Title: *Agreement displacement in Basque*
Draft: June 27, 2006.

This work discusses the results of research conducted 2005-2006 on agreement displacement in Basque at the University of the Basque Country (Gasteiz). This is a first, unfinished draft. Comments, perspectives, related works past or new, are deeply appreciated. Though this is the first complete version, it has greatly profited from presentation during a course at the University of the Basque Country (Gasteiz), June 5-9 2006, and I am already greatly indebted for comments that I have been able to take up into the text at this point; the real acknowledgments for help and encouragements, I will add later.

Milan Rezac

Table of Contents

Preface

This is a "testbed" study of a particular view of certain agreement morphology. The view is that certain φ-feature relationships reflect narrow syntactic dependencies established for φ-features, and so reflect, and can inform about, the properties of narrow syntactic computation, such as locality and cyclicity.

The view is pitted against some of the massive variation found in the agreement morphology of more than a hundred local varieties (contiguous within about a half-dozen dialect groupings) of one language, Basque. This gives an up-close look to what happens to an agreement system with particular properties -- say φ-feature based alternation between object and subject control of agreement -- as one moves from dialect to dialect, what agreement phenomena there are, which ones remain constant, which ones vary, and whether they vary according to patterns or essentially arbitrarily. These facts inform the theories of the three phenomena of "eccentric agreement" or "agreement displacements" proposed here, as the expected consequences of specific theory of narrow syntax under within minimalist guidelines.

At the same time, the amount and types of simply seem to call the whole idea into question at first: beautiful orderly paradigms that reflect the proposed syntax, and on which much work of this sort is based, stand beside others that still show the same patterns in far more limited contexts, and beside others too that simply show an apparently senseless mess. This raises a problem that requires attention in the further development of the basic program: parametrization. Parametrization visible to the syntax, at the level of lexical items, say in the φ-feature structure of probes and goals, that yields syntactically real patterns; and parametrization at the level of spell-out, say of vocabulary item insertion, that can yield a mess.

The idea that φ-feature agreement reflects a core syntactic dependency has come into prominence only fairly recently. The essence of the argument, as I see it, runs as follows. Agreement on a target controlled by a goal embedded at a potentially unbounded distance turns out to be subject to the same locality properties as A-movement of the said goal would be, if it moved; for example, there must not be an intervener capable of controlling the agreement (1). This had been handled by covert A-movement of the goal and agreement stated over the resulting configuration. However, the postulated covert A-moved (or chain formation) turned out to have no LF effects, unlike overt A-movement and other postulated covert movement, so it turned out to be just a device to capture the observation that φ-agreement at a distance is sensitive to the same locality properties as A-movement. This leads to an exploration that puts remote agreement at the core and views movement as an additional operation contingent on it, reversing the earlier idea that movement of interpretable material is at the core and remote agreement is its failure to take place overtly.

(1)   a.   In this oven there seem$_i$ to have been baked (*<u>someone</u>/*<u>John</u>) three cakes$_i$.
       b.   In this oven three cakes$_i$ seem$_i$ to have been baked (*<u>someone</u>/*<u>John</u>) t_i.
       c.   In this oven someone$_i$ seems to have been baked t_i three cakes.

The change in perspective, adopted in Chomsky (2000), elevates the status of φ-feature agreement as a source of evidence about syntactic dependencies (all the while keeping in mind

---

₁

that the posited operation, Agree, is clearly not the only source of φ-agreement).The relationship between φ-agreement targets and controllers is often more complicated than the vanilla one-to-one relationship found in English. One target can have multiple potential or even concurrent controllers, and one goal can provide the values for multiple targets. In Basque (in the non-present), a person agreement slot is controlled by the transitive object (absolutive) if it is not 3rd person, and otherwise by the transitive subject (ergative), if the latter is $1^{st}/2^{nd}$ person. This makes a pattern in which control of the agreement morphology oscillates between object and subject according to the 1/2>3 and object>subject hierarchies. Patterns such as this in the target-goal correlation can be explored as a source of evidence about the narrow syntactic computation that was proposed to establish them.

Such a view stands beside two major concurrent approaches to such agreement oscillations, not necessarily meant to account each for the same phenomena. One is the purely morphological approach that posits that when such oscillations of target-controller correlations are found, the target fully agrees with multiple controllers, or that there are multiple targets each agreeing with a single controller that assemble together, and the oscillation is the result of the morphological component that spells out the resulting multiple φ-feature bundles. This is the view of the Potawatomi person hierarchy in Halle and Marantz (1993). Bobaljik and Branigan (forthcoming) develop this approach to the spurious antipassive in Chukchi, and point out one of its predictions: there is no reason why the oscillations should have any synchronic rhyme and reason to them, and rather than basing the oscillation on 1/2>3, one can expect, and one finds in Chukchi, subject>object in in non-participial tenses *3.S>1.S, *2>1, participial tenses additionally *1>2, *1/2>3, *3.S>3, beside object control otherwise. The other approach is to view such oscillations as consequences of movement to positions in the functional architecture for interpretive reasons, developed particularly in the work of Jelinek (overviewed in Jelinek and Carnie 2005), but also Rice and Saxon (2002), Déchaine (2001): for example, 1st and 2nd person must move higher than 3rd person for interpretive reasons (presupposition, discourse participation, etc.), and this movement feeds agreement control. Here too the predictions have been clear: the oscillation should have interpretive consequences, for example for anaphora binding.

