out as \(m\), (18)a. When 1.SG EA controls the \(\pi\)-probe, which occurs when the IA is 3rd person, 1.SG is spelled out as \(v\), (18)b.

(18) a. m-xedav-s 3\(\rightarrow 1 = 1.I\)
1.1-see-x 1.1.I
'He sees me.'
   b. v-xedav 1\(\rightarrow 3 = 1.II\)
3.2-see-x 3.2.II
'I see him.' (Georgian)

Second cycle morphology shows up when the \(\pi\)-probe is valued on the second cycle, which is from the \(v_{II}\) projection c-commanding the EA; first cycle morphology when it is valued on the first cycle, on \(v_{I}\) (\(\leftrightarrow\) symbolizes Agree; \([F'], [F'']\) are the \(\pi\)-feature specifications of EA, IA; \([uF]\) is the \([u-3-2]\) probe of Georgian):

\[
\begin{align*}
[v_{II}] & \quad \text{EA} & \quad [v_{I}] & \quad [v] & \quad \text{IA}] \\
[uF] & \leftrightarrow [F'] & [uF] & \leftrightarrow [F'']
\end{align*}
\]

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 modeled
as sensitivity of vocabulary insertion to the difference between two occurrences of $v$: a lower one (the lowest being the head), and the higher one that the lower occurrence projects. The two differ configurationally and a vocabulary insertion rule for one may therefore refer to one or the other, as an instance of contextual allomorphy. This presupposes two important points, both of which are natural in Bare Phrase Structure: first, that the projection of $v$ be a potential vocabulary insertion site (see Béjar 2003, in prep.); second, that the lowest projection (the head) remains differentiated from higher ones (labels) configurationally, in the same way as the lowest copy in a chain differs from higher ones. Adjacent projections of $v$ are sufficiently local to one another to trigger alternations that can be captured by vocabulary insertion rules of the form in (20):

(20) Spell-out the feature set $\Gamma$ of $v_h$ by the exponent $\sigma$ in the environment of $v_{h+1}$

Thus, in the example of Georgian 1.SG agreement (18):  

(21) a. First-cycle vocabulary item: $m \leftrightarrow [3-2-1] // [\_\_\_]$  
    b. Second-cycle vocabulary item: $v \leftrightarrow [3-2-1] // [\_\_\_], [\ldots]$

As with allomorphy in general, the occurrence of this alternation cannot be said to be predicted. Whether or not a language manifests it is a language-specific and (at least synchronically) idiosyncratic property, up to complete absence (in Basque, Nishnaabemwin, or Mohawk). Cyclic expansion merely creates a plausible conditioning environment for such allomorphy. Its plausibility rests on the locality between the target and trigger of the alternation (Lieber 1980, Sproat 1985, Bobaljik 1995, Adger et al. 2001), which follows automatically from the intrinsic locality between subsequent projections. Other variants of this kind of allomorphy might be expected. For example, vocabulary insertion on $v_{II}$ could be sensitive not only to the presence of the adjacent projection $v_I$, but also to particular features on $v_I$; or $v_{II}$ could be realized by morphology distinguishing it from $v_I$ independently of Agree. Examples of both follow in section 4.

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. 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.

3.3 Nishnaabemwin: The core probe

In this section we show how the cyclic Agree mechanism developed above derives the basic pattern of PH-driven agreement displacement from Nishnaabemwin, a fully articulated [$u$-3-1-2] probe language as discussed above (the 2nd person being the mostly highly specified). The Nishnaabemwin singular agreement paradigm is given in Table 4.  

---

8 We thank an anonymous reviewer for helping us to clarify the formalism.  
9 Our discussion of Nishnaabemwin focuses exclusively on the agreement pattern known as the "independent order". In the "conjunct order", there are no PH effects (Rhodes 1976). We take this to mean that the $v$ probe introduced in the conjunct is flat. Certain tangential details of the independent paradigm, relating to further articulation of the person geometry, are deliberately not addressed here. First, the language distinguishes proximate 3rd and obviative 3rd' person, where 3rd outranks 3rd' (cf. note 12). Both behave identically in interaction with first and second person,
lies in the prefix agreement morphemes in small capitals, a unique agreement slot that is so far
fully predicted by the system; we call it the core agreement slot. Its spell-out is $n$- for [3-1],
g- for [3-1-2], and $w$- for [3] (proximate).

Table 4: Singular paradigm for Nishnaabemwin (core agr. small caps, theme suffix underlined)

<table>
<thead>
<tr>
<th>EA→IA</th>
<th>2</th>
<th>1</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This prefix cross-references person of the EA when it is more highly specified than the IA,
and of the IA otherwise. Table 5 summarizes the derivations for this paradigm. Agree for an
individual segment is represented by —. First and second cycle Agree, that is Agree taking place
from the first ($v_{1}$) vs. second ($v_{2}$) projection of $v$, are represented by — to the right of the probe
(first cycle) and to the left of the probe (second cycle).

Table 5: Cyclic Agree for the Nishnaabemwin singular paradigm

<table>
<thead>
<tr>
<th>EA→IA</th>
<th>2</th>
<th>1</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

as given in the table for 3rd person. The agreement prefix in 3(')→3(') combinations is $\mathcal{O}$; the theme suffix is -igw in
inverse contexts (3'→3) as given in the table, and -aa for direct contexts (3→3'). 3→3' is not discussed in the text; it
is fully predicted by our system. Second, beside 1st and 2nd person, there is a 1st person inclusive, leading to 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 from its inability to co-occur with another 1st or
2nd person argument, just as 1→1 and 2→2 cannot co-occur; it can be notated 21. It also has the well-known
property of behaving like a 2nd person with respect to person agreement, but like a 1st person with respect to number
agreement. This asymmetry is beyond our scope, and not clearly relevant. What is relevant is that the 21 category
behaves as we predict a category with 2nd person features should do: it fully deactivates a $[u3-1-2]$ probe. Cf.
The shaded cells in the paradigm are those with only one Agree step, with the IA; the probe has no segments left that can Agree with the EA. The remaining cells are those where the \([u-3-1-2]\) characteristic probe of Nishnaabemwin 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 data, as seen in section 4.

