Emergence and evolution of exceptional valency patterns in ergative languages: a case study

Jérémy Pasquereau
University of Massachusetts
jepasque@hotmail.com

Denis Creissels
University of Lyon
denis.creissels@univ-lyon2.fr
http://deniscreissels.fr

1. Introduction

This paper analyzes the possible origin and the evolution of a particular type of exceptional valency pattern found in East Caucasian languages. It deals specifically with a group of East Caucasian languages, the Andic languages (Andi, Akhvakh, Bagvalal, Botlikh, Chamalal, Godoberi, Karata, and Tindi), spoken in the Western part of Daghestan. The Andic languages constitute a genetic unit within the Avar-Andic(-Tsezic) branch of the East Caucasian language family.

The data we present comes from our own work on two Andic languages, Akhvakh and Karata – (Creissels 2010), (Pasquereau 2010), (Pasquereau 2011), and from the consultation of dictionaries and grammars of Andic languages – (Kibrik 1996), (Kibrik 2001), (Magomedova 1999), (Magomedova 2003), (Magomedova 2004), (Magomedova & Abdulaeva 2007), (Magomedova & Khalidova 2001), (Saidova 2006).

The paper is organized as follows. Section 2 provides basic information about the grammatical structure of Andic languages. Section 3 provides illustrations of exceptional case frames in Andic languages. Section 4 examines the question of the emergence of the exceptional case frames <ERG, LOC> and <ERG, ALL>. Section 5 and 6 concentrate on the particular cases of 'listen' and 'bite', respectively. Section 7 puts forward some concluding remarks.

2. The basics of Andic morphosyntax

In Andic languages, the syntactic function of noun phrases is expressed by case suffixes. Verbs agree in gender and number (not in person) but verb agreement is redundant with case marking, since verbs consistently agree with their nominative argument, and never express agreement with non-nominative noun phrases. Constituent order is remarkably flexible and plays no role in the expression of argument structure.

Andic languages have rich case systems, and in particular, very elaborate spatial case systems. The nominative (alias absolutive), also used as the quotation form of nouns, is characterized by a zero ending. The other cases are marked by suffixes attached either to a stem coinciding with the nominative form, or to a special stem traditionally called *oblique stem* (indicated by $-_o$ in the glosses of examples). Spatial case endings consist of two formatives, a topological marker (glossed TPL) and a directionality marker.

As illustrated by Ex. (1), the coding of core syntactic roles in Andic languages is consistently ergative:

– S (*boʁoše ãdo* in (1a)) and P (*istaka* in (1b)) are in the nominative case (alias absolutive), contrasting with A (*wašode* in (1b)) in the ergative case.
– Transitive verbs do not agree with A, but agree with P in the same way as intransitive verbs do with S: in (1a), the intransitive verb shows human plural agreement; in (1b), A is masculine singular, whereas P is neuter singular, and the verb shows neuter singular agreement.

(1) a. *b-oʁoše ãd-o   atobuši-ɬi-g-e b-eq'-īri.*   Akhvakh
    H⁺-most person-PL   bus-N₀-TPL-LOC H⁺-come-PF.H⁺
    'Most people came by bus.'

   b. *wašo-de   istaka   b-iqʷ-aj-ē   godi.*
    boy₀-ERG   glass   N-break-CAUS-ADV.N   COP.N
    'The boy broke the glass.'

The valency frames of Andic verbs canonically include at least a slot for a nominative NP (representing in particular the sole argument of monovalent verbs and the P argument of prototypical action verbs). According to the account of verb valency included in (Kibrik 2001: 369-376), in Bagvalal, this rule can be viewed as exceptionless. However, most Andic languages have a restricted set of verbs used in constructions in which no slot for a nominative NP can be posited.

3. Verbs without nominative arguments

In this section, we provide illustrations of Andic verbs used in case frames that do not include a slot for a nominative NP. Note that, in the absence of a nominative argument, the verb forms that express gender-number agreement show default neuter singular agreement.

3.1. Monovalent verbs

Ex. (2) illustrates the case of a monovalent Akhvakh verb which, depending on the situation referred to, may occur either in the canonical case frame <NOM>, as in (2a), or in the exceptional case frame <LOC>, as in (2b).

