Archis: the challenge of an extreme agreement system

Marina Chumakina and Greville G. Corbett
Surrey Morphology Group

Sandro Vasil’evič is widely known for his work in phonetics and phonology, particularly with regard to the languages of Daghestan, where the difficulties of the sound systems tower over the linguist like the mountains where the languages are spoken. His fundamental work has had a major role in making it possible to study other linguistic topics in Daghestanian languages, so it is fully appropriate that we should dedicate an article on agreement to Sandro Vasil’evič. The interest of the agreement system of Archi is apparent from reading Kibrik, Kodzasov, Olovjannikova & Samedov (1977) and Kibrik (1977a, b). That grammar is packed with interesting phenomena, sometimes dealt with quite briefly and laconically. Our purpose is to show just how challenging the agreement system of Archi is. We do so within a recent typological framework, partly highlighting material from the 1977 grammar, and partly presenting material from recent fieldwork and new analyses.

It will be helpful to settle on terminology, to give a framework within which we can discuss the Archi data (Corbett 2006: 4-5). We shall call the element which determines the agreement (typically the absolutive argument) the **controller**. The element whose form is determined by agreement (say the verb) is the **target**. The syntactic environment in which agreement occurs (the clause for instance) is the **domain** of agreement. And when we indicate in what respect there is agreement, we are referring to agreement **features**. Thus number is an agreement feature, it has the values: singular and plural (with further values like dual and paucal in some languages). This is diagrammed with a simple English example in Figure 1.

![Figure 1: Framework of terms](image)

We distinguish features from conditions. Features are directly reflected in agreement. There can be other factors (like word order) which can have an effect on agreement but are not directly reflected like features. Such factors are called agreement **conditions**. Thus within a particular domain, a target agrees with a controller in respect of its feature.

---

1 The support of ESRC under grant RES-062-23-0696 is gratefully acknowledged.
specification (that is, the features and their values); this agreement may be dependent on some other condition being met.

Within each of the five areas delineated by these terms Archi shows phenomena of interest, as we shall see, taking the areas in turn.

1. Controllers

For describing Archi controllers, we will pay attention first to the word class of the controller, and second to its syntactic position. In Archi, nouns, pronouns and nominalized adjectives (i.e. all elements that can function as the head of an NP) can control agreement. Nouns and pronouns inflect for number and case. There is agreement in number and gender; case does not participate in agreement. As we shall discuss in more detail in §4, Archi has two numbers (singular and plural) and four genders. The first and second genders are for male and female humans respectively, other nouns belong in the third and the fourth genders, with a tendency for words denoting animals to be in the third gender. We indicate these genders with Roman numerals. The gender of a noun is established by agreement.

Archi is a thoroughly ergative language, so agreement is always controlled by the absolutive argument:

(1) /uni0295umar-dada usdi-mij-wu w-immi < w > χu-li
   PN-uncle(I)[SG.ABS]  stand.ISG.PFV-LOC.CVB-and  LSG-remain<LSG,>PFV-EVID
Uncle Umar was standing right where he was.

In (1) the agreement of the verb wimmiwχuli ‘remained’ is controlled by the absolutive NP /uni0295umar-dada ‘uncle Umar’, which belongs to gender I.

(2) zari han uw-li was
   1SG.ERG  what(IV)[SG.ABS]  do.IV.SG.PFV-CVB1  2SG.DAT
What have I done to you?

In (2) the agreement of the verb uwli ‘do’ is controlled by the absolutive NP han ‘what’, which belongs to gender IV. The agent zari ‘I’ is in the ergative case, and does not control the agreement.

When the controller is a first or a second person pronoun (which does not vary for grammatical gender) it still controls the agreement, as in (3):

(3) to-w-mi-s un d-aku
   that.one-1.SG-OBL.SG-DAT  2SG.ABS  ILSG-see.PFV
He has seen you (female).