(22) Direct/inverse contexts

a. \textit{Inverse context} (shaded): the IA checks the characteristic probe of a language as fully as possible, so that the EA cannot Agree with it at all. For Nishnaabemwin, \((2)/1/3 \rightarrow 2; (1)/3 \rightarrow 1; 3 \rightarrow 3;\) for Basque, \(1/2 \rightarrow 1/2, 3 \rightarrow 1/2/3;\) for Swahili, all. There is no Agree of the core \(\pi\)-probe of \(v\) with the EA in inverse contexts.

b. \textit{Direct context} (unshaded): the EA is more specified than the IA, so that after the characteristic probe has Agreed as fully as possible with the IA, it Agrees for its unchecked segments with the EA. For Nishnaabemwin, \(2 \rightarrow 1/3, 1 \rightarrow 3;\) for Basque, \(1/2 \rightarrow 3;\) for Swahili, not possible. In direct contexts, the core \(\pi\)-probe of \(v\) Agrees with both the IA and EA, for different segments.

The core agreement slot is not all there is to the Nishnaabemwin paradigm. There is a second slot, underlined, whose realization refers to features of both EA and IA. However, setting aside the \(2 \rightarrow 1\) form \(i\), there is a direct/inverse pattern: in direct contexts the suffix depends on the EA (1\textsuperscript{st} person \(in\), 3\textsuperscript{rd} person \(igw\), while inverse contexts show an invariable, default-like exponent (\textit{aa}). In the next section, we argue that this is the correct generalization, and that there is an \textit{added probe} agreeing with the EA in all and only inverse contexts in Nishnaabemwin, Mohawk, and Basque. It is added to Agree with the EA, whose person features need to be licensed, in just inverse contexts because just there the core probe is fully valued from the IA, and does not reach the EA at all.

4 Person licensing and added probe

4.1 Inverse contexts: Person licensing failure and repair

The insertion of extra agreement morphology only in the inverse contexts turns out to fit a more general pattern characterizing languages with PH sensitivities. Inverse contexts are typically distinguished from direct contexts by special morphology; one may speak of disruption to the core paradigm. In Nishnaabemwin, Mohawk, and Basque, the core agreement slot is controlled by the IA, but an extra agreement slot appears for the EA. In the same inverse configurations where the extra agreement appears in Mohawk, also a \([u-3-2-1]\) probe language, Kashmiri instead puts the IA into a special Case, which we will call R-Case (homophonous with the dative), and its sole agreement slot is controlled by the EA rather than the IA. We will speak of both of these phenomena as repair strategies for reasons that will become clear: namely, the
added probe and the R-Case strategies (Table 6). We examine them in detail in sections 4.2 and 4.3.

Table 6: Repair strategies in inverse contexts (for a 3-2-1 probe language)

<table>
<thead>
<tr>
<th>Direct context</th>
<th>Added probe language</th>
<th>R-Case language</th>
</tr>
</thead>
<tbody>
<tr>
<td>(EA &gt; IA: 1/2→3)</td>
<td>EA-controlled AGR only</td>
<td>EA-controlled AGR only</td>
</tr>
<tr>
<td>Inverse context</td>
<td>IA controlled AGR +</td>
<td>EA-controlled AGR only</td>
</tr>
<tr>
<td>(EA ≤ IA: 3→1/2/3)</td>
<td>EA-controlled AGR</td>
<td>IA has special R-Case</td>
</tr>
</tbody>
</table>

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 independent passive, which is not limited to inverse contexts. It is therefore significant that the mechanics of agreement displacement developed in section 3 independently characterize two classes of computations by whether or not the core π-probe Agrees with the EA; inverse derivations are those where it does not (see Table 5). A consequence of the EA failing to Agree is a violation of the Person Licensing Condition PLC (13), developed to account for the Person Case Constraint: the π-features of the EA are not (Case) licensed through π-Agree. It is these derivations that are characterized by repair strategies. Therefore, we propose that the special character of inverse contexts arises from the fact that the EA never Agrees with the core probe, and this is what must be repaired.

The exact character of PLC violations is suggested by the nature of the repair strategies. An added probe language adds an extra agreement slot for the EA in the inverse. We found our analysis on this added probe strategy, and propose that inverse contexts permit to converge a mechanism that adds a π-probe to Agree with the EA. An R-Case language puts the IA in a special Case and has its core probe Agree with the EA rather than with the IA as would be expected in an inverse context. Since the R-Case strategy is coextensive with the added probe strategy, it is desirable to seek a unified analysis. On our proposal, it emerges as simply an alternative spell-out.

The boundary conditions on a unified mechanism behind both strategies are the following: (i) The EA must enter into Agree, either by what is spelled out as (i-a) an added π-agreement slot distinct from that of the core probe (Nishnaabemwin, Mohawk, Basque), or (i-b) the sole π-agreement slot of a language, whose control by the EA is anomalous for an inverse context (this will be the situation in Kashmiri). (ii) Under option (i-b), the IA must receive a Case different from the Case it would receive in direct contexts, the R-Case of Kashmiri. (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 constrains our analytic options, because the application of (ii), that is Case assignment by v to the IA, occurs at the point where the core probe only has the IA in its search-space, yet it must be so restricted that it succeeds only in inverse contexts (iii), without making reference to properties of the yet un-Merged EA.

We posit the following mechanism behind both strategies: the ability of the core π-probe on v to add a probe, (23), reflected most transparently in an added probe language. In any particular derivation, the core probe may or may not have this property P, and thus add a probe or not. We
leave open whether the availability of P itself needs to be parameterized. The \([\nu-3-2-1]\) probe languages in our data have P available (as added probe or R-Case). Of the \([\nu-3-2]\) probe languages, Basque does so, with the situation unclear in Georgian, Karok, and Erza Mordvinian.

(23)  P: If the core probe \(\alpha\) on AGR has P, a probe is added to AGR upon Agree by \(\alpha\).

