

Research report

The Iberian-Caucasian Connection in a Typological Perspective

(Birgit & Gad Rausings Stiftelse för humanistisk forskning)

2000-2001

Arthur Holmer

The purpose of the project has been to examine the nature of the grammatical similarities between Basque and the Caucasian languages (primarily Georgian), and to attempt to determine whether or not these grammatical similarities may serve as evidence of genetical relationship or historical contact. The main idea is the following: assuming a basic probability for any given pair of languages to share certain features, it follows that the more features they can be shown to share, the less probable it is that the similarities are coincidental. If similarities are not coincidental, they are presumably the result of genetic relationship or historical contact.

It has however been shown in typology that certain properties of language tend to be interrelated: a well-known example is that languages with SOV word order almost always have postpositions instead of prepositions. In this respect, what is relevant for any statistical comparison of features between languages is whether or not the extant similarities are interrelated, the intuition behind this being that if the features are not interrelated, it strengthens the possibility of historical or genetic contact (since we are dealing with a greater number of independent similarities). In contrast, if the features are interrelated, we can reduce the number of independent shared features, weakening the probability of historical or genetic contact.

Therefore, the methodology has encompassed the following points:

1) a deep structure analysis of Basque and a Caucasian language (Georgian) with the purpose of determining whether the similarities carry over to the deep structure or only are manifest on the surface;

2) an analysis of the nature of ergativity and other features (in particular multiple agreement) to ascertain whether or not these features are typologically dependent on one another;

3) a typological survey of a sample of other ergative languages (viewed on the surface) to determine whether or not the features shared by Basque and Georgian are also shared by other ergative languages for which no claims of genetic relationship have been put forward;

Point 3) above is intended as a purely empirical study to offset the primarily theoretical focus of points 1) and 2). Briefly, if we are unable to pinpoint a theoretical reason for an interdependence between ergativity and other features, such an interdependence can still be shown to exist if it is

represented in all or a majority of the sample languages. The results of the investigation are given below:

1. Deep structure analysis of Georgian and Basque

1.1 Ergativity

The most salient shared feature between Georgian and Basque is ergativity. Thus, in both Georgian and Basque, the subject of a transitive verb bears a case which is different from that of both the intransitive subject and the object (1)

- | | | | | | | |
|---|----|-------------------------|--|-----------------------|-----------------------|------------|
| 1 | a. | Gizon-ak
man-DEF.ERG | liburu-a
book-DEF.ABS | ikusi
see | zu-en.
3sA.3sE-PST | (Basque) |
| | | ‘The man saw the book.’ | | | | |
| | b. | Gizon-a
man-DEF.ABS | etorri
come | zen.
3sA.PST | | (Basque) |
| | | ‘The man came.’ | | | | |
| | c. | K'ac-ma
man-ERG | c'ign-i
book-ABS | nax-a.
see-AOR.3sg | | (Georgian) |
| | | ‘The man saw the book.’ | | | | |
| | d. | K'ac-i
man-ABS | shemo-vid-a.
in.hither-come-AOR.3sg | | | (Georgian) |
| | | ‘The man came in.’ | | | | |

It is true that Georgian displays a split-ergative pattern, whereby the ergative alignment only occurs with tense / aspect forms belonging to a certain series (including the simple past tense and the optative). This is not a serious objection, however, since it can be shown that the closely related language Laz has a purely ergative system, indicating that split-ergative and ergative systems can easily be derived from each other (in the case of the Kartvelian languages, the ergative system of Laz is a structural simplification of the split-ergative system of Georgian).

An analysis of the ergative systems of Basque and Georgian indicates that the similarity only obtains on the surface. As shown in Holmer (1999) and Holmer & Vamling (forthcoming), the cases ergative and absolutive in Georgian do not correspond to their Basque counterparts. The relation is effectively illustrated as follows:

Structural Case	(Spec-IP) (NOM)	I°-ECM (ERG)	V° (ACC)
Georgian	ABS	ERG	
Basque		ERG	ABS

The important difference between Basque and Georgian is that Georgian absolutive is structurally equivalent to nominative, in that it is only assigned in finite clauses, whereas Basque absolutive is structurally equivalent to accusative (being assigned to the direct object of any verb, finite or non-finite). Further evidence can be seen in the fact that Basque clauses in progressive aspect, formed with the auxiliary *ari*, realize both the subject and the object in absolutive (2a), a fact which is compatible with an analysis of ABS being assigned by the verb (i.e. structurally equivalent to accusative) but not with an ABS-as-NOM analysis. In contrast, in Georgian, there are no constructions with double absolutives. Instead, the object in a corresponding construction is realized in DAT (2b).

- 2 a. Gizona ardoa edaten ari zen
 man-ABS wine-ABS drink-IPF PROG 3sA.PST
 ‘The man was drinking wine.’
- b. K'ac-i c'ign-s xed-av-s.
 man-ABS book-DAT see-IPF-3sg
 ‘The man sees the book.’

Interestingly enough, the double-absolutive pattern found in Basque can also be found in the isolate Burushaski (Pakistan), cf Lorimer (1935). Thus, as far as the case-assignment pattern is concerned, Basque is more akin to Burushaski (3) than to Georgian. More data about Burushaski will be shown in section 4.1.

- 3 a. jε ma mAsqaiyΔm (Lorimer 1935b:12)
 1sg-(ABS) 2pl-(ABS) 2pl-kill-1sg
 ‘I will kill you(pl).’
- b. i:se pfut jε ma:r d-i:-uʃ-Δm
 that spirit-(ABS) 1sg-(ABS) 2pl-for D-3sgMASC-turn.out-1sg¹
 ‘I’ll turn out that spirit for you.’ (Lorimer 1935a:278)

¹ For an explanation of the gloss “D”, cf section 4.1.3.

1.2 Active configuration

Both Georgian and Basque display an active configuration as far as ergativity is concerned: in both languages, there are intransitive verbs which behave as if they were transitive, in realizing the subject in ergative case (4). Furthermore, the types of verb involved are quite similar, being those which express activities and certain bodily functions (such as sleep, yawn, sneeze etc).

- 4 a. Amaia-k dantza-tzen du. (Basque)
 Amaia-ERG dance-IPF 3sE.3sE.PRES
 ‘Amaia dances.’
- b. Nino-m da-amtknar-a. (Georgian; Harris 1981:40)
 Nino-ERG PRF-yawn-AOR.3sg
 ‘Nino yawned.’

While this type of system (termed split-S by Dixon 1994) represents a distinct sub-class of ergative systems, the distinction between two types of intransitive verb classes (unergative and unaccusative) can be found in many other languages. Thus, in German (and in other languages with auxiliary selection), unergative verbs (such as *arbeiten* ‘to work’, *tanzen* ‘to dance’, *gähnen* ‘to yawn’) are construed with the auxiliary *haben* (as are transitives), whereas unaccusative verbs (such as *fallen* ‘to fall’ and *kommen* ‘to come’) are construed with the auxiliary *sein*. This, while the active system is shared by Basque and Georgian (and here has an effect on case marking), it reflects a distinction between unergative and unaccusative verbs found in a very wide variety of languages (it could be said that German and French have an “active” system of auxiliary selection). This distinction is expressed in Government and Binding terms such that the thematic roles of Agent and Patient are projected in different syntactic positions (specifier and complement of the verb phrase, respectively).

1.3 Polypersonalism

Both Basque and Georgian display multiple agreement on the verb: the verb agrees with the subject, the direct object, and the indirect object (in some cases). This feature is shared with a substantial number of other Caucasian languages. This phenomenon is referred to in the Caucasian literature as *polypersonalism*.

This, combined with ergativity, is one of the strongest hints that there may be a connection between Basque and Caucasian languages (since both features are otherwise nonexistent in Europe). For this reason, a substantial part of the work has gone into comparing the Basque and Georgian agreement systems.

Both Basque and Georgian realize multiple agreement as a combination of prefixes and suffixes. However, here the similarities between the systems end. Basque consistently cross-

references absolutive arguments with prefixes and dative and ergative arguments with suffixes². Georgian, on the other hand, cross-references 1st and 2nd person arguments with prefixes and 3rd person arguments with suffixes (with the exception of datives, which are also cross-referenced with prefixes).

Basque: ABS - ROOT - DAT - ERG

Georgian: 1/2 - DAT - ROOT - 3

In Georgian, therefore, case and grammatical role distinctions are not reflected in the position of the affix, rather in its form: subjects are cross-referenced by one series of prefixes (e.g. *v-* for 1st sg.) and objects by another series of prefixes (e.g. *m-* for 1st sg.). Further, the *v*-series is used for both transitive and intransitive subjects, implying that the alignment of the agreement system is accusative³. The 1/2 position in the Georgian verb comprises only one slot. If both subject and object are 1st or 2nd person, the two prefixes cannot co-occur, and the prefix cross-referencing the subject is generally omitted.

Thus, the criteria determining the choice of form and position of affixes differ greatly in Basque and Georgian:

	FORM	POSITION
Basque:	case	case
Georgian:	grammatical role	person

Clearly, the polypersonalism of Basque and Georgian represents two quite different systems. What is shared is the fact of polypersonalism itself.

In section 3), a possible interdependence of polypersonalism and ergativity is addressed. In section 4) dealing with empirical surveys, it will further be shown that polypersonalism cooccurs with ergativity in a wide variety of languages.

