

## The two verbal stems in Andi

Steven Kaye, Collège de France (stevenjkaye@gmail.com)  
and

Timur Maisak, IL RAS / HSE University, Moscow (timur.maisak@gmail.com)

### 1. Andi in its linguistic context

- a minority language of southern Russia
- approximately 20,000 speakers across several varieties, all located within the area labelled A
- belongs to the large, indigenous East Caucasian (= Nakh-Daghestanian) family
- one of about 10 languages making up an ‘Andic’ branch of the EC family
- historical attestations date back only to the 19<sup>th</sup> century; some modern descriptions exist
- published texts are a Gospel translation (*Luke*) and a set of world folktales (*Tales*)



Map by Wikimedia user *Jeroenscommons*, licensed by CC-BY-3.0, modified by T. Maisak

### 2. Overview of the talk

Here we will aim to show the following:

- Synchronically, the Andi verb paradigm shows a predictable morphological distribution which has no functional correlate.
- It is possible to make sense of this situation historically, despite the lack of substantial historical data.
- Crucial to an understanding of the current system is the scenario explored by Haspelmath (1998), on the layering of so-called *old presents*: each successive imperfective expression, as it enters the paradigm as a new present form, pushes the preceding one towards the periphery of its earlier semantic domain.
- The Andi verb system preserves notably many distinct diachronic layers of this kind.

### 3. A first look at the Andi verb

Andi verb forms can be synthetic, as in (1), or periphrastic, as in (2):

- (1) ho-b ho-š-t:i **hit'om-ado**, din ho-b **hit'om-ado**  
 DEM-IV DEM-M.OBL-ERG say-PROG I DEM-IV say-PROG  
 {We are having an argument:} 'He says this, and I say this.' (Dirr 1906, texts)
- (2) adam-di **hit'om-ado** **b-ik'o-ddu** sonso-qi :  
 people-ERG say-PROG IV-be-PRF each.other-APUD  
 'People were saying to each other: ...' (Luke 4:36)

The relationship between these types is an interesting question in itself. However, for now we concentrate on the synthetic paradigm.

Synthetic forms can be divided into two groups, illustrated in Table 1 for the verb 'comb':

'comb'	Stem 1 <i>roxo-</i>	Stem 2 <i>roxu-</i>
	<i>roxo-∅</i> AOR 'combed'	<i>roxu-dja</i> FUT 'will comb'
	<i>roxo-rado</i> PROG 'combs, is combing'	<i>roxu-do</i> HAB '(generally) combs'
	<i>roxo-ddu</i> PRF 'has combed'	<i>roxu-daviddu</i> SNP 'still isn't combing (!)'
	<i>roxo-r</i> MSD '(action of) combing'	<i>roxu-dobł:ij</i> ANT.CVB 'before combing'
	<i>roxo-o!</i> IMP 'comb!'	<i>roxu-dos:ub!</i> NEG.IMP 'do not comb!'

Table 1. Some synthetic forms of the verb 'comb' in Gagatli Andi

This distribution of stems over the paradigm is consistent across verb lexemes, although the formal relationship between the two stems is not entirely predictable:

	VOCALIC					NASAL				
	<i>cook</i>	<i>water</i>	<i>count</i>	<i>stay</i>	<i>say</i>	<i>lock</i>	<i>open</i>	<i>rejoice</i>	<i>live</i>	<i>ask</i>
<i>stem 1</i>	-eža-	lento-	ago-	-ek:u-	ruLi-	daši-	arχon-	-erhan-	-aχun-	rac:'in-
<i>stem 2</i>	-eži-	lenti-	agu-	-ek:u-	ruLi-	daš-	arχin-	-erhin-	-aχun-	rac:'in-

Table 2. Some examples of formal relationships between verb stems

Neither stem can be assigned a specific value in synchrony: they are not aspectual stems (seen widely in other branches of East Caucasian), nor do they encode tense, polarity, finiteness... directly.