Addition of a probe by P to the core probe on a projection of \(v\), say \(v_3\), creates a modification of \(v_3\), and therefore requires insertion of the added probe on the next higher projection, \(v_{II}\), assuming that once inserted into the structure a term may not be modified without projecting (Rezac 2002, cf. Chomsky 2000:126). An added probe (convergent only in the inverse) and the second cycle of the unmatched segments of the core probe (occurring only in the direct) thus have this similarity: they are both on \(v_{II}\), standing in some relation to an earlier probe on \(v_1\). This immediately makes an added probe unavailable in constructions lacking an external argument for \(v\): assuming the first Merge of \(v\) creates \(\{v, \{v, \text{VP}\}\}\), the next higher projection of \(v\), where the added probe could be inserted upon Agree by the core probe of \(v\), namely our \(v_{II}\), is created by Merge of the EA which yields \(\{v, \{EA, \{v, \{v, \text{VP}\}\}\}\}\), and does not exist without it (see note 7).

The locus of difference between added probe and R-Case language is spell-out:

(24)  P spell-out:

(i) In Nishnaabemwin, Mohawk, and Basque, Agreement morphology spells-out both probes of \(v\) once fully valued, using potentially distinct morphology for the two probes; in Kashmiri, it spells out only the probe valued on the highest projection of \(v\). (Details of spell-out follow in the respective sections.)
(ii) Case assigned to IA by \(v\) can have a spell-out that reflects whether the core probe of \(v\) has P (Kashmiri R-Case).

(24) does not resort to any mechanisms not independently available. P and its spell-out options unify the identically distributed but disparately realized behavior of added probe and R-Case. Nothing forces overt spell-out of any probe or Case: Georgian is like Basque in the syntax, but it does not spell out the added probe. Richness of morphology may bear on (24). Kashmiri (R-Case) has poor agreement and rich case morphology, while Nishnaabemwin (added probe) lacks case. Interplay between the two may have as consequence lack of languages that spell out both R-Case and the added probe.

Consider now how desideratum (iii), the limitation of P to inverse contexts, can be derived without look-ahead from the \(v\)-IA Agree step to properties of the EA. The locus of the problem is at step 3 of the direct/inverse derivations in (25) for a probe with P. Here for a \([\nu-3-2-1]\) probe language, added probes are boxed, the still-active segments (not matched on a previous cycle) of a probe at any point in the derivation are in bold, while those that have matched previously and are thus deactivated are struck through.

(25)  \[I \rightarrow 2 \text{ (direct) } \quad 2 \rightarrow I \text{ (inverse)}\]

a.  Step 1: Agree with IA
b. Step 2: Add EA, project $v$ to $v_{II}$, and add a probe on $v_{II}$ following (23)

c. Step 3: Agree with EA from $v_{II}$

\[
\begin{array}{c|c|c}
\hline
v_{II} & \text{EA} & v_{II} \\
[u_2] & [u_2] & [u_3] \\
[u_1] & [u_1] & [u_3] \\
\hline
\end{array}
\]

\[
\begin{array}{c|c|c}
\hline
v_{II} & \text{IA} & v_{II} \\
[u_2] & [u_2] & [u_2] \\
[u_1] & [u_1] & [u_1] \\
\hline
\end{array}
\]

At step 3, 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. 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 $v_{II}$; this Agree is indicated as $\sim$. The added probe also Agrees with the EA, since nothing prevents it. This double Agree by two probes with a single EA segment takes place in all and only direct contexts.

It is here that we localize what goes wrong. We propose that the resulting structure crashes because a single phrase-structural locus, $v_{II}$, ends up having two probes on it that match from it and receive the same value from the EA. This leads to a crash because two identically valued elements that are not phrase-structurally distinguished stand in no asymmetric relationship that could be used by the mapping to PF to achieve the asymmetry required for the application of spell-out rules, for the establishment of linear order, or indeed (as an anonymous reviewer observes) to distinguish them as two separate elements in the first place. This problem does not arise if the two identical elements are structurally differentiated by placement on different projections of $v$; this can establish their linear order (e.g. under a top-down vocabulary insertion algorithm). The problem also does not arise if there are two distinct symbols on a phrase-structural locus.\(^\text{11}\)

\(^{10}\) Multiple Agree with the EA violates the Activity Condition: a $\phi$-feature (set) gets deactivated upon Agree for further Agree through Case assignment (Chomsky 2000:122-3, 128, Rezac 2003). In an earlier draft we proposed that this condition goes wrong in the derivations we wish to ban (direct contexts and unaccusatives). However, the condition is problematic for multiple Agree with a single controller, both within a clause (C and T in Dutch dialects, (Carstens (2003), van Koppen (2005)), and in cross-clausal agreement (Polinsky and Potsdam (2001), Branigan and MacKenzie (2001), Rezac (2004), Bhatt (2006)). We concur with Jonathan Bobaljik (p.c.) in having doubts about its status. In an alternative approach where the added probe is on a separate AGR head such as T above the EA (cf. Nichols 2001:533 note 19, recalled to our attention by an anonymous reviewer), the Activity Condition could be deployed to limit Agree with EA to inverse contexts only, when the core probe on $v$ has not Agreed with it. In direct contexts the high probe would have no goal, violating the Match Requirement of (11). Putting the added probe on a separate AGR head from that of the core probe makes more difficult the unification with R-Case. However, fundamentally, the cyclic Agree mechanism has as large a role to play in this alternative as in the approach we adopt.

\(^{11}\) In the 3–3 combination, while $v_{II}$ also has two identically-valued probes, the core probe has been valued on $v_1$ and is not modified on $v_{II}$, so it may be ignored on $v_{II}$ and spelled out from $v_1$. Not so for the example in the 1–2
In the next two subsections, we examine in detail the two derivational strategies used for avoiding PLC violations in inverse contexts.