The verb akus ‘see’ takes dative and absolutive arguments. In (3) the absolutive pronoun un ‘you’ refers to a female, and controls the agreement of the verb.

Within the NP, the head noun can control its attribute:
I (female) was dancing at my brother’s wedding.

Here the noun ušmu ‘brother’, is the head of the NP wis ušmin ‘my brother’s’. It is in the genitive, but controls the agreement of the pronoun wis ‘my’. The absolutive pronoun zon ‘I’ controls the agreement of the auxiliary verb erdi ‘be’.

I (female) danced at a good wedding.

In (5) the noun oq ‘wedding’ controls the fourth gender singular agreement of its attributive hibatut ‘good’ in the noun phrase hibatut oqlit ‘at (literally: ‘on’) a good wedding’.

2. Targets

In terms of targets, the Archi agreement system is interesting for several reasons: the word class of the targets, the lexical range covered by agreement, and the problem of multiple exponence.

2.1. Word classes showing agreement

Cross-linguistically, adjectives and verbs are the most familiar word classes that function as agreement targets. In Archi, however, members of almost all word classes have a morphological slot for agreement. There are the following word classes: nouns, adjectives, verbs, adverbs, postpositions and particles. Numerals can be distinguished on semantic grounds, but morphosyntactically some numerals behave like nouns, and some like adjectives.

As we saw above, verbs and adjectives show agreement with the absolutive argument of the clause or with the head of the NP. In the singular, both verbs and adjectives distinguish four genders. In the plural, verbs make a distinction between humans (comprising first and second genders) and non-humans (comprising third and fourth genders). In this respect Archi verbs demonstrate the behaviour typical of other Dagestanian languages like Bagwalal, Tsakhur and Dargi. Adjectives do not distinguish gender in the plural. Gender-number marking is realized with suffixes, prefixes and infixes.
Table 1. Archi gender-number markers for verbs and adjectives:

<table>
<thead>
<tr>
<th></th>
<th>verb</th>
<th>adjective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>PL</td>
</tr>
<tr>
<td>I</td>
<td>w- / &lt;w&gt;²</td>
<td>-b</td>
</tr>
<tr>
<td>II</td>
<td>d- / &lt;r&gt;</td>
<td>-r</td>
</tr>
<tr>
<td>III</td>
<td>b- / &lt;b&gt;</td>
<td>Ø</td>
</tr>
<tr>
<td>IV</td>
<td>Ø</td>
<td>-t</td>
</tr>
</tbody>
</table>

Some examples:

<table>
<thead>
<tr>
<th></th>
<th>verb cas ‘praise’, perfective</th>
<th>adjective hasdu ‘real, reliable’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SG</td>
<td>PL</td>
</tr>
<tr>
<td>I</td>
<td>cu</td>
<td>boco</td>
</tr>
<tr>
<td>II</td>
<td>doco</td>
<td>haśdur</td>
</tr>
<tr>
<td>III</td>
<td>boco</td>
<td>co</td>
</tr>
<tr>
<td>IV</td>
<td>co</td>
<td>haśdut</td>
</tr>
</tbody>
</table>

This is quite a familiar picture. But in Archi, other word classes show agreement as well. Formally, they all behave like verbs in making four distinctions in the singular and two in the plural, but the realization of agreement is different:

Table 2. Archi gender-number markers for adverbs, particles, pronouns and postpositions

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>w- / &lt;w&gt;²</td>
<td>-b</td>
</tr>
<tr>
<td>II</td>
<td>d- / &lt;r&gt;</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>b- / &lt;b&gt;</td>
<td>Ø / -t’</td>
</tr>
<tr>
<td>IV</td>
<td>Ø / -t’</td>
<td></td>
</tr>
</tbody>
</table>

Even for these less usual targets, the controller of the agreement is the always the absolutive argument of the clause. Here are some examples:

**Adverb:**

(6) o < b > q’ta-tu-b balanced dit: a < b > u b-eryin

Past trouble gets forgotten quickly (Kibrik 1977a: 186).