The remainder of the talk will provide an account of this situation, centring on a closer look at the relationship between the core synthetic forms PROG *roxorado* 'combs, is combing'; FUT *roxudja* 'will comb'; and HAB *roxudo* '(generally) combs'.

Note that all of these can combine with a past-tense copula, as in example (2) above; but the meanings of the resulting formations are not always so predictable.)

#### 4. The ‘old presents’ scenario

Haspelmath (1998) observes that on functional grounds, you might expect the basic present indicative form in a verb system to be morphologically simple, in comparison with more functionally ‘marked’ forms. But this is not always the case, cf.:

Udmurt: PRS.IND *myn-išk-o* ‘I am going’ vs. FUT.IND *myn-o* ‘I will go’

Cairene Arabic: PRS.IND *bi-yi-ktib* ‘he is writing’ vs. PRS.SBJ *yi-ktib* ‘that he write’

The rationale for situations of this kind is diachronic:

- At an earlier stage, the paradigm features a ‘present’ form with relatively broad function: in Udmurt this was a general non-past *myno*, covering both present and future senses.
- The language also makes available a more semantically specific or expressive imperfective formation: in Udmurt, an iterative in *-išk-*.
- When the latter formation enters the core paradigm of the verb, it does not immediately occupy the full domain of usage that was proper to the ‘old present’; instead, it naturally becomes established first in ‘actual’, dynamic present uses.
- As a result, rather than disappearing altogether, the old present may survive, restricted to the periphery of its earlier domain.

Observe that this scenario does not rely on the particular process whereby a ‘new present’ enters the system. This sequence of events has the following ramifications:

- The new range of usage of the ‘old present’ may be best defined *negatively*
- Interactions between the newly morphologized expression and the lexical semantics of certain verbs may give rise to lexical irregularities, which preserve the earlier value of the old present form.
- There is nothing preventing the same process taking place again and again within the same system, even where the diachronic effects of previous iterations of the process are still visible.

We identify evidence for multiple cycles of this kind in the prehistory of the Andi verb system. Specifically, progressing backwards in time:

- 1) Most recently, the PROG form represented by *roxorado* entered the paradigm at the expense of an ‘old present’, which is now labelled as the FUT form *roxudja*.
- 2) This form *roxudja* itself marginalized an ‘older present’, which survives as the HAB *roxudo*.
- 3) In a few verbs, there even remain traces of an ‘oldest present’, pre-dating the establishment of PROG, FUT and HAB.

We now turn to the evidence for this layering.

## 5. The FUT form in *-dja* as an ‘old present’

The central and most obvious function of the form in *-dja* is to express future reference (covering both prediction and intent):

- (3) hoč’o s:edu kabdi-b-š:u-j ič:i-dja!  
 most before enter-PTCP-M.OBL-DAT give-FUT  
 ‘I will give it to the man who arrives first!’ (*Tales: The king’s proud daughter*)

However, there are various indications that future value is not the full story. When it appears as part of a periphrastic form, FUT need have no future value:

- (4) čon-il=lo hege-l-Li guwwati-llo-č’u-k:u sir-dja j-ok’wo-d:u  
 animal(III)-PL=ADD DEM-OBL-GEN strength-OBL-CONT-EL fear-FUT III.PL-be.PL-PRF  
 ‘The animals used to fear its strength’ (*Tales: The lion and the mouse*)

The same form used non-finitely, as a participle, is not futurate, but distinctively *imperfective*:

- (5) q’aj=logu hinu-k:u daš-dja Ɂ<sup>w</sup>amars:a-lo j-ik’ol-d:u, ...  
 belongings=with inside-EL lock-FUT trunk-IN.LAT F-put.SG-CVB  
 ‘Putting her, with her things, in a trunk lockable from the inside...’ (Salimov 2010: 222)

Finally, the *dja* form is the default present form with one class of finite verbs, the modals:

- (6) χ:unt’-u-Li boši-llo-Li milki qala=lo b-ik’u-du b-uk:u-dja  
 pig-OBL-GEN young-OBL-GEN house(IV) castle=as IV-be.SG-INF IV-must.SG-FUT  
 ‘A piglet’s house must be like a castle.’ (*Tales: The three little pigs*)

- All this makes sense if the form in *-dja* was previously a more general non-past form, whose most ‘actual’ function of ongoing present activity was taken over by *-r(ado)*.