² The only exception is the phenomenon of *ergative displacement*, which occurs in the past tense when the absolutive argument is 3rd person, in which case the ergative argument is cross-referenced by a prefix instead.

³ In a subset of tense/aspect categories, including perfect and pluperfect, both case and agreement are subject to *inversion*. When inversion takes place, the grammatical marking of subjects and objects is reversed: objects surface in NOM, and are cross-referenced by the *v*-series of affixes (i.e. subject agreement), while subjects of transitives and unergatives surface in DAT, and are cross-referenced by the *m*-series of affixes (i.e. object agreement). Subjects of unaccusative verbs surface in NOM and are cross-referenced by the *v*-series of affixes, behaving in this respect like objects. Thus, in the tense/mood categories subject to inversion, the alignment of the agreement system is ergative - otherwise it is accusative.

1.4 Prefix agreement

Both Basque and Georgian use a combination of prefixes and suffixes to realize verb agreement. In Europe in general (with the exception of other Caucasian languages), verb agreement is always realized by suffixes alone. For this reason, prefix agreement is a feature which in the European area is only shared by Basque and the Caucasian languages, making it an interesting possible hint of a shared grammatical system. However, the similarity again seems to be a surface phenomenon: as we have seen above, the actual function and distribution of the prefixes differs greatly, the shared feature being the mere existence of prefix agreement.

What remains to be investigated, then, is the possible origin of prefix agreement. Agreement in general is often historically the result of the fusion of clitic pronouns with the verb - this development can occasionally be observed taking place synchronically, such as in the Austronesian languages. Hence, prefix agreement as such is a possible consequence of the existence of proclitics (a feature which is shared by languages in the Mediterranean area in general)⁴.

1.5 Paradigm trimming

In both Georgian and Basque, a verb can agree with both a direct object and an indirect object. In both languages, this is only possible if the direct object is 3rd person. This holds quite strictly in Basque, but there are variants of Georgian where it does not hold (cf Vamling 1988). In variants of Georgian where the direct object is not 3rd person, the indirect object must be. Thus, both systems seem to share the feature that it is impossible for verbs to cross-reference two objects unless at least one of them is 3rd person (and presumably unmarked, 3rd person object agreement usually being the least marked, as it is in both Georgian and Basque). As such, it constitutes a restriction on the amount of marking a verb may take. This phenomenon is referred to in the literature as *paradigm trimming* (cf Addis 1993).

This seems to be a universal. In languages with clitic pronouns, such as Spanish, the same restrictions seem to hold for the clitic system. Thus, it is impossible to express a third person dative argument and a non-third person argument by clitics (5a, b). Neither is it possible to combine two clitics if neither is 3rd person (5c). The only possible combination of clitics is as in (5d), where the direct object is 3rd person, or as in (5e) where both objects are 3rd person.

- | | | | | | | | | |
|---|----|------------------------|-------|---------|----|------------------------|-------|---------|
| 5 | a. | *Te | le | dí. | b. | *Le | te | dí. |
| | | 2sOBJ | 3sDAT | gave-1s | | 3sOBJ | 2sACC | gave-1s |
| | | ('I gave you to him.') | | | | ('I gave you to him.') | | |

⁴ The development of a separate category of agreement from clitics may reflect a possible ergative preference for polypersonal agreement systems (discussed in sections 3 and 4.4).

- c. *Él me te dió. d. Él me lo dió.
 he 1sOBJ 2sOBJ gave-3s he 1sOBJ 3sACC gave-3s
 ('He gave me to you / you to me.') 'He gave it to me.'
- e. Se lo dió.
 3s.REFL.ACC 3sACC gave-3s
 'he gave it to him...'

Thus, paradigm trimming does not seem to be valid evidence for genetic or historical contact, since it represents a universal.

1.6 Preverbal *wh*-slot

Both Basque and Georgian obligatorily place a *wh*-word immediately before the main verb, regardless of whether this position is clause-initial or not. Note that in Germanic languages the *wh*-word is also immediately followed by a main or auxiliary verb, but here the *wh*-word is always clause initial. What is particularly salient in both Basque and Georgian is that this position is not necessarily clause-initial (6).

- 6 a. Liburua **non** dago?
 book-(ABS) where? 3sA-be
 'Where is the book?'
- b. C'ign-i **sad** aris?
 book-ABS where be-3s
 'Where is the book?'

There does not seem to be any uniform analysis of this phenomenon. Empirically, it is not restricted to these languages, but is found in Turkish as well (cf Kennelly 1999), and optionally (but not obligatorily) in Mongolian, both of which are languages with SOV word order (as are Basque and Georgian).

Structurally speaking, a pre-verbal *wh*-position seems to indicate that there is at least one level in the language which is head-initial. This is because a fixed linear relation between a *wh*-position and the verb can be most easily derived by placing the *wh*-word in the Specifier of a given phrase XP and the verb in its head X (in which case the correct linear ordering only results if the complement of X is to the right of X, i.e. if XP is head-initial). Researchers working on Basque agree on this point, but differ as to the nature of the head-initial phrase (Ortiz 1989 refers to it as CP, whereas Laka 1994 refers to it as NegP or Σ P).

Given that the languages which display this ordering are otherwise consistently head-final, it is remarkable that they should be characterized by the existence of exactly one head-

initial projection. However, there seems to be some kind of empirical connection between head-finality and the hierarchically speaking relatively low level of the phrase responsible for the interpretation of wh-relations (“low” in the sense that the clause contains more material outside -and to the left of- the wh-word).

It is clear, at any rate, that the linear ordering, while unusual in Europe, is quite common for head-final (or SOV) languages which have a fixed wh-position (i.e. which in generative terms have overt “wh-movement”). Thus, the pre-verbal wh-position does not serve as convincing evidence for a particular Basque-Caucasian connection, but is rather connected with the SOV order found in these languages (regardless of how the construction is to be accounted for).

1.7 SOV word order

Basque and Georgian share SOV word order as a feature. This, again, is not so common in Europe. However, SOV word order is statistically the most common word order pattern in the world (figures vary, but most estimates place SOV, verging on 50% of the world’s languages, in first position, closely followed by SVO, with approximately 45% of the world’s languages). In fact, the Caucasus lies within a geographical belt stretching from Siberia to Turkey, Persia and India where virtually all indigenous languages have SOV word order.

The fact that Basque is SOV is symptomatic of its lack of relationship with the head-initial Indo-European languages which surround it (i.e. French and Spanish), but not of any distinct relation with Caucasian languages.

2. The structure of Georgian verb morphology

Since Marantz (1984) and Baker (1988) it has become widely, although not universally, accepted in generative circles that morphology and syntax are interrelated phenomena, more specifically, that the order of morphemes realized on an inflected verb reflects the hierarchical order of the phrases in the clause. Intuitively, this can be described such that the verb raises from head position to head position within the clause, receiving new morphological material at each step: if a given functional head is relatively low in the hierarchy of the clause, the verb will pass through the position relatively early in the derivation, and thus the corresponding morphology will be realized relatively close to the root. Conversely, if a functional head is relatively high, we expect the corresponding morphology to be realized correspondingly far from the root of the verb. This interrelation is termed the *Mirror Principle* in Baker (1988).

While the complexity of the Georgian verbal system makes it difficult to unambiguously determine the structure of the Georgian clause, certain salient features of the inflection system can be outlined.

Firstly, it seems clear that there are two levels responsible for subject agreement, the level corresponding to 1/2 person subject prefixes, and that corresponding to 3rd person subject suffixes (the suffix agreement category can conceivably be divided into two categories, since it

seems to be comprised of two slots: the first slot typically distinguishes 3rd person from 1/2 person, whereas the second slot distinguishes 3rd person singular from 3rd person plural in certain tense forms, as well as cross-referencing plurality of any argument). Note that the hypothesis that there are at least two levels has nothing to do with the realization of morphology as prefixes or suffixes: rather, it has to do with cooccurrence restrictions: a 1/2 person object morpheme can not cooccur with a 1/2 person subject morpheme, since they occupy the same slot - on the other hand, a 1/2 person object morpheme can cooccur with a 3rd person subject suffix, suggesting that these two categories do not interfere with each other, if the subject is 3rd person.

Secondly, the order of morphemes indicates that suffixed agreement is generated by a functional head (or heads) which are hierarchically higher than Tense - this is evidenced by the fact that the suffixed agreement morphemes are more peripheral than the most clearly identifiable tense morpheme *-d-* 'IMPERFECT' (7).

7	a.	v-q'ep-d-i	b.	q'ep-d-a	c.	qep-d-nen
		1sg-bark-IMP-1/2		bark-IMP-3sg		bark-IMP-3pl
		'I barked'		'S/he barked.'		'They barked.'

It might be argued that Mood, when distinguishable from Tense, is either at the same level as suffixed agreement, or higher, since differences in Mood - in particular between the simple past (Aorist) and the Optative - is reflected by differences in the values of the agreement morphemes. Thus, the Aorist morpheme can vary according to the person of the subject⁵:

8.		<u>1/2</u>	<u>3</u>
	AORIST	-e	-a
	OPTATIVE	-o	-o

However, even here we find cases where agreement is more peripheral than the vowel marking Aorist or Optative, cf (9).

9	a.	m-nax-o-s	b.	m-nax-o-n
		1sgO-see-OPT-3sS		1sgO-see-OPT-3pS
		'May s/he see me.'		'May they see me.'