(2) a. *miq'i   q'ʷaraɬ-ēhe   godi.*   Akhvakh
    road   become_narrow-ADV.N   COP.N
    'The road became narrow.'

   b. *miʕa-q̄-e   q'ʷaraɬ-ēhe   godi.*
    nose-TPL-LOC   become_narrow-ADV.N   COP.N
    'My nose is blocked.' lit. 'In the nose became narrow.'

Ex. (3) illustrates the Karata verb *bibāɫa* 'cry', used in the exceptional case frame <ERG>.

(3) *gugu-l   ɫabc'e   bib-ē.*   Karata
     cuckoo-ERG   three_times   cry-PF
     'The cuckoo sang three times.'

3.2. Bivalent verbs

Ex. (4) illustrates a bivalent Akhvakh verb used in the exceptional case frame <ERG, GEN>.

(4) *hu-s̄ʷ-e   daru-ɫi-īi   ħalbiχil-āri.*   Akhvakh
     DIST-M$_o$-ERG   medecine-N$_o$-GEN   try-PF
     'He tried the medecine.'

3.3. Trivalent verbs

Ex. (5) & (6) illustrate a trivalent verb used in the exceptional case frame <ERG, ALL, GEN>, in Karata and in Akhvakh.

(5) *hu-ɫi-īi   čela žo-ɫi   mač-uwa du-g-a   de-de*   Akhvakh
     DIST-N$_o$-GEN   other day-N$_o$[LOC] tell-POT   2SG-TPL-ALL   1SG-ERG
     'I will tell you about this another day.'

(6) *k̄'ʷāmsʷam   hedela-ɫi-ī bas-imīse!*   Karata
     trivial.N   thing-N-GEN tell-PROH
     'Don't speak about trivial things!'

4. The emergence of the case frames <ERG, ALL> and <ERG, LOC>

In this section, we examine the case of verbs used in the case frames <ERG, ALL> or <ERG, LOC>. In Andic languages these case frames are typically found with verbs expressing the following meanings: 'look at' – Ex. (7), 'listen', 'bite', 'pinch' – Ex. (8), 'sting' – Ex. (9). For examples with 'listen' and 'bite', see Sections 5 & 6, where the case of these verbs is examined in more detail.

(7) *wašo-de   di-g-a   eq̄-ari.*   Akhvakh
     boy$_o$-ERG   1SG$_0$-TPL-ALL   look_at-PF
     'The boy looked at me.'

(8) *o-š̆ʷ-i   č'un-o di-č'-i.*   Tindi
     DIST-M$_o$-ERG   pinch-PF 1SG$_o$-TPL-LOC/ALL
     'He pinched me.'

(9) *di-g-e* *ƚ'ižʷali-de* *č'in-āri*   Akhvakh
     $1SG_o$-TPL-LOC   wasp$_o$-ERG   sting-PF
     'A wasp stung me.'

Godoberi also has a verb *obadi* 'kiss' with the case frame <ERG, LOC>, but all the other Andic languages for which we have data express this meaning by means of a noun *oba* or *ōba* 'kiss' combined with a verb 'do', and consequently the exceptional case frame of Godoberi *obadi* 'kiss' must probably be viewed as the result of the univerbation of a *do*-compound.

The other cases are less easy to explain, but the variations observed in the expression of the other meaning listed above in Andic languages support a hypothesis already suggested by Charachidzé (1981) for Avar, according to which these exceptional valency frames may result from the reduction of the regular frames <ERG, NOM, ALL> or <ERG, NOM, LOC>, characteristic of verbs expressing meanings or the type 'X applies/holds Y on Z' (and found in particular with the verbs expressing 'hit', with the hittee in the locative/allative and the instrument in the nominative).

The point is that not all Andic languages have bivalent verbs with meanings such as 'look at', 'listen', 'bite', 'sting', 'pinch'. Among the Andic languages, it is also common to find such meanings expressed by means of less specific verbs combined with three NPs. For example, Tindi expresses 'sting' as *eq̄ʷa k̄ʷēƚ'a*, lit. 'hit the sting (on someone)', and 'bite' as *saldi bix̌ʲiƚʲa*, lit. 'hold the teeth (on someone)', with respectively *eq̄ʷa* 'sting (noun)' and *saldi* 'teeth' occupying the nominative slot.