In turn, there are indications that the FUT form in *-dja* is not independent of the HAB form in *-do*, but morphologically based upon it. Consider some negative forms in Andi:

	affirmative	negative
aorist	<i>roχo</i>	<i>roχo-s:u</i>
participle	<i>roχo-b</i>	<i>roχo-s:u-b</i>
masdar (verbal noun)	<i>roχo-r</i>	<i>roχo-s:u-r</i>
progressive	<i>roχo-rado</i>	<i>roχo-rado-s:u</i>
habitual	<i>roχu-do</i>	<i>roχu-do-s:u</i>
future	<i>roχu-dja</i> [ < *-do-ja ]	<i>roχu-dos:ja</i> [ < *-do-s:u-ja ]

The pattern seen here suggests strongly that the FUT *-dja* can be decomposed etymologically into two parts, the first being identical with HAB *-do*.

This suggestion is corroborated by the Zilo dialect of Andi, where the HAB forms are (unusually) *-e* and *-e-s:u*, and again appear to underlie the FUT forms: *-ija*, *-es:a*.

## 6. The HAB form in *-do* as an ‘older present’

Along with the morphological indications that FUT (*-dja*) is diachronically based on HAB (*-do*), there is a range of evidence that HAB was once a more general non-past form. Its usual value is indeed habitual:

- (7) iš:i-j                      baba-di                      **q’or-do**                      r-elora                      łoč:’u-do-la  
 we.EXCL-GEN.F                      mother-ERG                      call-HAB                      V-thin                      voice(V)-OBL-SUP.ESS  
 ‘Our mother calls in a high voice!’ (*Tales: The goat and the kids*)

But in (8) it is used for a deliberative question, referring to a single, current situation. There is no special link between this and habitual function; but they make sense as separate remnants of a broader functional domain.

- (8) ha,                      χalgi,                      din                      e<b>=ɸil                      **i-do,**  
 PTCL                      people                      I                      what<IV>=WH                      do-HAB  
 ‘Oh, people, what am I to do?’ (Dirr 1906)

Exceptionally, the verb *-ilin-* ‘go’ uses the HAB form in place of PROG, which is not found:

- (9) i<j>a=l                      men                      **j-ilin-no,**                      hiri                      rag<sup>w</sup>ara?  
 where<F>=WH                      you.SG.ABS                      F-go-HAB                      red                      hat  
 ‘Where are you going, Red Riding Hood?’ (*Tales: Little Red Riding Hood*)

What is more, HAB underlies several other morphological formations in Andi. The ordinal numeral transparently contains a fossilized ‘HAB participle’ in *-do-b* from verb *li-* ‘be, become’:

- (10) łob-**lidob**                      zubu                      he-w                      w-uč:’-idja=ɸoddu  
 three-ORD                      day                      DEM-M                      M-stand.up-FUT=QUOT  
 ‘He will rise on the third day [< \*the day that is three], it says.’ (*Luke 24:46*)

The prohibitive *-dos:ub* appears to represent the negative of this same participle, which no longer exists; a relationship between participles and imperatives is already established for Andi.

- (11) **i-dos:ub**                      di-j                      t’u,                      **k:’w<sup>ab</sup>-dos:ub**                      den  
 do-PROH                      I-DAT                      evil                      kill-PROH                      I.ABS  
 ‘Don’t do me any harm, don’t kill me’ (*Tales: The lion and the mouse*)

Finally, comparative evidence demonstrates that the *-do* form is ancestral to the remarkable ‘Still Not Present’ in *-daɸiddu* (Maisak & Verhees forthcoming):

- (12) hegeł:u-k:u                      b-uč:’in-nu                      **bild-daɸiddu**  
 there-EL                      III.SG-escape-INF                      be.able-SNP  
 ‘[the lion] is (surprisingly) not able to escape from there’ (*Tales: The lion and the mouse*)

The underlying syntagm combines *-do* with an intrinsically ‘negative’ verb (such as ‘avoid’), suggesting that it had general imperfective value when this expression became established.