4.2 Added probe

The following partial paradigm of the transitive singular from Mohawk (Iroquoian, Lounsbury 1953, Beatty 1974, Postal 1979, Baker 1996) exemplifies the added probe phenomenon more clearly than Nishnaabemwin, to which we return presently. Control of the core agreement slot in small caps clearly varies between the IA and the EA according to a [u-3-2-1] probe, creating the direct/inverse context split. This is the only agreement in a direct context. All and only inverse contexts have an extra agreement slot, underlined: an added probe. Both agreement slots use the same morphology, \( k \) [3-2-1], \( hs \) [3-2], \( wa \) [3] (we return to the allomorph \( ku \) marked \( \ddagger \) directly).\(^{12}\)

Table 7: Singular paradigm for Mohawk (added probe underlined)

<table>
<thead>
<tr>
<th>EA→IA</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>—</td>
<td>( ku)-see ( \ddagger )</td>
<td>( k)-see</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/2-see</td>
<td>1-see</td>
</tr>
<tr>
<td></td>
<td></td>
<td>'I see you.'</td>
<td>'I see him.'</td>
</tr>
<tr>
<td>2</td>
<td>(h)s-k-see</td>
<td>—</td>
<td>( hs)-see</td>
</tr>
<tr>
<td></td>
<td>2-1-see</td>
<td></td>
<td>2-see</td>
</tr>
<tr>
<td></td>
<td>'You see me.'</td>
<td></td>
<td>'You see him.'</td>
</tr>
<tr>
<td>3</td>
<td>( wa )-k-see</td>
<td>(h)s-(w)a-see</td>
<td>( hra-wa)-see &gt; ( hra-o)-see</td>
</tr>
<tr>
<td></td>
<td>( 3.INV)-1-see</td>
<td>2-3.INV-see</td>
<td>( 3.M)-DFLT-see</td>
</tr>
<tr>
<td></td>
<td>'He sees me.'</td>
<td>'He sees you.'</td>
<td>'He sees him.'</td>
</tr>
</tbody>
</table>

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 8, for any \([u-3-2-1]\) language like Mohawk, where the added probe is boxed.

---

\(^{12}\)In \( 3\rightarrow 3 \) combinations in Mohawk we pick the combination of different 3rd person specifications such that EA ≤ IA (masculine→neuter), just as in Nishnaabemwin (obviative→proximate, note 9). There are also direct \( 3\rightarrow 3 \) forms when EA > IA, and these turn out as expected: only the core agreement slot valued to [3]. Intransitives split according to whether their sole argument controls the same agreement morphology as the IA in the \( 3\rightarrow IA \) or as the EA in the EA→3 transitives, which does not correlate with unaccusativity (Baker 1996:212-3). The former class includes the 3rd person added probe exponent \( wa \) exactly as if there was a 3rd person external argument. This suggests for them a default EA as in derived causative analyses of unaccusatives (Davis and Demirdache 2000), perhaps comparable to the se of Romance inchoatives like Spanish \( abrir\ se \) inchoative 'open' < \( abrir\) transitive 'open' (Baker 1996:201-2). The latter class for us must base-generate the subject higher than the core \( \pi \)-probe probe, without necessarily identifying the position with that of the transitive EA (cf. Travis 2003).
Table 8: Cyclic Agree and added probe in Mohawk

<table>
<thead>
<tr>
<th></th>
<th>EA→IA</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>—</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>EA</td>
<td>[3] [u3]</td>
<td></td>
<td>IA [3] [u3]</td>
</tr>
<tr>
<td></td>
<td>v</td>
<td>[u3]</td>
<td></td>
<td>— [3]</td>
</tr>
<tr>
<td></td>
<td>IA</td>
<td>[u1]</td>
<td></td>
<td>— [u1]</td>
</tr>
<tr>
<td>2</td>
<td>EA</td>
<td>[3] [u3] [u3]</td>
<td></td>
<td>— [3]</td>
</tr>
<tr>
<td></td>
<td>v</td>
<td>[u2]</td>
<td></td>
<td>[2] [u2]</td>
</tr>
<tr>
<td></td>
<td>IA</td>
<td>[u1]</td>
<td></td>
<td>— [u1]</td>
</tr>
<tr>
<td>3</td>
<td>EA</td>
<td>[3] [u3] [u3]</td>
<td></td>
<td>— [3]</td>
</tr>
<tr>
<td></td>
<td>v</td>
<td>[u2] [u2]</td>
<td></td>
<td>[2] [u2]</td>
</tr>
<tr>
<td></td>
<td>IA</td>
<td>[u1] [u1]</td>
<td></td>
<td>[1] [u1]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The added probe is a derivational option that depends for its existence on the EA \(\leq\) IA relationship of inverse contexts; it does not occur independently, as passive morphology does for example. In the theory proposed in section 4.1, the probe is inserted on the projection \(v\) after the core probe Agrees with IA, and thus the lower bound of its insertion is the projection \(v_{II}\) above the lowest position of \(v\). From \(v_{II}\), EA is the first goal, and the added probe therefore correctly Agrees with it and has no interaction with IA. Placing the added probe on \(v_{II}\) rather than a higher AGR head localizes it on the same lexical item, \(v\), as the core probe, which permits unification of the added probe with the R-Case strategy that must be determinable at \(v\) (see section 4.3). It also leads to the explanation advanced above for the limitation of the added probe to inverse contexts: in a direct context adding a probe leads to a crash because the core and added probe receive identical values from the EA on the same phrase-structural locus \(v_{III}\), while in inverse contexts the core probe is valued from IA on \(v\) and the added probe from EA on \(v_{II}\).

In the Mohawk paradigm, the \(\ddagger\) marked \(1\rightarrow2\) combination has a portmanteau \(ku\) deviating from the expected \(k\), but differently from the added probe. Inverse contexts with added probe show two prefix slots using the same spell-out rules; \(1\rightarrow2\) \(ku\) is rather an allomorph of \(k\) [3-2-1] not containing \(hs\) [3-2]. Our system singles out the \(1\rightarrow2\) combination as the only direct context where the two arguments are both [participant]. The use of a special morpheme here will reappear in Nishnaabemwin. In restricting portmanteaus to a particular agreement configuration, Mohawk and Algonquian differ from languages such as Cherokee or Kiowa where portmanteaus occur throughout the paradigm. It is striking that this \(\phi\)-configuration is precisely the one that would have a single probe successively agree with multiple arguments that are both [participant], and we posit that the recurrence of portmanteau morphology here reflects this fact. Portmanteau morphology arises when features of more than one syntactic terminal are spelled-out by a single vocabulary item. Two paths that can lead to portmanteaus are the occurrence of two \(\phi\)-sets on a single spell-out terminal (position of exponence), bundled in the syntax or by morphological operations like merger and fusion (Halle & Marantz 1993, Bonet 1995, Embick and Noyer 2001, Harbour 2003); and contextual allomorphy whereby a morpheme is the primary exponent of one element and a secondary exponent of another (Noyer 1997, Bobaljik 2000). In \(1\rightarrow2\) direct contexts, a single probe gets valued on two projections of \(v\), on the lower \(v_I\) from the IA and on the higher \(v_{III}\) by the EA. This seems like a natural environment for contextual allomorphy or
fusion/merger. Under the allomorphy analysis, the contexts for insertion of $k$ versus $ku$ are as follows in (26). The reason why Mohawk and Algonquian restrict portmanteaus to [participant] direct contexts rather than direct contexts in general is plausibly due to their more marked status, though it remains to be determined whether UG actually disallows insertion rules differentiating the context of $[3 (\ldots)]$ from $[(\ldots)]$.