Here the adverb ditabu ‘quickly, soon’ modifies the verb beryin ‘forget’ but agrees with a noun balah ‘trouble’.

---

² We follow the Leipzig Glossing Rules in notating infixes in angle brackets.
Particle:

(7) arša horo:k ej>b>u iškul da<b>þu
Archि.INESS long.ago very<III.SG> school(II)[SG.ABS] open<III.SG>,PFV
A school was opened in Archи very long time ago (Kibrik 1977a: 326).

Here the particle ejbu ‘very’ modifies the adverb horo:k ‘long ago’ but agrees with a noun iškola ‘school’.

Pronoun:

(8) d-ez un malgan
Il.SG-1SG.DAT 2SG.ABS be.dear
You (female) are dear to me (male).

The dative pronouns dez ‘to me’ agrees in gender with the referent of the absolutive pronoun un ‘you’.

Postposition:

(9) to-w-mi-s sin-t’u ḗ“ak-du-t duχri-q’a-k e<b>q’en
that.1-1SG-OBL.SG-DAT know-NEG near-ATR-IV.SG village(IV).SG.INTER-LAT <III.SG>up.to
b-i-tru-b deq’
III.SG-be.PRS-ATR-III.SG road(III)[SG.ABS]
He does not know the way to the next village (Kibrik & Samedov 1977: 227).

The postposition ebq’en ‘to’ governs the word duχri-q’a-k ‘village’ but agrees with the absolutive argument deq’ ‘road, way’

In contemporary theories agreement is defined in terms of syntactic domains. Archи adverbs, pronouns and particles present a challenge, because they lack a clear syntactic link to the controller. Adverbs and particles have been accounted for as having syntactic and semantic scope over the whole clause. Dative pronouns seem to be more problematic since they are governed by the verb, and their syntactic link to the absolutive argument is not clear. The most difficult case to account for is the postpositions: they clearly form a phrase with the preceding noun, determining its case, but at the same time agree with a different noun, where the syntactic link is clearly absent. This has never been addressed by syntactic theories.

2.2. Lexical range covered by agreement

Although all Archи word classes show agreement, there is not a single word class with a hundred percent of its members doing so.

The largest coverage is by adjectives: almost all of them show agreement. In the Archи dictionary created in the Surrey Morphology Group (Chumakina, Brown, Corbett & Quilliam, 2008) there are only 33 adjectives out of 446 that do not agree. Those adjectives are defined as ‘underived’ in Kibrik (1977a).

As for verbs, a little less than a third of them show agreement: 399 out of 1248 in the dictionary.

Hardly any adverbs show agreement: 28 out of 397, one particle out of four, and one postposition out of 34.
Personal pronouns demonstrate a rather complex picture: first person pronouns show agreement when in the genitive and dative cases, first person plural inclusive also show agreement in the ergative. In table 3 below we show only the forms of genders III and IV.

Table 3. Partial paradigm of the first person pronouns

<table>
<thead>
<tr>
<th>case</th>
<th>gender</th>
<th>SG</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>EXCL</td>
<td>INCL</td>
</tr>
<tr>
<td>ABS</td>
<td>zon</td>
<td>nen</td>
<td>nen-t'-u</td>
</tr>
<tr>
<td>ERG</td>
<td>III</td>
<td>nen-a-b-u</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>nen-t'-u</td>
<td></td>
</tr>
<tr>
<td>GEN</td>
<td>III</td>
<td>b-is</td>
<td>bolo</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>is</td>
<td>olo</td>
</tr>
<tr>
<td>DAT</td>
<td>III</td>
<td>bez</td>
<td>b-ela-b-u</td>
</tr>
<tr>
<td></td>
<td>IV</td>
<td>ez</td>
<td>el</td>
</tr>
</tbody>
</table>

Demonstratives, which are used as third person pronouns, mark the gender of their antecedent, compare:

A. to-r lo d-aq'a
   that.one-II.SG girl(II)[ABS.SG] II.SG-come.PFV
   This girl came.

B. to-r d-aq'a
   that.one-II.SG II.SG-come.PFV
   She came.

This situation presents a challenge for defining word classes, as there is no word class in Archi (except nouns) where all members exhibit the same morphosyntactic behaviour, and in the case of the personal pronoun behaviour is not consistent even within the paradigm of a given word.