The beginnings of the enterprise where patterns of agreement are taken as direct evidence about the properties of the computational system go back before the change in perspective about the establishment of agreement without movement, particularly Laka (1993) and Hale (2001). Ensuing work has made large-scale claims on the basis of such patterns about the building blocks and the structure of the feature bundles that enter into syntactic dependencies, the nature of the "matching" relationship between them, the locality of this relationship, and the delimitation of search-space by the construction of syntactic objects.[2] The research program is developed in the minimalist framework, and seeks to reduce or derive rather than add to existing mechanisms. For all that, its empirical coverage is impressive: in a recent synthesis, Béjar (2003) finds the pattern of Basque object>subject, 1/2>3 oscillation in languages as widely separated from Basque and each other as Karok (Hokan), Georgian (Kartvelian), Erzya Mordvinian (Uralic), and posits the same basic mechanisms underlying the more complex person hierarchies such as the object>subject, 2>1>3 pattern of Nishnaabemwin (Algonquian).

---

[2] This list is long and the references will appear later. I have in mind principally: person-number separation in the Person Case Constraint (Taraldsen 1995, Sigurðsson 1996, Anagnostopoulou 2003, Béjar and Rezac 2003), person-number separation in Basque ergative displacement (Laka 1993, Fernández and Albizu 2000, Fernández 2001), articulation of the person to capture person hierarchies (Béjar 2003), locality conditions over probes that are such articulated feature geometries (Béjar 2003, Béjar and Rezac 2004), cyclic search space expansion (Rezac 2003).

This work is embedded in the context of this research program. I try to push the programme adopted see how far it will go. The view is simple: the properties of syntactic dependencies involving φ-features, reflected as a particular type of φ-agreement, follow from (i) the construction mechanism for syntactic objects where a timing unit is commensurate with the application of an operation; (ii) a concept of locality that follows from a construal of matching as identity; (iii) syntactic parametrization that consists in varying the φ-feature content of probes and goals (cf. work on parametrizing the binding theory, Reuland and Reinhard 1995, Reuland 2001, 2005ab). However, the empirical database has so far emphasized connecting patterns across unconnected languages to understand the system that unites them. This plays some role here, but it is a supporting role to the main theme: a detailed look at agreement displacements across some of the myriad Basque dialects whose principal agreement system has been documented. This leads to the research goals posed at the beginning. First, a theory of the agreement displacements found in these dialects and their bearing on the theory: "ergative displacement", much treated in the literature of this program; "dative displacement", far more rarely; and "absolutive displacement" in Person Case Constraint contexts, undocumented. The other goal of the research, more tedious, but perhaps more necessary, is to see what lies in the details. What are the limits of systematic patterns; how does the syntax go from more general systems to more limited patterns by means of parametrizing the properties of syntactic terminals; how to deal with arbitrary exceptions and gaps at spell-out? What happens to a system as it transits, descriptively, between Basque ergative displacement to the Chukchi spurious antipassive? And are there reasons not to assume that in the end it's syntactic at all, or even farther, that it's not all been memorized?

By no means all such questions will receive insightful answers here. Much of it is furthermore tedious, in different ways. To the theoretician, for much as I have tried to cull empirical discussion into the appendices, annexes, or altogether, a great, great deal of it remains, and often enough, brought to bear on a particular theoretical issue, it leads to little theoretical gain. And to the (generative) Bascologist, whom even the dead ends discussion may fascinate, but perhaps not deciding between labels and head-movement. Both are often confused by a goal that does not sit well with either, to put in a systematic but detailed form what I've been able to learn about Basque agreement displacement, both my own research and notes in technical literature often in Basque, French, or Spanish, as a resource that others may use. I leave the reader to ponder the following point, raised in evaluating the contribution to the development of film of the pre-eminent Czech inventor, Jára Cimrman (see Wikipedia entry):

> O švédském spisovateli, Augustu Strindbergovi, se dočteme, že byl přistižen, jak vyfukuje kouř z dýmky do umývadla s vodou, skoušeje zda touto cestou nevznikne zlato. A my si říkáme, hle jak podivínský nápad se zrodil v genialním mozku! Ale není správnějš si říci, nebyl to Strindberg, kdo experimentálně zjistil, že vypouštěním tabákového dýmu do vody zlato nevznikne? Někdo to totiž vyskoušet musel. Někdo musel slepou uličku lidského poznání ohledat, a ohlásit světu: Tudy ne, přátelé! Dá se říci, že pravým pionýrem slepých uliček by Jára Cimrman. (*Vražda v salónním kupé*, Smoljak et al.)