Starting from the hypothesis of valency frames originally including three slots, at least two reduction scenarios can be imagined: conventionalization of the ellipsis of the nominative argument, or fusion of the nominative argument with the verb. In Andic languages, there is evidence for the conventionalization-of-ellipsis scenario in the case of other verbs with exceptional case frames, but not for those examined here. By contrast, some of them at least show evidence for the fusion scenario. The clearest cases are those of 'listen' and 'bite', examined in Sections 5 and 6.

5. Emergence and evolution of verbs 'listen' in Andic languages

Three situations are found among Andic languages with respect to the expression of 'listen':

– Some Andic languages express 'listen' by means of a construction involving a nominative noun phrase with the meaning 'ear' in addition to those encoding the two participants. For example, Godoberi expresses 'listen' as *hãt'uk'ja rik̄i*, literally 'fix the ear (on someone/something)' – Ex. (10). Formally, this construction is an instance of the regular valency pattern <ERG, NOM, ALL> with *hãt'uk'ja* 'ear' filling the nominative slot.
– Others have a verb 'listen' with the exceptional case frame <ERG, ALL>: Tindi *anix̄ʲiƚʲa* – Ex. (11), Chamalal *woƚuk'la*;
– A verb 'listen' with the regular case frame <NOM, ALL> is found in three Andic languages: Akhvakh *hãdax̄uruᴌa* – Ex. (12), Karata *ãdukaƚa* – Ex. (13) , Bagvalal *aštila*.

(10) *wašu-di*   *imu-q̄-i*   *hãt'uk'ʲa* *r-ikk-i* *r-ukk-ida.*   Godoberi
son$_o$-ERG   father$_o$-TPL-LOC/ALL   ear   N⁺-hold-INF N⁺-must-IPF
'The son must listen to his father.'

(11) *di-q̄-a*   *anix̄ʲ-ā*   *hik'ʲi* *o-š̌ʷ-ī*   Tindi
1SG$_o$-TPL-LOC/ALL   listen-IPF   NEG   DIST-M$_o$-ERG
'He does not listen to me.'

(12) *waša* *imo-g-a*   *hãdax̄-ari.*   Akhvakh
boy   father$_o$-TPL-ALL   listen-PF
'The boy listened to his father.'

(13) *waša* *imo-χa-r*   *ãduk-e.*   Karata
boy   father$_o$-TPL- ALL   listen-PF
'The boy listened to his father.'

Interestingly, 'fix the ear on' is the obvious etymology of Akhvakh *hãdax̄uruʟa* (compare with *hãde* 'ear', *bix̄uruʟa* 'fix'), in spite of the fact that the nominative case assigned to the NP representing the listener does not correspond to what could be expected from this etymology.

The variation in the expression of 'listen' in Andic languages provides therefore evidence supporting the reconstruction of the following evolution:

– at a first stage, the coalescence of a trivalent verb occurring in the regular frame <ERG, NOM, ALL> with a noun occupying the nominative slot creates a bivalent verb with the exceptional case frame <ERG, ALL>;
– at a second stage, attested by Akhvakh and Karata, the exceptional valency pattern resulting from this evolution may be regularized into <NOM, ALL>.

6. Variations in the case frame of 'bite' in Andic languages

Among Andic languages, variations similar to those observed in the expression of 'listen' are attested for the expression of 'bite' too, with however a different case frame in the languages in which the case frame of 'bite' has been regularized:

– Some Andic languages express 'bite' by means of a construction involving a noun phrase with the meaning 'tooth' in addition to those encoding the two participants: as mentioned above, Tindi expresses 'bite' as *saldi bix̄ʲiⱡʲa*, lit. 'hold the teeth (on someone)'.
– Others have a verb 'bite' with the exceptional case frame <ERG, ALL>: Karata *q̄'ʷaraⱡa* – Ex. (14), Akhvakh *q̄'eleč'uruʟa* – Ex. (15), Chamalal *q̄'āna* – Ex. (16);
– A verb 'bite' with the regular case frame <ERG, NOM> is found in two Andic languages: Godoberi *q̄'ami* – Ex. (17),[1] Bagvalal *saⱡila* – Ex. (18).

---

[1] Note however that the two available sources on Godoberi give contradictory indications about this verb. Saidova (2006) provides several examples of its use, all with the case frame <ERG, NOM>, whereas Kibrik & al. mention it as a verb used in the case frame <ERG, LOC>, but give no example.