2.3. **Multiple exponence.**

Even if targets behave similarly (agreeing or not agreeing), there are still complexities in the syntax-morphology interaction. Thus targets may have more than one agreement slot. Many targets mark agreement in two places, some in three, and a few in four. The agreement can be with the same controller, as in example (10):

(11) mahlo-wu b-imma <b> aq'u
   household(III)[SG.ABS]-and III.SG-leave<III.SG>PFV
   ... and left the household (to someone)

Alternatively, agreement can be with two different controllers, as with participles, see section 3.3. below.
3. Domains

The agreement domain is the syntactic environment in which agreement occurs. Normally agreement is defined within a syntactic constituent: a noun phrase or a clause, for example. In Archi, we can distinguish three types of agreement in terms of domains.

3.1. Agreement within the syntactic constituent

We already saw the agreement within an NP (with adjective or first person genitive pronoun), with clause-level adjuncts (adverbs, dative and ergative personal pronouns), and in a clause with the absolutive argument. There are some constructions where both arguments of a transitive verb take the form of the absolutive (the bi-absolutive construction). These constructions will be discussed in more detail in §5.

3.2. Agreement outside the syntactic constituent

These are the situations when the controller of the agreement is outside the phrase. We saw this in the case of postpositions (see example 9), where postposition agrees with a noun other than the noun it is syntactically linked to, and in the case of particles (see example 7), where the particle modifies the adverb but agrees with the absolutive argument of the clause.

Another instance of agreement outside the syntactic constituent is the agreement outside the clause, the interesting phenomenon known as ‘long distance agreement’, in Russian ‘prozračnoe soglazovanie’. The first reference we have found to it is Kibrik in Bergel´son, Zaliznjak & Kibrik (1982: 49). Here is an example from Archi:

(12) laha-s kľ’an b-eker čamasdak b-ukmu-s
girl.SG.OBL-DAT like III.SG.AUX.IPFV date(III)[SG.ABS] III.SG-eat-INF
(This) girl likes eating dates.

Here the auxiliary verb beker agrees with the absolutive čamasdak ‘date’ of the dependent clause. Sentences like this always have a variant where the matrix verb takes IV gender singular agreement (with the whole clause):

(13) laha-s kľ’an ker čamasdak b-ukmu-s
girl.SG.OBL-DAT like AUX.IPFV[IV.SG] date(III)[SG.ABS] III.SG-eat-INF
(This) girl likes eating dates.

Long distance agreement is also possible in situation when the dependent predicate does not have a morphological slot for agreement:

(14) laha-s me⁷ c’abu-s kľ’an b-eker
girl.SG.OBS-DAT whey(III)[SG.ABS]drink-INF like III.SG-AUX.IPFV
(This) girl likes drinking whey.

The conditions favouring long distance agreement and the matrix verbs showing it vary from language to language. Thus, in Archi the verb bijetas ‘begin’ does not allow it, whereas in another Lezgic language, Tsakhur, the verb ‘begin’ does.

For Archi, we can say that long distance agreement is preferred given a particular word order, namely when the dependent clause is located within the main clause:
Example (15B) shows long distance agreement: the auxiliary *ebdi* agrees with the absolutive of the dependent clause *k"ar* ‘thread’. This dependent clause *k"ar abk\l as* is inserted in the main clause *lahas sinši ebdi* ‘the girl knows how’ (the verb ‘know’ takes dative).