To summarize, suffixed agreement is located in a relatively high position in the clause structure. As far as prefixed agreement is concerned, it is more difficult to study its interaction with any other category, since the only other inflectional category found in prefixed position is the

⁵ Example 8 is a representative subset of the type of variation found. For other verbs the corresponding values are *-i* (1/2 person) and *-a* (3 person).

preverb which indicates aspect or Aktionsart (and which is sometimes harnessed for temporal distinctions), cf (10).

10	a.	v-c'er 1sg-write 'I write.'	b.	da-v-c'er Preverb-1sg-write 'I will write.'
----	----	-----------------------------------	----	---

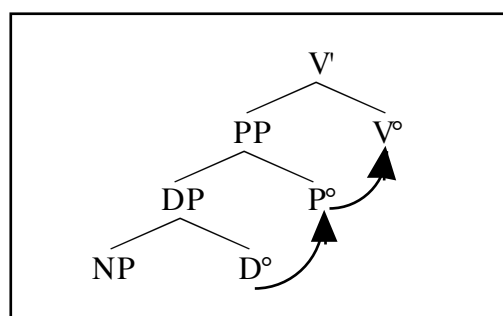
Nash-Haran (1992) suggests that the prefixed agreement morphemes are clitics rather than true agreement, basing her argument on a theoretical incompatibility between prefixation morphology and head-final structure. In the framework adopted here, such a theoretical incompatibility is not recognized. Nevertheless, other facts in the language seem to point to a cliticization or incorporation analysis:

1) Since prefix agreement is less peripheral than the preverb, it would seem to be generated in a position which is remarkably -perhaps even impossibly- low in the structure. Hence, it is preferable with an analysis where the relevant features of the subject are incorporated into the verb at the beginning of the derivation.

2) The fact that objects are cross-referenced in the same slot in the verb supports the hypothesis that they are incorporated at the same level as subject agreement. The fact that objects agreement excludes subject agreement suggests that features of the object are incorporated first, blocking the subsequent realization of any subject features incorporated into the verb. This ordering is what would be expected if agreement consists of features incorporated into, rather than cross-referenced by, the verb.

3) Various types of applicatives in Georgian (known in the literature as *objective* and *superessive version* respectively) are plausibly analysed as incorporation of a preposition into the verbal head (following the mainstream analysis of applicatives adopted by Baker 1988). The version affix differs in form depending on the person of the indirect object, suggesting that version morphology is the result of incorporation, first of the D° (determiner) head of the object NP into the P° head of the prepositional phrase, and subsequently of the entire complex P° into V°. The assumed derivation is shown in Fig 1.

Fig 1. Object feature incorporation in applicatives



4) Finally, it seems a remarkable coincidence that the types of NP represented by prefix agreement are exactly those which do not distinguish morphological case (namely 1/2 person pronouns in ERG, ABS and DAT). Contrary to the situation in English and most other languages which only have morphological case marking for a subset of argument types (where only personal pronouns are overtly case-marked), in Georgian all NP's, with the notable exception of 1st and 2nd person pronouns, have morphological case. 1st and 2nd person pronouns only distinguish morphological cases which are not cross-referenced by verb agreement. In terms of an incorporation analysis, this remarkable situation is in fact expected, assuming that the D° head, which is the head which generally bears the case assigned to the NP, is the one whose features are incorporated onto the verb.

For the above reasons, we discount data for prefix agreement in our analysis of the functional categories of the clause. Instead, we note that Georgian has one (or possibly two) categories of suffixed agreement responsible for subject agreement (on an accusative basis) and this category is located higher than Tense and probably also higher than Mood.

The situation in Georgian can be compared fruitfully with the situation in Basque. In Basque, as we noted in section 1.3, the position of the affix corresponds to the case role of the argument it refers to, and does not depend on the category of person. It is possible to posit two levels of ergative agreement in Basque (at least historically), one of which differentiates between 1/2 plural and other persons, while the other specifies the person more exactly, cf (11a-c), (similarly to the situation in Georgian), and two levels of absolutive agreement, one of which differentiates between singular and plural absolutes, whereas the other specifies the person more exactly, cf (11d-f), (again, in analogy with Georgian). In the following discussion, the agreement category which does not exactly specify the person will be referred to as *vague* agreement, for want of a better term.

- | | | |
|----|---|---|
| 11 | a. g-en-u-en
1pERG-1/2plERG-have-PRET
‘we had ... it’ | b. h-u-en
2sERG-have-PRET
‘thou hadst it’ |
| | c. z-u-en
3sERG-have-PRET
‘s/he had ... it’ | d. ga-it-u-zu
1pABS-ABS.PL-have-2s.POL.ERG ⁶
‘you have ... us’ |
| | e. d-u-zu
3ABS-have-2s.POL.ERG
‘you have ... it’ | f. d-it-u-zu
3ABS-ABS.PL-have-2s.POL.ERG
‘you have ... them’ |

Interestingly enough, the more specific type of agreement in Basque is more peripheral than the vague type, suggesting that its hierarchical position in the clause is higher than the vague type. This is the diametrical opposite of the situation in Georgian, assuming that we have interpreted the Georgian data correctly.

Furthermore, it is clear that the PRET morpheme *-en-*, when present, is more peripheral than any type of agreement, suggesting that Tense is hierarchically above all categories of Agreement in Basque⁷, as opposed to the situation in Georgian.

For the reasons outlined above, it is clear that the structure of the Georgian clause is in many ways different from that of the Basque clause. At the same time, there are surprising parallels, which are the more remarkable since they seem to indicate that certain features of Georgian are the mirror image of corresponding features in Basque. It is unclear what weight can be given to such parallels, but it should be noted that the distinctions they involve (singular vs. plural and 1/2 vs. 3rd person) are cross-linguistically frequently involved in agreement pattern asymmetries. As will be discussed in section 4.4, it is conceivable that the existence of various parallel levels of agreement is symptomatic of a general tendency for ergative languages to favour polypersonal agreement systems, and a further tendency to renew such agreement systems by whatever means available, if loss of relevant distinctions at one level should render this necessary.

3. *The nature of ergativity*

In previous and ongoing work, I have noted that the most straightforward analysis of various types of ergativity involves the interaction of two separate factors, namely:

⁶ POL indicates that this is a polite reference to the 2nd person singular, similar to German *Sie*.

⁷ Similar arguments supporting this view can be found in Laka (1993, 1994).

- a) in an ergative language, the head I° (or the lowest head of a split inflection phrase) is capable of assigning case to the specifier of its complement;
- b) in an ergative language, there is no intermediate category between I° and the VP; thus the case assigned by I° is always assigned to the Agent;

If both of these are satisfied, the alignment of the language is ergative. If either of these is not satisfied, the alignment of the language is accusative. If a) is satisfied, but not b), the result is that case is assigned to a position structurally below the finite verb, but that the argument occupying this position is not necessarily an agent - this produces an accusative language with (prototypically) VSO word order. This is, with slight modifications, the analysis generally adopted for Celtic languages such as Irish.

The two parameters presented above can display a total of 4 different permutations of values. This would imply that, all things being equal, the probability of an ergative alignment occurring in a language should be approximately 25%. Interestingly enough (although possibly coincidentally), estimates of the number of ergative languages in the world suggest a figure around 25% (cf Dixon 1994).

The various types of ergativity found are modifications on this theme, and are covered in detail in Holmer & Dooley-Collberg (in preparation) - parametric differences include whether or not there is a discrete case NOM assigned in a finite clause, whether or not an Agent can raise from a position where ERG would be assigned to a position where NOM is assigned, whether or not the assignment of any case can be blocked (as ERG is by the thematic suffix *-av / -ob / -eb* etc in Georgian). These parameters are capable of accounting for the various systems of ergative and ergative-like languages which are found in the world.

It should be noted that syntactic typology does not recognize any strict dichotomy between ergative and accusative *languages* - rather, certain *constructions*, or certain features of a language, can display an ergative alignment. In fact, most ergative languages have a substantial number of accusative properties as well. Thus, while accusative languages are generally rather uniform in their syntactic alignment, ergative languages (or languages displaying some kind of ergative characteristics) are a very heterogeneous group. And nowhere is this heterogeneity more evident than in connection with the distribution of subject properties.

In an accusative language the subject is syntactically the sole obligatory argument (if there is one), the argument which primarily is cross-referenced by verb agreement, the argument which bears the morphologically least marked and distributionally most common case form (nominative). It is also semantically the prototypical agent, the entity which is assumed to be capable of control and volition and the entity which is most commonly human and/or animate. Clearly, then, accusative languages like German and English have subjects - all subject properties are concentrated on a single argument in the clause.