(26) a. $k \leftrightarrow [3-2-1] // [\_\_]_v$  b. $ku \leftrightarrow [3-2-1]_v // [\_\_]_v [3-2]_v$

The cross-linguistic recurrence of portmanteau morphology in [participant] direct contexts is notable given that contextual allomorphy is generally unforced. We leave to future research why it seems prevalent or even forced here. Intuitively, one might expect this to relate to $v_I$ being a reprojection of $v_L$. Being copies of the same item distinguished only by differently valued versions of the same probe, it may be that the morphology must somehow reconcile their inconsistent feature structures, recalling the morphological resolution of case conflict via syncretism (McCreight 1988). See further Béjar (in prep).

Mohawk shows clearly the correlation of added agreement morphology with the independently characterized inverse contexts. The added probe is responsible for PLC licensing of the EA, which is required precisely there. [$u$-$3$-$2$] probe languages like Basque can also morphologically signal inverse contexts, indicating the added probe. In Basque PH-driven agreement displacement in (2), there is extra agreement morphology $in(d)$ beside the underlined agreement slot that has not been discussed. An example comes from the dialect of Bolívar (Bizkaian, Yrizar 1992: 502), given in Table 9. The core agreement prefix in small caps is controlled by EA in direct contexts (clear), and by the IA in inverse contexts (shaded), defined with respect to the [$u$-$3$-$2$] probe of Basque. The underlined morphology is what we posit to reflect the added probe (Gómez (1994: 109) also relates it to the Algonquian theme marker, as we do below).\textsuperscript{13}

Table 9: Added probe in a Bizkaian Basque (core agr. small caps, added probe underlined)

<table>
<thead>
<tr>
<th>EA$\rightarrow$IA</th>
<th>1</th>
<th>2</th>
<th>1.P</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>--</td>
<td>s-iñdd-u-te-n</td>
<td>--</td>
<td>N-eb-en</td>
</tr>
<tr>
<td>2</td>
<td>N-iñdd-u-su-n</td>
<td>--</td>
<td>G-iñdd-u-su-n</td>
<td>S-eb-en</td>
</tr>
<tr>
<td>1.P</td>
<td>--</td>
<td>s-iñdd-u-gu-n</td>
<td>--</td>
<td>G-eb-en</td>
</tr>
<tr>
<td>3</td>
<td>N-iñdd-u-en</td>
<td>s-iñdd-u-en</td>
<td>G-iñdd-u-en</td>
<td>eb-en</td>
</tr>
</tbody>
</table>

The characteristic probe determines when the added probe shows up. It appears in Basque in $1 \rightarrow 2$ because that is an inverse context with respect to a [$u$-$3$-$2$] probe (and IA controls core agreement), but not in Mohawk because it is a direct context with respect to a [$u$-$3$-$2$-$1$] probe.

\textsuperscript{13} The realization of this "theme marker" varies across the Basque dialects; it is analyzed in Rezac (2006), along with the other aspects of Basque agreement mentioned here. The relevance lies not in the presence of morphology between the prefix and the root (following a sustainable traditional analysis, $e$ in $eb$ in Table 9 is a theme marker), but in the properties to which it is sensitive and their delimitation. These are: the direct-inverse split, as predicted; the number feature of EA in inverse contexts, which our system would model and properly restrict to the inverse if the added probe includes number; and the possibility of cyclic expansion ("ergative displacement"), which is restricted to the non-present (cf. Algonquian, note 9). This can be modeled in various ways, e.g. by [present] T selecting $v$ with a flat probe. The remaining agreement morphology on the right periphery results from agreement of or cliticization to a higher AGR head such as T.
(and EA controls core agreement). We end up with a strong relationship between the cyclic Agree account for core agreement, which determines 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.

Nishnaabemwin provides a striking illustration of the range of possible variation with respect to the spell out of an added probe. So far, we have discussed the core agreement slot, realized as the prefix (in small caps) in Table 4, here repeated.

Table 4: Singular paradigm for Nishnaabemwin (core agr. small caps, theme suffix underlined)

<table>
<thead>
<tr>
<th>EA→IA</th>
<th>2</th>
<th>1</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>G-see-i  ‡</td>
<td>G-see-aa</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>G-see-in</td>
<td>N-see-aa</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>G-see-ig</td>
<td>N-see-ig</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-see-DFLT</td>
<td>2-see-DFT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-see-DFLT</td>
<td>1-see-DFLT</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'You see me.'</td>
<td>'You see him.'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'I see you.'</td>
<td>'I see him.'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>'He sees you.'</td>
<td>'He sees me.'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W-see-igw-n</td>
<td>3-see-3.INV-OBV</td>
<td></td>
</tr>
</tbody>
</table>

Our interest now is the suffix position closest to the root, known as the theme suffix (underlined in Table 4). In our analysis, this position is the realization of the second projection of v, and it is here that the added probe in Algonquian is spelled out. 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 vI or vII depending on whether the probe was deactivated on the first or second cycle. Thus, the prefix does not have a structural locus as the spell out of a particular syntactic head. Empirically, it seems to be always the case that the partial agreement on vI in direct environments does not receive a dedicated exponent, only none or portmanteau as discussed for Mohawk, suggesting that vocabulary insertion only discharges valued features of a probe on the projection where the probe is deactivated, be that vI or vII.