Long distance agreement is common in Daghestanian languages. Archi was used as an illustrative example of long distance agreement in Testelec (2001: 401). Long distance agreement has also been reported in Bagwalal, Godoberi, Tabassaran, Tsakhur, and Tsez (see Corbett 2006: 65-66 for references). The examples given here were elicited during our fieldwork in Archi village in 2007.

3.3. Agreement with more than one controller

As example (12) shows, participles take an intermediate position: they show suffixal agreement with the nouns they modify, and prefixal agreement with their absolutive argument:

(16) [*[s\l hru b-i-tu]-r lo*]

    cunning(III)[SG.ABS]   III.SG-be.PRS-ATR-II.SG child(II)[SG.ABS]

Cunning girl

Literally this phrase means ‘girl who has cunning’. The participle *bitur* ‘having’ agrees with its absolutive *s\l hru* ‘cunning’ by the prefix *b-* (III gender), and with its nominal head *lo* ‘girl’ by the suffix *-r* (II gender).

4. Features

The agreement features of Archi are of varying interest. Number is relatively straightforward, with the values singular and plural. According to the research done to date, the use of these number values seems largely unsurprising. Gender (often called ‘class’) presents greater interest. As we noted earlier, Archi has four major genders. Nouns are assigned to these genders according to a ‘predominantly semantic’ system. Genders I and II are easy to describe: nouns denoting male humans are assigned to gender I, and those denoting female humans to gender II. Genders III and IV are more difficult. There are clear semantic groupings: for example, non-human animates are mainly assigned to gender III. Abstracts are typically found in gender IV. However, nouns denoting inanimate concrete objects are found in both genders III and IV; it is hard to find regularities, semantic or formal, to account for the assignment of many of these nouns (see Kibrik, Kodzasov, Olovjannikova & Samedov 1977: 55-66 for details, and Corbett 1991: 27-29 for discussion of their data). Thus the system is predominantly
semantic, in that the rules are semantic in nature, but they do not cover the entire lexicon in the way the strict semantic assignment system does in a language like Godoberi.

Some phonological regularities in gender assignment were suggested in Kibrik, Kodzasov, Olovjannikova & Samedov (1977: 64-66): nouns with initial $b$-, $m$- and final $-n$, $-u$, tend to be in gender III. We can test these suggestions using counts from the Archi dictionary (Chumakina, Brown, Corbett and Quilliam, 2008). The basic statistics are as follows:

- total number of nouns in the dictionary: 2417
- total number of nouns belonging to gender III: 909 (37.6%)
- total number of nouns belonging to gender IV: 1153 (47.7%)

Genders I and II, not surprisingly, account for a minority of the nouns. They are semantically assigned and we set them aside in order to concentrate on the suggestion of phonological assignment. Of the remaining nouns, those in genders III and IV, we can see that the balance is in favour of gender IV. If we consider just the nouns in these two genders (2062 in all), then 44.1% are in gender III and 55.9% in gender IV. These are the baseline figures to keep in mind when assessing the suggestions for phonological assignments. The following table takes the nouns belonging to each of the four phonological classes suggested by Kibrik, Kodzasov, Olovjannikova & Samedov, and gives the percentage of those in gender III from the nouns in genders III and IV.

**Table 4. Suggested phonological assignments**

<table>
<thead>
<tr>
<th>Nouns in:</th>
<th>total in genders III and IV (i.e. the sum of the less semantic genders)</th>
<th>in gender III</th>
<th>% in gender III against total</th>
<th>% in gender III against third column (sum of genders III and IV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b$-</td>
<td>161</td>
<td>133</td>
<td>85</td>
<td>52.8</td>
</tr>
<tr>
<td>$m$-</td>
<td>170</td>
<td>149</td>
<td>81</td>
<td>47.6</td>
</tr>
<tr>
<td>$-n$</td>
<td>217</td>
<td>173</td>
<td>49</td>
<td>22.6</td>
</tr>
<tr>
<td>$-u$</td>
<td>123</td>
<td>68</td>
<td>53</td>
<td>43</td>
</tr>
</tbody>
</table>

It is the final column which is most enlightening. We established that in the genders that are not assigned purely by semantic criteria, that is III and IV, something over half of the nouns are in gender IV. Against that background, the fact that 63.9% of the nouns in $b$- are in gender III suggest that there is an effect. In some instances the $b$- is a fossilized overt gender marker (Kibrik, Kodzasov, Olovjannikova & Samedov 1977: 65). The figure for final $-u$ is also impressive. The other figures are less convincing.