One important characteristic of an ergative alignment is that the concept of *subject* is diluted - the case-marking system no longer selects the subject as a privileged argument. Subject

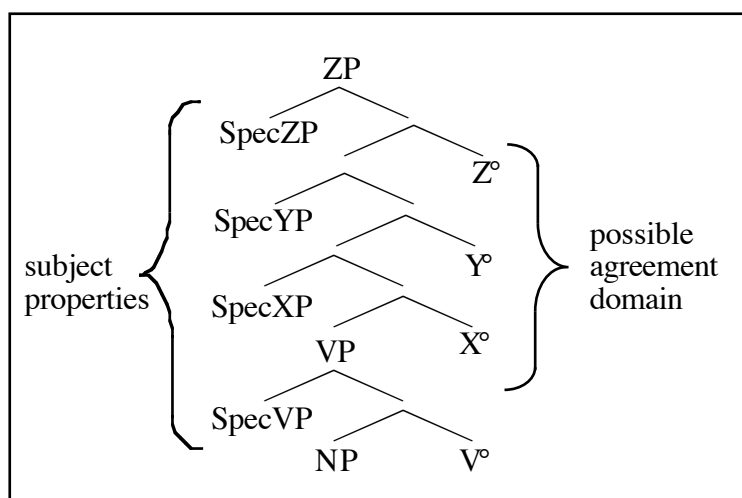
properties are to a greater extent spread out among the arguments in the clause. This also implies that there is, among the arguments, no obvious single candidate for preferential treatment with respect to verb agreement. Rather, we expect verbal agreement to be more evenly distributed among the arguments of the clause than would be the case in an accusative language like Latin or German.

In generative terms, subject properties are often described as properties relating to Specifiers of phrases. For example, Guilfoyle, Hung & Travis (1992) discuss two types of “subject position” in Austronesian languages responsible for the split in subject properties in these languages (particularly the Philippine languages and Malagasy). They argue that some subject properties concern the Agent position SpecVP, whereas other subject properties concern the “Topic” position SpecIP, and the split in subject properties is caused by movement of the Patient to SpecIP, resulting in both arguments being located in Specifier positions.

In contrast to the situation in accusative languages, where all evidence suggests movement of the single “subject” through each of the functional Specifier positions in turn, languages with ergative characteristics (and ensuing subject property splits) are crucially assumed to involve the movement of *more than one argument* into the functional domains of the clause. Since agreement as a category is based on grammatical relations in the clause as a whole rather than on argument structure (typical evidence includes the realization of agreement on auxiliaries and raising verbs rather than on the main verb), it is also part of the functional domain of the clause (the result of an agreement relation between a functional head and a Specifier housing the argument in question). If, then, ergative languages prototypically involve the multiple movement of arguments to the functional domain of the clause, the prerequisite for multiple agreement (multiple Spec-head relations within the functional domain) is satisfied, and it is to be expected that such languages should tend to develop polypersonal agreement systems.

An example of a possible structure is given in Fig. 2, highly simplified and abstracted away from language-specific particularities. It shows the relation between split subject properties and a tendency to polypersonal agreement (in an accusative language with no subject property splits, the prototypical position for the object is the NP position at the bottom of the structure, which is outside the domain where agreement is expected). Note that the labels XP, YP and ZP are intended for convenience of exposition, in an actual given language they would correspond to Tense, Agreement, Mood, or other morphological categories.

Fig. 2. The relation between verb agreement and subject properties



Thus, grammatical theory -whether general or specifically generative- leads to the expectation that ergative languages should be more liable to polypersonal agreement than are accusative languages. Note that this is question of general tendencies - it is certainly not impossible for accusative languages to have polypersonal agreement (Swahili being a case in point), not is it unknown for ergative languages to have single agreement (Kurdish being a typical example⁸). Nevertheless, a cursory overview of a wide spectrum of languages gives the impression that there is such a distributional tendency. We will see more evidence for this in the following section, where a typological survey of a small sample of 3 ergative languages is discussed - all have polypersonal agreement. In contrast, a common, possibly prototypical, scenario for an accusative language developing agreement is that it would tend to develop single (subject) agreement. This is for instance the case in the Mongolian language Kalmyk, spoken to the West of the mouth of the Volga, in southern Russia⁹.

The structural account also clearly illustrates the contrast between the relative homogeneity of accusative languages and the heterogeneity of ergative languages. If a language type involves movement of several arguments to the functional domain of the clause (as is the case in ergative languages), there are various possible alternatives as to which argument passes through which Specifier position, leading to various possible constellations of subject properties. If a language only involves movement of a single argument through all the functional Specifiers of the clause in turn (as in an accusative language), the range of possible variation is greatly reduced, leading to greater typological homogeneity. Thus ergativity, structural heterogeneity and polypersonal agreement are structurally expected show some kind of statistical correlation, as are accusativity, structural homogeneity and single agreement. The expectations are borne out by the samples presented below, but cannot be confirmed statistically without an investigation of a

⁸ Interestingly enough, in the Kurdish ergative pattern, the verb agrees with the *object*, not with the subject.

⁹ Jan-Olof Svantesson (p.c).

much larger sample of both accusative and ergative languages, which is outside the scope of this study.

4. Typological survey of a sample of ergative languages

The sample of ergative languages chosen for comparison comprises the isolate Burushaski (Pakistan), the Paleo-Siberian language Chukchi (Russia) and the Eskimo-Aleut language Inuit (Greenland). These have been chosen to avoid Caucasian areal features. The aim of this survey is simply to examine whether or not they empirically seem to share features which characterize Basque and Georgian. No in-depth analysis of the parameters determining the grammatical systems of these languages has been made, except in certain crucial cases.

4.1 Burushaski

4.1.1 Basic typological facts

The data on Burushaski is based primarily on Lorimer (1935a) and certain examples are quoted from Lorimer (1935b). Since the approach is corpus-based rather than elicited, some unfortunate gaps occur which make certain generalizations impossible to confirm. However, the basic facts are as follows.

Burushaski has unmarked word order SOV with an optional marked order OSV (with an emphasized object). It has three basic cases: ABS, ERG and GEN, with GEN serving as the basis for further case suffixation such as dative and various postpositional forms. The three cases are only wholly differentiated with nouns or pronouns of FEM gender - with other genders, ERG and GEN syncretize. With some pronouns (particularly plurals), all three cases syncretize. Furthermore, since Burushaski allows omission of arguments cross-referenced by the verb, it is difficult to find authentic examples where the case alignment is unambiguous. However, this much is clear:

In tenses and moods derived from the “Past Base” (formally identical with the verb root), the alignment is ergative. The subject of a transitive verb is realized in ERG, whereas the subject of an intransitive verb is realized in ABS. In tenses and moods derived from the “Present Base” (in most cases the root plus a formative suffix - usually involving *č*) the alignment is accusative - that is to say, both subject and object are realized in ABS (there being no distinct ACC case)¹⁰. However, it appears from a great number of examples that the accusativity of a present-base clause is optional, thus, it is not at all uncommon for the subject of a present-base transitive verb to be realized in ERG (12a). In contrast to Basque and Georgian, Burushaski does not have an active alignment, as evident from (12b), where a typical unergative verb takes an absolutive subject:

¹⁰ This is illustrated in example 3, section 1.1.

- 12 a. ja λkəɾ ε-sqai.-λm (Lorimer 1935a:67)
 1sg.ERG self-(ABS) 3sgINAN-kill-1sg
 ‘I will kill myself.’
- b. ne hir yált-i (Manning 1996:3)
 DET.MASC man yawm-3sM
 ‘The man yawned.’

With passive participles the agent is realized in either ERG or GEN - since these syncretize for all genders except FEM, and since no examples were available with a FEM Agent, it is impossible to determine whether the general rule requires ERG or GEN. If the former, Burushaski patterns with Basque, if the latter, with Georgian.

4.1.2 The Burushaski verb

The verb has a wide range of tense/aspect/mood (T/A/M) forms, based either on the Past Base or on the Present base. There is a clear parallelism between some of the formations (many of which are periphrastic and make use of the auxiliary):

Fig. 3. Burushaski tense patterns

	PAST BASE	PRESENT BASE
+m	simple past tense	future tense
+AUX (pres)	perfect tense	present tense
+ AUX (pret)	pluperfect tense	imperfect tense

Infinitives, imperatives, gerunds and nominalizations of various kinds are generally formed from the Past Base. All T/A/M forms are formed by suffixation.

4.1.3 Agreement in Burushaski

There are two agreement slots in the Burushaski verb, one prefix slot and one suffix slot. The suffix slot occurs on the verb itself if the conjugation is synthetic, and on the auxiliary, if the conjugation is periphrastic. The prefix slot is not available for all verbs, although it seems to be the rule rather than the exception for transitive verbs. The distribution of the affixes is illustrated schematically below:

- 13 a. **(PREF)-V** **AUX-SUF**
 Λ-či bai-i (Lorimer 1935a: 267)
 1sOBJ-give AUX.PRES-3sMASC
 ‘He has given it to me.’
- b. **(PREF)-V-SUF**
 a-ku-či-č-Λm (Lorimer 1935a: 266)
 NEG-2sOBJ-give-IPF-1sSUBJ
 ‘I will not give it to you’

The suffix always cross-references the subject of the verb, regardless of whether the verb is transitive or intransitive. The prefix, on the other hand, cross-references the object of a transitive verb and the subject of an intransitive verb. With intransitive verbs (of the type that do have prefixed agreement) the subject is cross-referenced twice in the same form (this is presumably the reason why the prefix slot is not so common with intransitive verbs):

- 14 a. gu:-ir-č-um-a (ibid.:218) b. u.-i:r-um-Λn (ibid.:219)
 2s-die-PF-M-2s 3p-die-M-3p
 ‘thou wilt die’ ‘they died’