> One reads of the Swedish writer, August Strindberg, that he was found blowing the smoke of his pipe into a sink filled with water, in order to see whether gold would not be created in this manner. And we tell ourselves, how loony an idea was born in the brain of

a genius! But should one not rather say: was it not Strindberg, who experimentally verified that through the emission of tobacco smoke into water, gold is not created? For someone had to try. Someone had to explore this blind alley of human knowledge, and announce to the world: It's not through here, friends! One can say that among true pioneers of blind alleys was Jára Cimrman.

HOW TO READ PARADIGMS:

The table below is a typical paradigm format used in the text that illustrate most of the conventions used. It is the paradigm of an applicative-transitive with ergative, dative, and absolutive agreement controllers, with absolutive singular (a paradigm called 3VS in the text, as explained there), past tense.
-The columns are for different ergative values
-The rows for different dative values.
-*Italicised* forms are always ergative displacement.
-Underlined forms always have "doubling" of the prefix by suffix, e.g. *n* by *ta*.
-**Bold** forms are in some respect strange.
-"-" represent impossible combinations.
-Nothing in a cell by contrast means that the data has a lacuna at this point.
-In paradigms containing $1^{st}/2^{nd}$ person dative agreement controllers, there are almost always two forms: the one to the left of the coma is a form without "dative displacement" (C3) and the one to the right of the coma is one with "dative displacement"; if either kind of formation does not occur, it is indicated by $\varnothing$.
-In simple transitive paradigms (no dative but an ergative beside an absolutive), "leismo" (C3:leismo) forms, if they exist, are in square brackets, and otherwise not reported. Thus, a 3.S>1.S form like *nau [dit]* indicates that the dialect may use for this combination either the non-leismo form *nau* or the leismo *dit*, while lack of either form, is indicated in the paradigm as *nau* and *[dit]* respectively.

TABLE: 3VS past paradigm of B-V-O:A.

DAT ERG	3.SG	3.PL	1.SG	1.PL	2R	2.PL
3.SG	otzan	otzain	*notzan*	*gotzan*	*sotzan*	*sotzain*
3.PL	osten	ostein	*nosten*	*gosten*	*sosten*	*sostein*
1.SG	$\varnothing$, nostan	$\varnothing$, nostain	-	-	*sostan*, $\varnothing$	*sostain*, $\varnothing$
1.PL	$\varnothing$, **noskun**	$\varnothing$, **noskuen**	-	-	*soskun*, $\varnothing$	*soskuen*, $\varnothing$
2.R	otzun, $\varnothing$	otzuen, $\varnothing$	*notzun*, $\varnothing$	*gotzun*, $\varnothing$	-	-
2.PL	otzuen, $\varnothing$	otzuen, $\varnothing$	*notzuen*, $\varnothing$	*gotzuen*, $\varnothing$	-	-

OTHER CONVENTIONS:
-Combinations of φ-features are indicated in the order ERG > DAT+ > ABS, where a dative always has + (except when I have forgotten, of course). Thus, 1.S>2.P+>3.P is applicative transitive, 1.SG ergative, 2.PL dative, 3.PL absolutive; 2.P+>3.P, is applicative unaccusative, 2.P dative and 3.P absolutive.

DIALECT ABBREVIATIONS: see file DIALECT LIST.

(#) NOTATION/ indicates what I suspect to be errors in the data, either perpetrated by me in creating the database on which this work is based, and which I have not yet had a chance to check, or in the source.

* NOTATION: In chapters C0 and C1, sections that are not relevant to the discussion, but belong there conceptually and will be referred to later, are marked by *.

FOOTNOTES VS ENDNOTES: The discussion uses only footnotes. Endnotes are used exclusively give lengthy citations when I found them useful, and will probably all disappear. If printing the chapters it is easy not to print those pages.

INTERNAL REFERENCES: References to other part of the work take the form of "C2" for "Chapter 2", "Appendix TM" for "Appendix on the Theme Marker", and XN, ANNEX for sections within and annexes to individual chapters, which have not yet been numbered.

EXTERNAL REFERENCES: References to other work are standard, except that: (i) The grammars of Lafon, Lafitte, Azkue are given without year (for Azkue, II refers to volume 2, etc.; § means paragraph number). (ii) References to Yrizar are by dialect name and volume, so Y-G1-234 means Verbo Gipuzkoano, tomo 2, p. 234; Y-D1, Y-D2 refers to the two dialectology volumes of 1981. (iii) References to the grammar edited by Hualde and Ortiz de Urbina are to the version I have, 2001 draft, by section numbers.