It has been claimed that verbal agreement in Archi can be adequately described just in terms of gender and number. We suggest that Archi also has a category of person. Our claim is based on the behaviour of first and second person plural pronouns, whose agreement pattern differs from that of third person pronouns. With personal pronouns, we might expect verbs to show human agreement (marker $b-/<b>$) when the referents are
human, and non-human agreement (zero marking) when the referents are non-human (see Table 2). For third person plural pronouns this expectation turns out to be correct:

(17) teb b-as:ar-ši b-i
that.one.ABS.PL HUM.PL-tremble.IPFV-CVB HUM.PL-be
They (humans) are trembling.

(18) teb a:šar-ši i
that.one.ABS.PL [NHUM.PL]tremble.IPFV-CVB HUM.PL-be
They (non-humans) are trembling.

However, with first and second person pronouns, verbs do not show the expected forms:

(19) nen a:šar-ši i
1PL.ABS [1PL]tremble.IPFV-CVB [1PL]be
We are trembling.

(20) ž/en a:šar-ši i
You (plural) are trembling.

We believe that these data argue for postulating the presence of the category of person; this conclusion is strengthened when we look at the data from gender resolution: see Chumakina, Kibort & Corbett (2007) for details.

5. Conditions

Conditions on agreement can apply when there are several agreement options; they are factors determining the choice. In §3.2 we noted ordering conditions that favour the usage of long distance agreement. Another agreement option in Archi concerns the agreement in the bi-absolutive construction. There are several conditions on this construction: the predicate must be analytical, and imperfective. The agreement is with both arguments:

(21) zon buq’ b-e<r>ku-r-ši w-i
1SG.ABS grain(III)SG.ABS III.SG-<IPFV>sort-IPFV-CVB 1SG-be.PRS
I (male) am sorting grain (Kibrik 1977a:187).

The auxiliary wi agrees in gender and number with the agent zon ‘I’, whereas the main verb berkurši ‘sort’ agrees in gender and number with buq’ ‘grain’. This is a morphosyntactic condition on agreement; the aspectual characteristics license the unusual agreement.

Archi also has the more familiar semantic agreement. In (22) the verb is plural, though the controller is singular:
Conclusion

The agreement system of Archi is as rich as its phonetic and morphological systems and presents severe challenges for current syntactic theories. We have presented an overview of agreement in Archi emphasizing those phenomena that are particularly difficult: agreement going beyond the expected syntactic domains, the problem of lexical coverage and the issue of establishing a grammatical feature (person) where the evidence is limited.

Abbreviations.

1 first person
2 second person
I first gender
II second gender
III third gender
IV fourth gender
ABS absolutive
ATR attribute
AUX auxiliary
CVB converb
DAT dative
ERG ergative
ESS essive
EVID evidential
GEN genitive
HUM human
IN in-localisation (inside hollow space)
INF infinitive
INTER inter-localisation (inside filled space)
IPFV imperfective
LAT lative
LOC locative
NEG negative
NHUM non-human
OBL oblique
PFV perfective
PL plural
PN personal name
PRS present
PST past
SG singular
SUP superlative

(22) os-mi-s χir os b-eš< r >šu-Ši-bi
one-SG.OBL-DAT behind one-SG.ABS HUMPL-IPFV run-IPFV-CVB HUMPL-be.PRS
One is running after another.
References.