Another important difference between the categories of prefix agreement and of suffix agreement is that the prefixes can be found on non-finite forms of the verb (participles and infinitives), whereas the suffixes are usually only found on finite forms (one possible exception being what Lorimer refers to as the “Static Participle”, which is of doubtful status: the only differentiated suffixed agreement morpheme is in the 1st person singular, and this form is identical to the past tense form). Relevant examples of prefix agreement with non-finite forms are given below (all examples are quoted from Lorimer 1935a:277; the D refers to a verbal prefix which is lexically determined and which regularly precedes prefix agreement):

- 15 a. dU-mu:-is-Λs b. dU-ku:-is-Λs
 D-3sFEM-take.out-INF D-2sg-take.out-INF
 ‘to take her out’ ‘to take thee out’
- c. d-i:-ts-un d. dU-mU-ts-un
 D-3sMASC-bring-PST.PRT D-3sFEM-bring-PST.PRT
 ‘having brought him’ ‘having brought her’

Another typical feature of the prefix agreement category is that it is identical in origin (and often in form) to possessive prefixes found on certain classes of nouns (including body-parts etc):

16	a.	a-riŋ 1sg-hand 'my hand'	b.	gu-riŋ 2sg-hand 'thy hand'	c.	mu-riŋ 3sgFEM-hand 'her hand'	(op.cit:129)
	d.	a-r-ΛS 1sg-send-INF 'to send me'	e.	go-r-ΛS 2sg-send-INF 'to send thee'	f.	mo-r-ΛS 3sgFEM-send-INF 'to send her'	(op.cit:270)

Thus, it seems clear that the polypersonal system in Burushaski consists of two quite different categories, only one of which is verbal agreement in the most prototypical sense of the word, the other being more reminiscent of object cliticization and/or possessive marking.

4.1.4 Paradigm trimming

Lorimer presents no data which bears immediate reference to the question of paradigm trimming: however, with the verb *-čīΛS 'to give' the prefix series refers to the indirect object, not to the direct object, which suggests that verbal object agreement does not cross-reference a direct object if there is an indirect object present.

17.	čΛp	a-ku:čī-č-Λm	(op.cit.:266)
	meat	NEG-2sg-give-IPF-1sg	
		'I shall not give thee the meat.'	

4.1.5 Wh question formation

The final point we touch upon is the word order in wh-questions - most examples quoted by Lorimer have the wh-phrase directly preceding the verb, regardless of whether it is clause-initial (18a) or clause-medial (18b). There is one example where the wh-phrase is not adjacent to the verb (18c).

18	a.	u:ŋ-ε 2sg-GEN	gu.-i:k 2sg-name	besΛn what?	b-ila? be-3sg.INAN	(op.cit:164)
						'What is your name?'
	b.	menΛn who?	b-ai be-3sg.MASC	K ₁ sər? K ₁ sər		(op.cit:162)
						'Who is K ₁ sər?'

- c. **ბესანე** ho:i su:-ča ba-iy-λm? (op.cit:164)
 how? vegetables bring-PRF be-1sg-PST
 ‘How was I to bring vegetables?’

Clearly, there is a strong tendency for the wh-phrase to directly precede the verb, even if this is not obligatory as in Georgian or Basque.

4.1.6 Summary of Burushaski features

To summarize, Burushaski shares on the surface all of the features which Basque and Georgian are known to share - ergativity (of some form), polypersonalism, SOV word order, paradigm trimming, preverbal wh-phrases (preferred, but not obligatory). A deeper analysis reveals that while the tense/aspect split is similar to that of Georgian, the case pattern is more akin to that of Basque, and that polypersonalism involves a mixture of verbal and nominal categories (as opposed to the uniform verbal system in Basque and Georgian).

4.2 Chukchi

4.2.1 Basic typological facts

Chukchi is a Paleosiberian language spoken in the extreme east of Siberia. The data quoted here is based primarily on Skorik & Spencer’s online *Chukchi grammar*¹¹. Chukchi has relatively free word order, with SOV as the least marked option (this tallies well with other word order facts, such as postpositions, REL-N order, etc).

Chukchi has a fully ergative case system for all tenses. It possesses a maximum of 10 cases (although some -in particular ERG- are subject to syncretism). Fortunately for our present purpose, ABS is clearly distinguished for all types of NP’s, and ERG syncretizes with LOC and INSTR, never with the marker of possession analogous to GEN.

In contrast to Basque and Georgian, Chukchi does not have an active configuration: there are no intransitive verbs which behave syntactically like transitive verbs. On the contrary, Chukchi has an antipassive construction which derives grammatical intransitives from transitives (19).

19. gəm t-ine-tejkə-rkən orw-etə (op cit:section 5:3)
 1sg.ABS 1sg.-ANTIPAS-make-PRES sledge-ALL
 ‘I am making a sledge.’

In such cases, the object is either omitted or realized in one of the oblique cases allative, instrumental or locative. It is never realized in ABS (in contrast to Basque or Burushaski). The distribution of ABS is restricted to once per clause, making it a structural equivalent of NOM

¹¹ http://privatewww.essex.ac.uk/~spena/Chukchee/CHUKCHEE_HOMEPAGE.html

rather than ACC. In this respect, Chukchi patterns with Georgian rather than Basque (but cf also section 4.2.4 below).

4.2.2 Agreement in Chukchi

4.2.2.1 The structure

Chukchi displays a rather complex kind of polypersonal agreement¹², characterized by a high degree of interaction between person, number, tense, aspect and mood. Thus, it is virtually impossible to make any simple generalizations about the positioning of morphemes (regardless of whether we choose case or grammatical function as the relevant criteria). A general outline follows below:

The structure of the Chukchi verb can be subdivided into the following slots:

A	B	C	D	E	F	G
	<i>ra</i>				<i>ər̥kə</i>	
	FUT		√		PROG	

Slot D is that occupied by the verb root, and slots B and F can be occupied by the future morpheme *-ra-* and the progressive morpheme *-ər̥kə-* respectively. The slots which are of relevance to verb agreement are thus slots A, C, E and G. As can be seen, there are two slots which are relatively peripheral (A and G), and two which are relatively central (C and E).

Chukchi has 10 finite tense/aspect/mood categories, namely aorist (simple past tense), present progressive, future, future progressive, imperative perfective, imperative imperfective, conditional perfective, conditional imperfective, stative present and resultative perfect. The imperfective and progressive forms are characterized by the presence of *-ər̥kə-* in slot F, and the perfective forms by its absence. The categories can be grouped as follows:

<i>-ər̥kə-</i> in slot F		
AOR	PRES	
FUT	FUT PROG	<i>-ra-</i> in slot B
IMPER PF	IMPER IPF	special morphology ¹³
COND PF	COND IPF	<i>-ʔ-</i> infix in slot B ¹⁴
STAT		<i>n-</i> prefix (replaces slots A&B)
RES		<i>ge-</i> prefix (replaces slots A&B)

¹² Chukchi verbs agree with the subject and with the direct object, but not with the indirect object (as opposed to the situation in Basque, Georgian and possibly Burushaski).

¹³ In particular, the agreement affixes in slot A are different from those in other categories.

¹⁴ Since FUT and COND exclude one another, it is impossible to say whether *-ʔ-* is located in slot B or replaces slot B, or is infixes immediately before or after slot B. For ease of exposition, let us assume that it is located in slot B.

For ease of exposition, only one imperfective form is discussed, namely the present progressive, leading to a total of 7 categories to be covered. These will not be discussed in detail, rather, we shall look at the behaviour of the various agreement affixes in the different categories.

As we have already mentioned, there is no general one-to-one correspondence between an argument and a corresponding affix position in Chukchi. Thus, the positioning of morphemes does not clearly reflect either grammatical function, person, number or even the case of the argument. The situation is further complicated by the fact that there is no obvious one-to-one relation between the form of affixes and their meaning either: thus, while *tə-* always corresponds to a 1sg subject and *mət-* always corresponds to a 1pl subject, *-tək* can refer to 2pl which is *either* an object or a subject (in the latter case only if the object is not 3rd person). Note that this holds *within* a given tense/aspect/mood paradigm:

20	a.	ine-lʔu-tək 1sO-see-2pS 'You (pl) saw me.'	b.	ne-lʔu-tək 3S-see-2pO 'S/he/they saw you (pl).'
----	----	--	----	---

The affix *-tək* at least conveys the meaning 2pl -with the grammatical function being dependent on other factors. But the affix *-ine-* in slot C (listed here as representing a 1sg object) can, in the stative present, have the following references:

- 1sSUBJ
- 1pOBJ
- 2sSUBJ (if OBJ = 1s, 3s or 3p)
- 2pSUBJ (if OBJ= 1s, 3s or 3p)

The only common denominator seems to be that at least one of the arguments is non-3rd-person. For the suffix *-gʔen* (slot G), on the other hand, the constellation is even less transparent: it has the following references:

AOR, FUT: 3sOBJ (if SUBJ = 1s, 2s, 1p, 3p)

COND: 1sSUBJ (if OBJ = 3s)
2sSUBJ (if OBJ = 1s, 3s, 1p)
3sSUBJ (if OBJ = 1s, 1p, 3p)

IMPER: 1sSUBJ (if OBJ = 3s)
2sSUBJ (if OBJ = 1p, 3p)
3sSUBJ (if OBJ = 1s)

At the same time, since it is present in 4 different T/A/M categories and not omnipresent in any, it is clearly *not* a T/A/M affix. Thus, any attempt at generalization must admit a large number of exceptions. Nevertheless, the following represents an attempt at some kind of generalization:

- 1) 1st person subjects are usually prefixed (slot A)
- 2) 1sg objects are usually prefixed (slot C)
- 3) 1pl objects are usually suffixed (slots E and G)
- 4) 2nd person subjects and objects are usually suffixed (slot G)
- 5) 3rd person ERG are usually prefixed (slot A)
- 6) 3rd person ABS are usually suffixed (slot G)

In many cases, a given argument is represented by \emptyset -morphology, its reference being deduced from the appearance of the rest of the verb paradigm. From the above, a couple of other generalizations crystallize:

a) the intermediate slots (C and E) are primarily used for object agreement, whereas more peripheral slots (A and G) can cross-reference either subjects or objects. This tallies well with the fact that slots C and E are not used with intransitive verbs;

b) The positioning of agreement affixes is clearly split along the following lines:

- 1) 1st person: affix position dependent on grammatical function and number;
- 2) 2nd person: affix position dependent on person
- 3) 3rd person: position dependent on Case (prefixes for ERG, suffixes for ABS)

In practice, this results in an accusative agreement alignment for 1st person, a neutral alignment for 2nd person, and an ergative alignment for 3rd person¹⁵ (in contrast with a purely ergative alignment for Basque and a largely accusative alignment for Georgian). In this respect, Chukchi displays features found in both Georgian and Basque (and is, with respect to the alignment of agreement, thus typologically more similar to both Georgian and Basque than these two languages are to one another).