(14) *χʷaj-ol q̄’ʷar-e di-č’o.*   Karata
  dog-ERG bite-PF 1SG$_0$-TPL[LOC]
  'The dog bit me.'

(15) *χʷe-de di-g-e   q̄’eleč’-ari.*   Akhvakh
  dog-ERG 1SG$_o$-TPL-LOC bite-PF
  'The dog bit me.'

(16) *χʷāj-d q̄’ān-nida o-s̄u-č’.*   Chamalal
  dog-ERG bite-PF DIST-M$_o$-TPL[LOC]
  'The dog bit me.'

(17) *χʷaji-di q̄’am-i ho-w.*   Godoberi
  dog-ERG bite-PF$_1$ DIST-M
  'The dog bit him.'

(18) *ʕamo-r salli di-b lela*   Bagvalal
  donkey$_o$-ERG bite.PF$_1$ 1SG$_o$[GEN]-N hand
  'The donkey bit my hand.'

A first interesting observation is that Godoberi *q̄’ami* and Chamalal *q̄’āna* 'bite' do not have the same construction but come from a common root that can be reconstructed as **q̄’am*. Consequently, the difference observed in their case frames cannot be attributed to their etymology, and a change must have occurred in the construction of one of them.

A second interesting observation is that Bagvalal *salīla* 'bite' is quite obviously cognate with Bagvalal & Chamalal *salʷ*, Karata *sale*, Tindi *salu*, Andi *sol* 'tooth'. This suggests that this verb results from the univerbation of a 'noun + verb' compound similar to Tindi *saldi biχʲiɬʲa*, in spite of the fact that the nominative case assigned to the NP representing the bitee does not correspond to what could be expected from this etymology.

The variation in the expression of 'bite' in Andic languages provides therefore evidence supporting the reconstruction of the following evolution:

- at a first stage, the coalescence of a trivalent verb occurring in the regular frame <ERG, NOM, ALL> with a noun occupying the nominative slot creates a bivalent verb with the exceptional case frame <ERG, LOC>;
- at a second stage, attested by Bagvalal, the exceptional valency pattern resulting from this evolution may be regularized into <ERG, NOM>.

The following observations can also be made about the verbs expressing 'bite' in Andic languages:

- Several Andic languages express 'eat' by means of verbs that are reflexes of a root **q̄’am*: Akhvakh *q̄’ōnuɬa* 'eat', Karata *q̄’amaɬa* 'eat', Bagvalal *q̄’anila* 'eat'. As mentioned above, this root also has reflexes expressing 'bite' in Godoberi and Chamalal. Interestingly, the irregular case frame observed with the reflexes of this

root expressing 'bite' is not found with the reflexes expressing 'eat', which always occur in the case frame <ERG, NOM> characteristic of verbs encoding prototypical transitive events.

– As illustrated by Ex. (15) above, the Akhvakh verb *q̄'eleč'uruɫa* 'bite' is used in the case frame <ERG, LOC> to encode prototypical biting events that may cause pain but do not result in a change of state of the bitee (for example, 'The dog bit me'). However, when referring to biting events that affect the physical integrity of the bitee (bite off a piece of something and eat it), *q̄'eleč'uruɫa* is also attested with the case frame <ERG, NOM> characteristic of the expression of prototypical transitive events, as in Ex. (19).

(19) *wašo-de* *ʕeče* *q̄'eleč'-ari.*
boy$_o$-ERG apple bite-PF
'The boy bit off a chunk of apple.'

7. Conclusion

Andic languages are among the languages making a wide use of lexicalized combinations 'noun + verb' in which the noun most commonly occupies the same syntactic slot as the patient in prototypical transitive predication. Since Andic languages are consistent ergative languages, the nominal element of such compounds is in the nominative case and governs verb agreement. Consequently, the univerbation of 'noun + verb' compounds in Andic languages yields verbs whose construction includes no slot for a nominative NP governing verb agreement, creating thus potential exceptions to the rule according to which, in Andic languages, the valency frames of verbs must minimally include a nominative term controlling gender-number agreement.[2]

The first conclusion that can be drawn from the data presented above is that, in the evolution of languages, exceptional valency patterns that emerge as the accidental result of lexicalization processes (in the case examined here, the univerbation of lexicalized 'noun + verb' combinations in ergative languages) may subsequently undergo a regularization process by means of a change in the encoding of one of the arguments: in the case of 'listen', an argument originally encoded as an ergative NP takes nominative marking, and in the case of 'bite', an argument originally encoded as a locative NP takes nominative marking.