- It seems likely that the existence of morphological stem 2 is entirely based on the *-do* form and its derivatives; modern-day HAB *-do* is older and ‘deeper’ in the system than both FUT *-dja* and PROG *-rado*.

## 7. Traces of an ‘oldest present’

Intriguingly, a handful of Andi verbs (including ‘look’, ‘smile’ and ‘cry’) seem to preserve yet another formation in this domain. This uses the suffix *-o* (vocalic verbs), *-on* (nasal verbs). Like all the forms considered so far, this can combine with a past copula:

- (13) **j-el:-o**=guža      **č’-on**    j-ok<sup>w</sup>o-d:u    enni-Li=gu...      boc:u-lo  
III.PL-smile-O=ADV look-O    III.PL-be.PL-PRF    REFL.PL.OBL-GEN=EMPH    sibling.OBL-IN.LAT  
‘They were looking at their... sibling and smiling’ (*Tales: The three little pigs*)

This *-o* also underpins another lexically specific formation: *-LO*, found on *-igo-* ‘come’.

- (14) rešno-lo      harč’o-m-ul –    bis:i-j      χos:arlir=ža      **b-igo-LO**  
heaven.OBL-IN.LAT look-IMP-PL    you.PL-DAT    liberation(IV)=EMPH    IV-come.SG-LO  
‘Look up to the sky – your liberation is coming indeed.’ (*Luke 21:28*)

We suggest that *-LO* in fact represents the ancient *o*-present of a separate lexeme, *-el-* ‘walk’, fused on to *-igo-* ‘come’. Parallels elsewhere in Andic (e.g. Godoberi; Kibrik 1996: 47) suggest a very old development.

## 8. Summing up

- The core non-past formations of Andi (and the analytic formations based on them) can be understood diachronically, in terms of Haspelmath’s ‘old presents’ scenario
- On formal and functional grounds, we identify the following progression, working backwards: PROG *-rado* < FUT *-dja* < HAB *-do* (< lexically specific *-o*)
- The purely morphological stem opposition articulating the paradigm appears to reflect a diachronic distinction, between forms based on HAB *-do* (stem 2) and others (stem 1)
- But many questions remain unanswered – including the mysterious origin of PROG *-rado* itself, which has not been addressed here.

## References

- Dirr, A. M. 1906. Kratkij grammatičeskij očerk andijskogo jazyka [A short grammatical sketch of Andi]. *Sbornik materialov dlja opisanija mestnostej i plemen Kavkaza* 36. Tbilisi.
- Haspelmath, M. 1998. The semantic development of old presents: New futures and subjunctives without grammaticalization. *Diachronica*, 15: 29-62.
- Kibrik, A. E. (ed.). 1996. *Godoberi*. Munich: Lincom Europa.
- Luke* = Anonymous. 2015. *Lukašdi bosam rerhanoli χabar* [The Gospel according to Luke]. Moscow: Institut perevoda Biblii.
- Maisak, T., and S. Verhees. Forthcoming. The Still Not Present in Andi: discerning the grammaticalization source. *Folia Linguistica Historica*, 45.1.
- Salimov, Kh. S. 2010. *Gagatlinskij govor andijskogo jazyka* [The Gagatli dialect of Andi]. Makhachkala: IJaLi DNC RAN.
- Tales* = Magomedova, P. A. & M. A. Alisultanova (eds.). 2010. *Dunjalla baχunnirs:i χalgilol gwanab mic’:illas:ol muχal* [Fairy tales from around the world in Andi]. Makhachkala: Nurul’ iršad.