Some affixes only occur in certain T/A/M categories, and when they do, some have an ergative alignment, some an accusative alignment, and some a split alignment, also involving a type of absolutive displacement (if ERG is 3rd person, ABS is realized in the position corresponding to ERG), which is almost the mirror image of Basque ergative displacement (if ABS is 3rd person, ERG is realized in the position corresponding to ABS). While it is outside

¹⁵ The direction of the split tallies well with the Nominal Hierarchy discussed by Dixon (1994:85). 1st and 2nd person are more likely to have an accusative alignment than is 3rd person.

the scope of this report to describe the various conflicting alignments of affixes in detail, it should be noted that Chukchi has a clearly mixed agreement system, and appears to combine several phenomena hitherto covered in both Basque and Georgian.

4.2.2.2 *The affixes*

At this point it would be fruitful to examine more closely the identity of some of the agreement morphemes involved. While most of the affixes are relatively opaque, four are absolutely identical to the corresponding pronoun:

<i>-gə̃m</i>	‘1sg’	<i>-muri</i>	‘1pl’
<i>-gət</i>	‘2sg’	<i>-turi</i>	‘2pl’

It is interesting to note that the distribution of these affixes is not uniform, despite, as we shall see, other clear similarities. Thus, *-muri* and *-turi* occur only in STAT and RES, where they have two different types of split-ergative alignment combined with an absolutive displacement:

	INTR	TRANS	
		STAT	RES
<i>-muri:</i>	1pSUBJ	1pSUBJ 1pOBJ (if SUBJ = 3)	1pOBJ (if SUBJ = 3)
<i>-turi:</i>	2pSUBJ	2pSUBJ 2pOBJ (if SUBJ = 3)	2p OBJ

The 1sg suffix *-gə̃m* occurs in these contexts, but also serves as a marker for 1sOBJ (if SUBJ = 3p) in all other tenses as well. The 2sg suffix *-gət* has the widest distribution, since, in addition to the aforementioned contexts, it also occurs as a generalized marker of 2sgOBJ for all tenses.

The existence of opaque and transparent affixes, the inconsistencies in their behaviour, and the general lack of uniformity in the conjugation system suggests that the various affixes have been added to the paradigm in different stages, in turn suggesting the possible development or modification of a polypersonal agreement system.

4.2.3 *Agreement and Tense*

Let us now look at the relation between Agreement and Tense. Following the Mirror Principle (Baker 1988), it seems clear that slots C and E are structurally below T° and Asp° respectively, while the other slots, being more peripheral, are arguably higher in the structure. Since slots C and E are primarily used for object agreement, it might be argued that Chukchi has one category of object agreement below T° and Asp°, and one category of subject agreement above these levels. However, since the functions of neither of the two levels is free from exceptions, and

since even the exact function of various of the affixes is not clear, this suggestion is hard to motivate on other than purely theory-internal grounds. We might possibly argue for more levels of agreement, one for each slot. Again, it is not clear that such a suggestion can be empirically motivated. In fact, even such an uncomplicated position as slot A (housing the subject morphemes *tə-* ‘1sgSUBJ’ and *mət-* ‘1pSUBJ’, which display an accusative alignment) is also the location of the 3rd person morpheme *ne* ‘3ERG’ which has an ergative alignment. Thus, we cannot define the alignment of each level at present, let alone the categories for which each level is responsible.

What we have seen so far suggests that Chukchi has various levels dealing with agreement, of shifting alignments, and that some persons and/or grammatical functions syncretize in various parts of the paradigm. The result is a split-ergative system of polypersonal agreement, realized by prefixes, suffixes and infixes.

4.2.4 Agreement and Case

We have already mentioned that Chukchi antipassives result in the Agent, rather than the Patient, being realized in ABS, with the Patient instead either being omitted or realized in one of a small group of oblique cases such as allative. This implies that Chukchi clauses only admit one instance of ABS being realized in a clause (analogous to the situation in Georgian, in contrast to the situation in Basque and Burushaski). Nevertheless, the structural cases in Chukchi, ERG and ABS, are not dependent on agreement for their assignation. While Chukchi does not admit of ABS-ABS constructions as in Basque or Burushaski, non-finite verbs, if heading a separate clause, do assign case to their arguments (both ERG and ABS, depending on transitivity).

21. *muri* *am-peljaw-a* *nenenet-e,* *ne-mŋol-more*
 1pl.ABS AM-leave-TA¹⁶ child-ERG 3-pine-1pl
 ‘Because the children leave us, we pine for them.’ (op.cit. section 7.2.4)

In Georgian, we recall, the assignation of ERG and ABS is the privilege of finite clauses (with agreement), whereas in Basque ABS can be assigned to any object, even if the subject is also assigned ABS (in an *ari*-construction). Hence Chukchi represents an intermediate type between Basque and Georgian.

4.2.5 Wh-question formation

Chukchi does not appear to have any marked preference for an immediately preverbal *wh*-position. While the *wh*-word is preverbal in all relevant examples in the data, it is often separated from the verb by adverbials (cf 22).

¹⁶ The meaning of the AM-TA circumfix in Chukchi is given in Skorik & Spencer as ‘just’ or ‘only’ - this is not particularly clear with the present example.

22. gət-əm e-**mik**-ke igər r-iwini-gʔe
 2s-EMPH NEG-**who?**-NEG today FUT-hunt-1/2FUT
 ‘Who will you not go hunting with today?’ (op.cit section 8.1.1)

The issue is further complicated somewhat by two other facts: Chukchi allows the double incorporation of modifiers into nouns, and of the resulting incorporation into the verb (23a) and Chukchi has a special wh-verb *-req-* ‘to do what?’ (23b).

- 23 a. tə-**tʔar**-qora-kənʔor-rkən
 1sg-**how.many?**-reindeer-lasso-PROG
 ‘How many reindeer am I lassoing?’ (op. cit section 5.4.6.1)
- b. **req**-ərkən-əm igirqej gə-nin ekəkʔ
 do.what?-PROG-EMPH right.now 2sg-POS son.ABS
 ‘What is your son doing right now?’ (op. cit. section 4.2.1.1)

In cases like these, the adjacency of the wh-word with the verbal head is either trivial (due to incorporation) or not an issue at all (since it *is* the verbal head).

Given the fact that Chukchi word order is relatively free, it is difficult to state any generalizations about the unmarked or preferred wh-position. At any rate, it is clear that Chukchi does not have any restriction of wh-positioning comparable to that in Basque or Georgian.

4.2.6 Chukchi summary

From a cursory view of Chukchi clause structure, we have seen that Chukchi displays characteristics intermediate between Georgian and Basque (including the interaction of case and agreement and the alignment of agreement morphology). In other cases, such as the ergative (as opposed to active) alignment and the word order in wh-questions, it represents a different type of structure from that shared by Basque and Georgian. Since Chukchi verbs do not agree with the indirect object, the question of paradigm trimming is not an issue here.

4.3 Inuit

4.3.1 Basic typological facts

We now turn to the final ergative language in our sample, the Eskimo-Aleut language Inuit. Data on Inuit has been taken from two sources, Bok-Bennema 1991 and Manning 1996. Inuit is spoken in an Arctic zone from Alaska to Greenland, and the data quoted is primarily from the highly standardized variant West Greenlandic, which is the official language of Greenland.

Inuit has relatively free word order, but SOV is unmarked, and like Chukchi, it possesses both antipassive and incorporation constructions (24 a,b respectively, cf 24 c). Inuit has no ergativity splits of any kind (as opposed to the situation in Georgian and Burushaski).

- 24 a. ujarqa-mik tigu-si-voq (Manning 1996: 93)
 stone-MOD take-ANTIP-IND.3sg
 ‘He took a stone.’
- b. Suulut ataatsi-mik ammassat-tor-poq (Manning 1996: 118)
 Suulut.ABS one-MOD sardine-eat-IND.3sg
 ‘Suulut ate one sardine.’
- c. ujarak tigu-vaa (Manning 1996: 93)
 stone.ABS take-IND-3sg.3sg
 ‘He took the stone.’