Spell out of the theme suffix on the other hand is obligatory for transitive verbs in Algonquian across the direct / inverse split, and we propose it is the systematic spell out of a specific syntactic terminal, vII. 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 and linear positions of functional heads, no such generalizations are possible about agreement morphology. This indicates that the spell out of agreement is typically not discharging (in Noyer's (1997) sense) a head, but merely the φ-features on a head (compare Distributed Morphology's fission operation), and discharge of the head is potentially distinct.

Despite being fundamentally distinct, spell out of the vII head interacts with spell out of the core probe in crucial ways. If we suppose (without further justification) that the core probe is discharged before vII, there may or may not be φ-features on vII when it comes to spell out vII itself. In direct contexts vII is the locus of a probe by virtue of second-cycle agreement. In inverse
contexts $v_{II}$ hosts the added probe. Direct and inverse derivations are given in (27), with
the corresponding theme marker in bold caps. Shaded parts show where core probe is discharged.
For each derivation, the remnant on $v_{II}$ after vocabulary insertion of the agreement morphology
for the core probe is indicated in the unshaded box. With this in place, generalizations as to the
form of the theme suffix start to emerge. The default $-aa$ surfaces whenever $v_{II}$ hosts the core
probe, discharged prior to $v_{II}$ itself, leaving a bare head to be spelled out. (29c) is an exception,
marked ‡ in the paradigm, but this we have already discussed; it falls into the class of
portmanteau morphology in [participant] direct contexts (Nishnaabemwin 2→1) and, like the
Mohawk case, we take it to reduce to allomorphy of the core probe in the context of a
[participant] valuation of the same probe on $v_{I}$. In the inverse contexts, the core probe is
discharged on $v_{I}$. The added probe on $v_{II}$ and the $v_{II}$ position itself are then discharged jointly, so
the valuation of the added probe affects the form of the theme marker: $v_{II}$ is realized as $-igw$
when it has an undischarged [u-3] probe, and as $-in$ when it has an undischarged [u-3-1] probe.
Thus in inverse contexts the theme spells out the added probe.

(27) Direct contexts       Inverse contexts (bold = added probe)

a. $v_{II}$ EA $v_{I}$ IA 1→3=1
   ((u3) [3] [u3]—[3] $AA$ [u3]—[3] $IGW$
   [u1]—[1] [u1] [u1]—[1]
   [u2] [u2] [u2]

b. $v_{II}$ EA $v_{I}$ IA 2→3=2
   ((u3) [3] [u3]—[3] $AA$ [u3]—[3] $IGW$
   [u1]—[1] [u1] [u1]—[1]
   [u2]—[2] [u2] [u2]—[2]

c. $v_{II}$ EA $v_{I}$ IA 2→1=2
   ((u3) [3] [u3]—[3] $I$ $u1$ [u1]—[1]
   [u1] [u1] [u1]—[1]
   [u2]—[2] [u2] [u2]—[2]

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 the three languages the added probe is a derivational strategy that appears
only in inverse contexts, which cyclic Agree characterizes, and avoids a violation of the PLC by
Agreeing with the EA. We have proposed that it appears on the projection $v_{II}$ from which second cycle of the core probe takes place in direct contexts, and the theory of section 4.1 formulates the property that permits the addition of the probe and derives why it appears in inverse contexts only.

It is worth briefly comparing the added probe proposal with an alternative: both EA and IA control separate probes, say on $T$ and $v$, whose exponence can make reference to them individually or as fused to yield portmanteau morphology (e.g. Bruening (2001:117ff.) for Passamaquoddy (Algonquian)). This kind of bird's-eye view completely misses the generalizations about the distribution of EA/IA control of agreement that our approach seeks to address. To take the core examples that belie the fusional view of EA/IA control: (i) Only EA or IA controls the core agreement slot, following the direct/inverse split. (ii) The added probe slot is limited to inverse contexts (clearly Mohawk, Basque) and controlled only by the EA (clearly in Mohawk which qualitatively differentiates allomorphy in the $1\rightarrow2\ ku$ combination from the added probe; the complexity of the Algonquian theme obscures the added probe/allomorphy distinction on the surface). These generalizations follow from our approach. The next two sections contribute to this point by finding an identical direct/inverse split in a different domain, Case assignment, and showing how it follows from the same underlying system.

### 4.3 R-Case

In Kashmiri, inverse contexts are special because the IA gets a superficially oblique Case, distinct from the Case it gets in direct contexts, and furthermore it does not control the $\pi$-probe as cyclic Agree suggests. We call the special Case the R-Case. 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 the morphological property P, the same as adds a probe, which converges only in inverse contexts (4.1).

Kashmiri (Indo-European, Wali and Koul 1997) has a nominative-accusative system in the present, and ergative-nominative in the past. Table 10 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. Beside the core agreement slot tracking the underlined argument, there are two clitic doubling series, one for the ergative subject and the unmarked object, which we gloss E/A, and one for the agreeing nominative subject, which we gloss N; there is also an independent dative clitic series, glossed D.

<table>
<thead>
<tr>
<th>Present/future</th>
<th>IA</th>
<th>Dative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present/future</td>
<td>Unmarked, N-clitic</td>
<td>Unmarked, E/A-clitic</td>
</tr>
<tr>
<td>Past/perfective</td>
<td>Ergative, E/A-clitic</td>
<td>Unmarked, N-clitic</td>
</tr>
</tbody>
</table>

The ergative occurs on EAs in the past/perfective, and it behaves as an inherent Case in Kashmiri (Mahajan 1989, Nash 1996, Woolford 1997, 2006, Nichols 2001): it does not interfere with assignment of the nominative (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, crucially, it is retained under passivization:

(28) Mohnas a:yi kəmi:z aslamni zəriyi dini.
'Mohan was given the shirt by Aslam.' (Kashmiri, Wali and Koul 1997:154)

Table 10 omits that there is a case morphologically identical to the dative that is borne by the IA of transitives in the present tense in all and only inverse contexts, as determined with respect to a [$u$-3-2-1] probe:

(29) a. bi chu-s-ath tsi parina:va:n. 1→2, direct
    I.N be.M.SG-1.SG.N-2.SG.E/A you.N teaching
    'I am teaching you.'

b. tsi chu-kh me parina:va:n. 2→1, inverse
    you.N be.M.SG-2.SG.N me.D teaching
    'You are teaching me.' (Kashmiri, Wali and Koul 1997:155)

This is the R-Case. 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. (30)a shows this, in the context of an independent inherent dative, and (30)b shows that R-Case dative disappears in passives, unlike the inherent dative in (28).