Moreover, the difference observed between 'listen' and 'bite' suggests that semantic factors condition this regularization process. In the case of 'bite', the substitution of nominative marking for locative marking results in the case frame <ERG, NOM>, which in Andic languages is used to encode prototypical transitive events involving an agent and a patient. By contrast, in the case of 'listen', regularization results in the case frame <NOM, ALL>, typically used to encode movement towards a goal, and the same is observed with 'look at'. A plausible explanation is that the ergative argument of 'bite' has more affinities with the prototypical agent than the ergative argument of 'listen' or 'look at', and the locative argument of 'bite' has more affinities with the prototypical patient than the allative argument of 'listen' or 'look at'. The ergative

---

[2] See (Haspelmath 1993: 178–180) for a description of this process in Lezgi.

argument of 'listen' or 'look at' is clearly not a typical agent, whereas the allative argument has clear affinities with the goal of movement. In other words, in the regularization process, arguments whose initial case marking has clear semantic motivations tend to maintain their case marking, whereas case marking is more likely to be modified for arguments whose semantic roles can be viewed as relatively ambiguous as regards their affinities with the prototypes underlying the use of cases.

This hypothesis is supported by the variation observed in the case frame of Akhvakh *q̄eleč'uruλa* 'bite', and by the fact that verbs glossed 'eat' cognate with verbs glossed 'bite' uniformly have the case frame <ERG, NOM>, since in the biting events of the type illustrated by Ex. (19) as well as in eating events, in contrast with prototypical biting events, the second participant shows a degree of affectedness similar to that of prototypical patients.

Abbreviations

—$_o$: oblique stem / ADV: adverbial / ALL: allative / CAUS: causative / COP: copula / DIST: distal demonstrative / ERG: ergative / F: feminine / FUT: future / GEN: genitive / H$^+$: human plural / IMP: imperative / INF: infinitive / IPF: imperfective / LOC: locative / M: masculine / nH$^+$: non-human plural / N: neuter / NEG: negation / PF: perfective / PL: plural / POT: potential / PROH: prohibitive / SG: singular / TPL: topological marker

References

Charachidzé, G. 1981. *Grammaire de la langue avar*. Paris: Editions Jean-Favard.

Creissels, D. 2009. 'Valency properties of Northern Akhvakh verbs.' http://www.deniscreissels.fr/public/Creissels-valency_classes_project_Akhvakh.pdf

Haspelmath, Martin. 1993. *A grammar of Lezgian*. Berlin & New York: Mouton de Gruyter.

Kibrik, Alexandr (ed.). 1996. *Godoberi*. München: Lincom Europa.

Kibrik, Alexandr (ed.). 2001. *Bagvalinskij jazyk: grammatika, teksty, slovari*. Moskva: Nasledie.

Korjakov, Ju. B. 2006. *Atlas of Caucasian languages*. Moscow: Piligrim

Magomedova, P. 1999. *Čamalinsko-russkij slovar'* [Chamalal-Russian dictionary]. Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.

Magomedova, P. 2003. *Tindinsko-russkij slovar'* [Tindi-Russian dictionary]. Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.

Magomedova, P. 2004. *Bagvalinsko-russkij slovar'* [Bagvalal-Russian dictionary]. Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.

Magomedova, P & I. Abdulaeva. 2007. *Axvaxsko-russkij slovar'* [Akhvakh-Russian dictionary]. Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.

Magomedova, P & Š. Xalidova. 2001. *Karatinsko-russkij slovar'* [Karata-Russian dictionary]. Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.

Pasquereau, J. 2010. *Spatial cases in Karata: spatial and non-spatial uses*. First year's master thesis. University of Lyon

Pasquereau, J. 2011. *Valency in Karata: a preliminary study*. Second year's master thesis. University of Lyon

Saidova, P. 2006. *Godoberinsko-russkij slovar'* [Godoberi-Russian dictionary]. Maxačkala: Dagestanskij Naučnyj Centr Rossiskoj Akademii Nauk.