In practice, this implies that ABS is only assigned once per clause in Inuit, parallel to the situation in Georgian, but in contrast to Basque. At the same time, Inuit differs from both Basque and Georgian in not having an active alignment, but rather a true ergative alignment. Thus, the assignation of ERG in Inuit is dependent on the actual transitivity of the clause, and agentive intransitive verbs in Inuit take ABS subjects (25).

25. Oli sinip-poq. (Manning 1996: 3)
 Oli.ABS sleep.IND.INTR.3sg
 ‘Oli sleeps.’

In fact, a transitive verb¹⁷ may be used intransitively as long as the object is realized in modalis case rather than ABS, the agent then being realized in ABS instead (26).

26. Jaani tuktu-mik taku-vuq. (Manning 1996: 15)
 Jaani.ABS caribou-MOD see-IND.3sg
 ‘Jaani sees a caribou.’

Inuit has 8 cases, the most relevant of which are (for our present purposes) ABS, ERG and modalis (MOD). The case termed here ERG is, in fact, identical to the form used to mark possession - we could therefore postulate a 9-case system with total syncretism between ERG and GEN. Henceforth, for reasons of clarity, the term ERG will be used in verbal contexts, and the term GEN in nominal contexts, under the understanding that we are in fact dealing with the same case.

¹⁷ This holds for a large subset of verbs termed *ambiguous verbs* by Bok-Bennena (op cit. 46). These include many (but not all) transitive verbs and unergative verbs.

4.3.2 Inuit agreement

Inuit has polypersonal agreement, with the verb agreeing with both the subject and the object, but not with the indirect object. The agreement pattern is relatively straightforward, with discrete affixes for each category, although the facts are clouded somewhat by the complex system of morphophonemic changes which affect the shape of affixes in various phonological contexts. For this reason, not all affixes are transparent on the surface.

All the relevant verb morphology in Inuit (mood¹⁸ and agreement) is suffixal - for this reason, the Mirror Principle (Baker 1988) is applicable for the derivation of the hierarchical structure of the clause. Assuming the Mirror Principle, agreement is the highest category in the Inuit clause, followed by mood and presumably tense/aspect. The relevant slots are illustrated below.

√	TENSE / ASPECT	MOOD	AGR1	AGR2
---	-------------------	------	------	------

Given that Inuit has two agreement slots, it is relevant to discuss the positioning of the affixes and their identity. There is a general consensus that the most transparent combinations synchronically place ERG in AGR1 and ABS in AGR2, although Manning (op cit: 100ff) argues that the order is in fact mixed, both synchronically and diachronically, with ABS-ERG ordering occurring (in some moods) if ABS is 3rd person - incidentally a type of ergative displacement similar to that of Basque, although less clearly visible.

In general, however, the ERG-ABS tendency holds true, and this in turn suggests that object agreement is the higher of the two agreement categories in Inuit (an interesting contrast to the Minimalist analysis of Inuit clause structure presented in Bobaljik 1993 - which is concerned with the syntax to the exclusion of agreement morphology).

The identity of the affixes is not immediately obvious, in part since the morphophonemic changes involved greatly alter the the shape of the individual morphemes. Furthermore, one of the agreement morphemes (-*si*- '2pl') syncretizes for both subject and object agreement. However, the general pattern of agreement seems to be ergative: thus, the same affix is used to cross-reference the subject of an intransitive verb and the object of a transitive verb. Example (27) illustrates this relatively clearly.

¹⁸ Tense/aspect distinctions are optionally expressed by suffixes preceding the mood/agreement ending, and are not discussed by Bok-Bennema as part of the inflection as such, presumably because they do not interact with the other categories, and because the category they represent often corresponds to adverbials of various kinds. For our purposes here, this does not affect the discussion, since they, like mood, are less peripheral in the inflected verb form than is agreement.

- 27 a. taku-vas-si-nga (Bok-Bennema 1991: 42)
 see-IND-2pl-1sg
 ‘You (pl) saw me.’
- b. ateqar-pu-nga Kaali (Bok-Bennema 1991: 218)
 named-IND.INTR-1sg Kaali
 ‘I am called Kaali.’

In many other cases, the morphophonemic changes effectively camouflage this relation between the transitive and intransitive agreement patterns.

Since both agreement categories are suffixal, it is impossible to identify which slot the intransitive agreement affix occupies - however, since the form of the affix in itself suggests an ergative agreement alignment, we have no reason to doubt that the positioning should also point in the same direction. We can thus safely claim that Inuit, with respect to agreement alignment, behaves like Basque rather than like Georgian or Burushaski (even to the extent of displaying a type of ergative displacement).

4.3.3 Inuit finiteness

Inuit ERG is not only the case used for Agents - it is also the case used for possessors. In fact, the parallelisms between possessive constructions and argument structure go remarkably far. This is illustrated in Bok-Bennema 1991:31 with data from the Labrador variety of the language (28).

- 28 a. kivga-t nunang-at
 servant-PL.GEN land-3pPOS
 ‘the land of the servants’
- b. kivga-t attuar-paat
 servant-PL.ERG read-3sABS.3pERG
 ‘The servants read it.’

Not only is the case of the possessor identical with ERG, the marking of the possessed noun is identical to the corresponding agent agreement on a verb. In fact, the parallelism perhaps stretches even further: in appositions, a coreferent pronoun is realized in a form identical to the agreement suffix cross-referencing ABS (cf 29c).

- 29 a. taku-vas-si-gut (Bok-Bennema 1991: 42)
 see-IND-2pl-1plABS
 ‘You saw us.’

- b. niri-var-put (Bok-Bennema 1991: 42)
eat-IND.3sgABS-1pl.ERG
'We ate it.'
- c. Amerikamiu-u-suq-gut (Bok-Bennema 1991: 85)
American-be-ones-1plABS
'We who are Americans / We Americans'

Given the striking parallelism between ERG and GEN, it is difficult to relate the assignation of ERG to verb agreement. However, we can relate it to agreement as a more general phenomenon, if we accept the possessor suffixes in constructions like (28a) as agreement. Whether or not this can be conveniently equated with case assignation in a functional projection of the clause is, however, another matter.

In control constructions, the object of the embedded clause can be realized in ABS. This might be taken to imply that ABS is assigned by the verb, analogously to the situation in Basque. However, the verb form in the embedded clause, termed *infinitive* in Manning (1996) and *gerundive* in Bok-Bennema (1991), is not, in fact, non-finite, as the term might suggest. Rather, it displays ABS agreement (30).

30. Miiqqat [_ Juuna ikiu-ssa-llu-gu niriursui-pp-u-t.
children.ABS (ERG) Juuna.ABS help-FUT-INF-3sg promise-IND-INTR-3pl
'The children promised to help Juuna.' (Manning 1996:124)

For this reason, it seems clear that both ERG and ABS are assigned by a morphological category of agreement, even if the exact mechanisms are not entirely clear.

4.3.4 Inuit wh-structure

Inuit has no fixed wh-position, wh-words are located in situ (as far as can be determined in a language with free word order). In this respect, Inuit possibly patterns with Chukchi, and at any rate differently from the situation evident in Georgian and Basque.

4.3.5 Inuit summary

We have noted that the ergative language Inuit has polypersonal agreement aligned on an ergative basis (like Basque, but unlike Georgian). The agreement system involves ergative displacement with 3rd person patients (in some moods), like Basque. ABS and ERG are assigned by agreement (like Georgian, but unlike Basque). Inuit has antipassives and incorporation, and does not have an active alignment, nor any ergative splits. The agreement system shows striking parallels with the possessor system, to the extent of suggesting a nominal

origin (if not nature) of Inuit clause structure (cf Manning 1996:20 for a review of the relevant literature).

4.4 *The parallel development of polypersonalism*

We have seen that Basque, Georgian, Burushaski, Chukchi and Inuit all display polypersonalism. However, there are striking differences in the behaviour of the polypersonal systems. The variation in the synchronic patterning of the agreement systems can be characterized as in the following table:

	Basque	Georgian	Burushaski	Chukchi	Inuit
Alignment	ERG	(ACC) ¹⁹	BOTH	MIXED	ERG
Erg. displacement	+	-	-	(ABS displ)	+
Ind. obj	+	+	(+)	-	-
Paradigm trimming	+	+	(+)	(-)	(-)
Possessive syncretism	-	-	ABS	-	ERG
Prefixes	ABS	1/2	ABS	1 / 3.ERG	-
Suffixes	ERG	3	SUBJ	2	ALL
No. of slots ²⁰	3	2	2	4	2
Suggested structure ²¹	T/AGR	AGR/T	AGR/T	AGR/T/AGR	AGR/T

We see a remarkable heterogeneity in the build-up of polypersonal agreement - the more so remarkable since all the languages share the two features of ergativity and polypersonalism. Furthermore, even within a given language, there is a striking variation in the origin and behaviour of the agreement affixes. The only feature which connects these 5 languages seems to be the link between ergativity and polypersonalism.

Put somewhat simplistically, the fact that these 5 languages all display ergativity and polypersonalism could hypothetically be explained by some kind of (presumably pre-Altaic or pre-Indo-European²²) historical contact. But if this were the case, we would either expect, on the one hand, far greater similarities in the build-up of the respective agreement systems, or (given a sufficient time-depth to account for the present diversity) the loss of polypersonalism in one or some of the languages involved.