(30) a. su kari-y tse me hava:l 3→2, inverse
    he.N do.FUT-2.SG.D you.D me.D handover
    'He will hand you over to me.'

b. tsi yi-kh me hava:l karni təm'sindi də's
    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.' (Kashmiri, Wali and Koul 1997:208)

Only structural Case EA and IA count for establishing an inverse context, as in (29) and (31)a. Neither the dative in (30)a, nor the ergative in (31)b, count, as expected of inherent cases inaccessible to $\phi$-Agree (Nichols 2001), and the IA is nominative.

(31) a. tsi vuch-a-kh me 2→1 present, inverse
    you.N see-2.SG-2.SG.N me.D
    'You will see me.'

b. tse vuch-u-th-as bi 2→1 past, direct
    you.E saw-M.SG-2.SG.E-1.SG.N me.N
    'You saw me.' (Kashmiri, Wali and Koul 1997:156)
In both direct and inverse configurations, the sole agreement slot of Kashmiri tracks the EA. We model this by having this agreement spell out the core probe (see the discussion for Nishnaabemwin, section 4.2). Thus the pattern in Table 11 emerges, where the left-hand side reflects the \( \pi \)-value of core agreement (always the EA), and the right-hand is the Case assigned to the IA: \( \emptyset \) in direct contexts, R-Case in inverse ones.

Table 11: Kashmiri Person Hierarchy effect

<table>
<thead>
<tr>
<th>EA → IA</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1/( \emptyset )</td>
<td>1/( \emptyset )</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2/D</td>
<td>2/( \emptyset )</td>
<td></td>
</tr>
</tbody>
</table>

The Kashmiri pattern is important to us for two 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 EA ends up controlling overt agreement where it should not and in turn the IA fails to do so, and this is coextensive with R-Case assignment to the IA.

The limitation of R-Case to inverse contexts, defined by the configuration of \( \pi \)-features on EA and IA, means that it is radically different from homophonous inherent Case. True inherent Case on a DP is introduced at base-generation, and it is not sensitive to person specification of another DP; it remains under passivization. The sensitivity of R-Case to the interaction of EA and IA agreement features suggests that it is Agree-related, structural Case. The task is then to account for its morphological differentiation from normal IA structural Case, and for its limitation to inverse contexts that gives this Case a last-resort, global economy flavor (cf. Nichols (2001:529 note 14)): it gets assigned to the IA only if the EA is not more highly specified, so the core probe of \( v \) would not reach it. We have proposed that R-Case is assigned when the probe on \( v \) has a property P whose consequences are illegitimate in direct contexts. Since the identity of assigned Case is determined by the identity of the probe, R-Case can be different from the Case assigned by \( v \) without P (Kashmiri "nominative"). This difference is simply not used in Basque, which has case morphology, or Mohawk, which does not. The consequence of P is the addition of a probe to Agree with EA, licensing it for the PLC. In spelling out of the consequences, Kashmiri differs from Mohawk in always spelling out only the highest \( \pi \)-probe of \( v \), and in having a special morphology for R-Case distinct from regular \( v \)-assigned Case. These parametric spell-out choices are encoded in (24).

There exist striking parallels to the R-Case phenomenon beyond PH phenomena, viewing it as the assignment of an atypical structural Case to a DP by a probe \([uF]\) when \([uF]\) "needs" for convergence to Agree with another DP. We discuss these in the appendix: 2-3 Retreat in K'ekchi' (Mayan, Berinstein 1985, 1990, cf. Davies and Sam Colop 1990, Hale 2001), whereby an IA gets a structural dative rather than absolutive just in case the EA needs the absolutive for A'-extraction; and Bobaljik and Branigan's (2006) approach to the ergative in Chukchi and to the causee dative in French which they make dependent on a probe's regular Case assignment (absolutive, accusative respectively) to another DP. These examples suggest a pattern in which an Agree relation between a \( \phi \)-probe and a DP can, because of some property P of the \( \phi \)-probe, allow the probe to Agree with a different DP; cf. further Rezac (2007).
Returning to R-Case in Kashmiri, we have proposed to unite the R-Case strategy there with the added probe strategy because both are derivational strategies coextensive with inverse contexts. R-Case suggests that choice of the strategy occurs at the point the core probe Agrees with the IA, and therefore at $v_I$ before the EA is added. We have thus analyzed the two together as the reflex of a single underlying mechanics, the presence of property $P$ of $v$ and the probe addition in which it results (23), differentiated in spell-out by morphological properties (24) of a language. 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.14

5 Conclusion

We have displayed a pattern of PH-driven agreement displacement, and argued that it follows from simple assumptions about the mechanics of the derivation, (32)a, (32)b, combined with an independently motivated understanding of $\phi$-features, (32)c:

   b. A fine-grained approach to cyclicity giving each operation its own cycle.
   c. A decomposition of $\phi$-features that associates with each person value a different feature structure and thus a different locality class.

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

(33) a. Partial agreement sensitive to specifications of goal arises from interaction between the articulation of the characteristic probe (giving PH-sensitivity: $[u-3]$ probe has no sensitivity, $[u-3-2]$ distinguishes $1^{st}/2^{nd}$ person from $3^{rd}$ person, and $[u-3-2-1]$ 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 IA 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 correspondingly the nature of the repair strategies.

14 Other inverse-only derivational strategies mechanism may exist. The Tanoan languages like Southern Tiwa (Allen and Frantz 1983, hierarchy 1/2 > 3) agree with EA and IA in direct contexts, and in inverse contexts have intransitive-like agreement with IA alone, put the EA in an instrumental-like PP, and add special inverse voice morphology. This strategy is coextensive with inverse contexts and thus derivational, and we would like to see in it another reflex of the added probe mechanism: the inverse morphology reflects a $v$ with an added probe, and the EA PP arises perhaps through selection by this $v$. In a different direction, 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 for it (all transitives are inverse contexts). This raises the interesting possibility that what is normally taken as the $\pi$-probe of T in languages like English is in fact due to $P$, with the accusative as R-Case.
These results 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.

We conclude by noting that the syntactic approach the PH-effects based on Agree is fully compatible with possible syntactic displacement correlates of PH-effects, though it correctly does not require them (section 2). In some Algonquian varieties, syntactic phenomena single out the EA of direct contexts and the IA of inverse contexts, and do so moreover only in clause types where agreement morphology shows the direct-inverse contrast (Rhodes 1994 on Ojibwe independent vs. conjunct orders). If these rely on a designated clausal position, it is occupied by the EA in direct contexts and by the IA in inverse ones (cf. Bruening 2001 for Passamaquoddy). Our Agree-based system can readily model this correlation of positions and Agree controllers given an analysis of movement as Agree followed by (internal) Merge of the goal of Agree (Chomsky 2000, 2005). In direct contexts, the EA controls core agreement, while in inverse contexts the IA does, leading to Merge of the one or the other according to which ends up controlling the core probe. The only mechanics required is linking of the trigger of movement, for example the EPP/OCC feature, to core probe valuation.

6 Appendix: R-Case in K'ek'chi and French

Here we consider the R-Case phenomenon, delimited as the assignment of an atypical Case to a DP by a probe $[uF]$ when $[uF]$ "needs" for convergence to Agree with another DP, that is in our proposal when P triggers the insertion of a second probe to the head that bears $[uF]$. We bring support for the phenomenon from 2-3 Retreat in K'ekchi, and then from Bobaljik and Branigan's (2006) analysis of Romance causativization, which posits the equivalent of R-Case.

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 A'-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:

(34) 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 you-D
    'That's the snake that will bite you.' (Berinstein 1990:10)
We analyze this as a requirement for A'-extraction to proceed from a non-derived position (cf. Rizzi 1982:chapter 4, 2004). Ergative agreement, which blocks A'-extraction of EA, must then be a property of T (Bobaljik 1993), and force externalization of the EA. When the external argument is to A'-extract, the ergative φ-Probe of T must be absent, and the external argument remain in [Spec, vP]. 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. 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 (34)a, but required here, (34)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, (35)a, against both various types of obliques and the dative goals of ditransitives (35)b. However, the dative IAs of 2-3 Retreat clauses can Control (35)c.16

\[(35)\]

a. Ch-o-a-tenk'a chi cuąnc sa' x-yal-al.
   TNS-1.AP-2.E-help in.order exist-INF in its-truth-2.POSS
   'May you help us in order to live in peace.' (Berinstein 1990:27)

b. *X-∅-in-q'ue acu-e₁ chi PRO₁ cuac.
   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)

---

15 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 A'-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). A'-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 φ-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 2-3 Retreat applied without A'-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 A'-extracted, we may suppose that it satisfies the EPP regardless of Agree with T, because such EPP-of-T satisfaction under A'-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).

16 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 1995:272-6); alternatively, it is the extra PP-like structure introduced by inherent Case that blocks such DPs from controlling (Bayer et al. 2001).
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, and spelled-out agreement is controlled by the EA. The IA therefore surfaces as a non-agreeing dative, but one which has entered into Agree with v. A possible reflex of the v-IA Agree is movement to a designated position (e.g. specifier of vP), since Berenstein (1990:14) also shows that 2-3 Retreat datives unlike all other datives are necessarily immediately post-verbal.

Bobaljik and Branigan (2006) 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 common pattern (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. (36) shows this for French: the sole argument of *venir 'come' in (36)a can only be accusative, as the direct object of *acheter 'buy' in (36)b, while the subject of *acheter must be dative (formed with à 'to').

(36) a. Luc a fait *venir les/*aux étudiants
   Luc-N has made to.come the/*to.the students-A
   'Luc has made the students come.'

   b. Luc a fait t1 acheter un livre aux/*les étudiants1
   Luc-N has made to.buy a book-A to.the/*the students
   'Luc has made the students buy a book.'

Bobaljik and Branigan propose a multiple-Case checking mechanics in configurations where a single head F0 (say v) must license two DPs: the closer DP1, the causee, gets a special Case KF-R from its relationship to F0, our R-Case, which leaves F0 free to assign its regular Case KF to the farther DP2, the theme (with obligatory evacuation of DP1 from the path between F0 and DP2 because of locality, cf. Chomsky 2000:131).17

(37) a. Causativized transitive

   __________
   F0   DP1-KF-R   DP2-KF

   b. Causativized intransitive

   __________
   F0   DP-KF

There is thus a convergence of evidence that an Agree relation between a φ-probe and a DP can, because of some property P of the φ-probe, allow the probe to Agree with a different DP; assignment of a special Case by the probe to the IA occurs when this happens. Other properties

17 Neither in French, nor in the (Chukchi) type of ergativity to which Bobaljik and Branigan apply their approach (with T as F0, EA as DP1 (ergative as R-Case) and IA as DP2 (absolutive as regular T-Case)), do PH effects play a role. We expect this behavior whenever the probe is flat: any DP1 would intervene for DP2 and so requires R-Case (cf. Hindi-Urdu -ko). In our assimilation of Kashmiri-type and French-type R-Case, we require that French v not have the option of giving the IA R-Case and Agreeing with the EA in regular transitive clauses; such a derivation would independently crash, since French also has a φ-probe on T that needs to Agree with the EA. This follows for example if F0 is the C of causative complements, a possibility entertained by Bobaljik and Branigan; their alternative where F0 is v works out if it is the v of causatives, but becomes more complicated if it is the regular v of (di)transitives.
of the language determine when this is required, but the impression given is that it occurs only as last-resort to allow a derivation to converge.

References


Morimoto, Yukiko. 2002. Prominence mismatches and differential object marking in Bantu. In Proceedings of the LFG02 Conference, ed. by Miriam Butt and Tracy Holloway King. Stanford, Calif.: CSLI. Available at:
http://csli-publications.stanford.edu/LFG/7/lgf02morimoto.pdf


Nash, Léa. 1996. The internal ergative subject hypothesis. In NELS 26, ed. by Kiyomi Kusumoto, 195-209. Amherst, Mass.: University of Massachusetts, GLSA.