¹⁹ Recall, however, that inversion constructions display an ergative alignment.

²⁰ Dative agreement is discounted here for Basque and Georgian for ease of comparison with the other three languages which lack indirect object agreement (the object being to compare the number of alternative slots for a given number of arguments, i.e. indirectly the complexity of the agreement system irrespective of the number of arguments cross-referenced).

²¹ Based on the relative positions of Tense and Agreement, and assuming Baker's Mirror Principle.

²² In the sense of predating the *spread* of Altaic and Indo-European languages across the relevant areas (and not necessarily predating the origin of the language families as such).

What the diversity (including the probable diachronic diversity within a single language, cf Georgian and particularly Chukchi) instead shows is that polypersonalism itself is quite stable, whereas its realization in a given case is not. The natural conclusion would therefore seem to be that ergativity puts pressure on a grammatical system to develop and maintain polypersonalism²³. In section 3 we have already speculated on functional and structural reasons for why this might be the case. Whether or not this can be statistically confirmed for a larger sample of languages is the topic of future research.

If this correlation can be shown to valid, then the great diversity between the languages in the sample would be expected - each language, with its own structural idiosyncrasies, would then be under pressure to develop a system of polypersonal agreement, making use of whatever mechanisms might be available, such as possessor marking (in Inuit and Burushaski), pronoun cliticization (to a certain extent in Chukchi) or incorporation of pronominal features into the verb (as possibly in Georgian, cf also Nash-Haran 1992).

While the above scenario would account for the known facts as an alternative to genetic relationship, it depends crucially on the assumption that ergativity in itself is a driving force in the development of polypersonalism, and this assumption, while plausible in the light of the sample languages reviewed here, must be checked against a much wider sample of ergative languages, with a control sample of accusative languages. I defer this issue to future research.

4.5 Summary

In the preceding sections, we have looked at various phenomena in Basque and Georgian, and compared these with three other ergative languages for which no relationship has been proposed. The results are briefly summarized in the following table:

	Basque	Georgian	Burushaski	Chukchi	Inuit
ACTIVE	yes	yes	no	no	no
SPLIT ERGATIVE	no	yes	yes	no	no
ABS-ABS	yes	no	yes	no	no
POLYPERSONAL AGR	yes	yes	yes	yes	yes
CASE DEP. ON AGR.	no	yes	yes	yes (?)	yes
AGR. ALIGNMENT	ERG	(ACC)	BOTH	MIXED	ERG
AGR/T	no	yes	yes	Agr/T/Agr	yes
PREFIXES	ABS	1/2	ABS	1 / 3.ERG	-
SUFFIXES	ERG	3	SUBJ	2	ALL
PREVERBAL WH	strict	strict	optional	free	in situ
WORD ORDER	SOV	SOV	SOV	SOV	SOV

²³ Polypersonalism seems to be in the process of developing in the ergative Iranian language Talysh (cf Comrie 1981: 175).

The table of features reveals some interesting facts: firstly, it seems clear that Basque and Georgian do not form any clear group to the exclusion of the other languages. Based on a simple count of the shared features, the languages typologically closest to Basque are Georgian, Burushaski and Inuit (all with 3 shared features) and the languages closest to Georgian are Burushaski and Inuit (both with 4 shared features), while Basque only shares 3 features with Georgian, and Chukchi shares between 2 and 5 features with Georgian, depending on how liberally we interpret the mixed alignment of Chukchi. The above issues have been discussed in detail in this paper.

The first two sections of the paper showed that many of the similarities which we find between Basque and Georgian do not carry over to a deeper level: while both languages have an ergative pattern, Basque ABS is structurally equivalent to ACC, whereas Georgian ABS is structurally equivalent to NOM. Both languages have polypersonal agreement but the alignment and build-up of the agreement system is entirely different. Thus, while the surface behaviour is similar, virtually nothing else is. This would seem to make it unlikely that the similarities are inherited from a common ancestor, as would be the case in a scenario of genetic relationship. A possible alternative scenario is outlined in section 3, and again in section 4.4.

More importantly, in my opinion, the empirical survey in section 4 shows that the features shared by Basque and Georgian are not particular to these languages, but rather are typical of ergative languages (or a subset of these) as a group. It follows that any attempt to establish a relationship between Basque and Georgian on typological evidence should also make reference to the patterning of the other three languages in the sample. This could lead to two possible conclusions: either the proposed language family or phylum should be extended to include the other three languages, or the typological features should be considered as possibly illustrating an implicational universal tendency (pending further investigation of a more representative sample of ergative languages). For geographical and historical reasons, the latter avenue seems the more realistic.

With the risk of getting involved in the issue concerning the relationship between the three languages families of the Caucasus (Kartvelian, Northeast Caucasian and Northwest Caucasian) a possible alternative to strengthen the genetic hypothesis would be to widen the investigation to cover languages from the other two Caucasian families. It might be the case that Georgian is not the most suitable (although the best known and the most publicized) candidate to represent the Caucasian languages with respect to possible contacts with Basque.

5. Conclusion

The results of the above investigation show that there is no clear typological evidence in favour of genetic or historical connections between Basque and Georgian, for the following reasons:

- a) certain similarities are simply surface phenomena;
- b) certain similarities are structurally interdependent;

c) certain similarities can be empirically shown to be interdependent;

It should be noted that this does not necessarily mean that Basque and Georgian are not genetically or historically related - it simply shows that the most salient typological facts they share do not serve as valid evidence (or at least evidence of a relation between Basque and Georgian to the exclusion of the other languages discussed), and any evidence for such a relationship must either be based on the possibility of other shared features not covered here, or on lexical correspondences. For establishing genetic relationships, lexical correspondences are greatly to be preferred.

References

- Addis, Kristin. 1993. 'Paradigm trimming in Basque.'. *Linguistics* 31:431-474
- Baker, Mark. 1988. *Incorporation - a theory of grammatical function changing*. Univ. of Chicago Press
- Bobaljik, Jonathan. 1993. 'On ergativity and ergative unergatives'. *MIT Working Papers in Linguistics* 19: 45-88
- Bok-Bennema, Reineke. 1991. *Case and agreement in Inuit*. (Studies in Generative Grammar). Berlin: Foris
- Comrie, Bernard. 1981. *The languages of the Soviet Union*. Cambridge: Cambridge University Press
- Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press
- Guilfoyle, Eithne, Henrietta Hung & Lisa Travis. 1992. 'SPEC of IP and SPEC of VP: Two subjects in Austronesian languages'. *Natural Language and Linguistic Theory* 10.3: 375-414
- Harris, Alice. 1981. *Georgian syntax. A study in relational grammar*. Cambridge: Cambridge University Press
- Holmer, Arthur. 1999. 'An active analysis of Basque ergativity'. *Fontes Linguae Vasconum* 81: 189.225. Pamplona: Gobierno de Navarra
- Holmer, Arthur & Sheila Dooley-Collberg. in preparation. 'Voice and argument structure.' To appear in the report volume of the research project *Parametric Typology*
- Holmer, Arthur & Karina Vamling. forthcoming. 'The parameters of Kartvelian ergativity'. To appear in the report volume of the research project *Parametric Typology*
- Kennelly, Sarah. 1999. 'The syntax of the P-focus position in Turkish'. In: Rebuschi, Georges & Laurice Tuller (eds.). *The grammar of focus*. Linguistik Aktuell (24). Amsterdam: Benjamins
- Laka, Itziar. 1993. 'The structure of inflection: a case study in X⁰ syntax'. In Hualde, José & Jon Ortiz de Urbina (eds). *Generative studies in Basque linguistics*. Amsterdam: Benjamins (pp. 21 - 70)

- Laka, Itziar. 1994. *On the syntax of negation*. (Outstanding dissertations in linguistics). New York & London: Garland Publishing
- Lorimer, D.L.R. 1935a. *The Burushaski language. Vol I. Introduction and grammar*. Oslo: Instituttet for sammenlignende kulturforskning
- Lorimer, D.L.R. 1935b. *The Burushaski language. Vol II. Texts and translations*. Oslo: Instituttet for sammenlignende kulturforskning
- Manning, Christopher. 1996. *Ergativity. Argument structure and grammatical relations*. Stanford: CSLI Publications
- Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA.: MIT Press
- Nash-Haran, Léa. 1992. 'La catégorie AGR et l'accord en géorgien'. *Recherches Linguistiques* 21:65-79
- Ortiz de Urbina, Jon. 1989. *Parameters in the grammar of Basque - a GB approach to Basque syntax*. Dordrecht: Foris
- Skorik P. & Andrew Spencer. *Chukchi grammar*. Available online at:
http://privatewww.essex.ac.uk/~spena/Chukchee/CHUKCHEE_HOMEPAGE.html
- Vamling, Karina. 1988. 'A feature-based analysis of Georgian morphology. Case study: object agreement'. In Thordarson, Fridrik (ed). *Studia Caucasologica I. Proceedings of the Third Caucasian Colloquium, Oslo, July 1986*. Oslo: Institute for Comparative Research in Human Culture. (pp. 309-319)

This report is submitted to *Birgit och Gad Rausings Stiftelse för Humanistisk Forskning* in grateful acknowledgement of the financial assistance which made the above research possible. Further spinoff research based on the above results may be published in future, in which case due acknowledgement will be made.

January, 2002

Arthur Holmer