

DGfS 2016 | 24.-26.2.2016
AG2: The syntax of argument structure

Relativization in a morphologically ergative language: a corpus study

Dmitry Ganenkov
University of Bamberg /
Institute of Linguistics RAS, Moscow
d.ganenkov@gmail.com

Lezgian: Introduction

- from the Nakh-Daghestanian (East Caucasian) language family
- spoken in Daghestan, Russia, and Northern Azerbaijan
- ca. 500 thousands speakers
- ergative alignment in argument case marking
- no agreement

Valency classes

intransitive verbs have one core argument in the **absolutive** case. In Dixon's (1994) theory of grammatical functions, the absolutive argument has the function S;

transitive verbs have their subject (A) in the **ergative** case and their direct object (P) in the absolutive case;

experiencer subject verbs have their subject in the **dative** (A), whereas the direct object (P) of such verbs is in the absolutive: 'see', 'hear', 'know', 'want, love', 'find', 'be afraid'

Subjects

Absolute, ergative, and dative subjects pattern together with respect to most subject properties:

- they bind subject-oriented long-distance reflexives
- they are controlled arguments in obligatory control
- they cannot host TAM-auxiliary clitics, in contrast to the absolute object
- etc...

Corpus of Standard Lezgian (CoSL)

- <http://www.dag-languages.org/LezgianCorpus/search/>
- first ever morphologically annotated corpus of a language from the Nakh-Daghestanian (East Caucasian) family
- composed of literary prose from 1930-2009 (67.5%) and newspapers articles from 2009-2012 (32.5%)
- tokens are annotated for lemma, part of speech, and morphological inflectional categories
- the corpus contains more than 4.8 million tokens, of which ca. 87% are provided with annotations

Data

Using the on-line search engine, three types of data on relativization in Lezgian have been extracted from CoSL yielding 11 different data sets:

- (i) a random sample of 2000 relative clauses,
- (ii) a random sample of 150 core argument (subject or object) relative clauses for each of the following five transitive verbs: *kxin* ‘write’, *ecigun* ‘put, build’, *t’ün* ‘eat’, *q:alurun* ‘show’, and *wehin* ‘throw’,
- (iii) a random sample of 150 core argument (subject or object) relative clauses for each of the following five dative subject verbs: *akun* ‘see’, *wan xun* ‘hear’, *žuvun* ‘find’, *k’an xun* ‘love, want, need’, and *čir xun* ‘know, find out’.

Annotation

Relative clauses in the 2000 sample are annotated for the **valency class** of the relativized verb and the **grammatical function** of the relativized argument inside the relative clause (S, A, P, Obl).

The 10 datasets for individual transitive and experiencer subject verbs were coded for the grammatical function of the relativized argument (A or P).

Big sample: results

Frequency of extraction from relative clause broken down by grammatical functions

S	A	P	OBL
787	296	637	280

The preference for relativization on P as compared to relativization on A is highly statistically significant (binomial test, $p\text{-value} < 2.2 \times 10^{-16}$)

Big sample: results

Frequency of extraction of A and P broken down by valency class

	Transitive	Experiential
A	229	67
P	543	94

Transitive verbs: a highly statistically significant preference for relativization on P (p-value $< 2.2 \times 10^{-16}$).

Experiential verbs do not show such a preference (p-value = 0.04012).

Chi-square test confirms the correlation between relativization on core arguments and valency class (p-value = 0.003037).

Small samples: results

Frequency of extraction of A and P in 150 clause samples of five transitive verbs

Verb	A	P	p-value
'write'	22	128	$< 2.2 \times 10^{-16}$
'put'/'build'	29	121	1.462×10^{-14}
'eat'	31	119	2.344×10^{-13}
'throw'	37	113	3.83×10^{-10}
'show'	41	109	2.614×10^{-8}

Small samples: results

Frequency of extraction of A and P in 150 clause sample of five experienter verbs

Verb	A	P	p-value
'find'	43	107	1.772×10^{-7}
'hear'	62	88	0.04087
'see'	66	84	0.1649
'want'	70	80	0.4625
'know'	72	78	0.6832

Frequency of relativization: summary

- The preference for relativization on P is highly statistically significant (p-value < 0.0000001) with transitive verbs.
- Experiential verbs do not constitute a homogeneous class.
- 'Find' is like transitive verbs with respect the P-relative preference,
- other experiential verbs have p-values above 0.01 threshold and thus do not show any statistically significant preference for relativization on either A or P.

Question

Do these counts have any relation to the syntax of Lezgian?

Are usage data (at least partly) motivated by the syntactic structure of Lezgian or do they have another source?

Question

Do these counts have any relation to the syntax of Lezgian?

Are usage data (at least partly) motivated by the syntactic structure of Lezgian or do they have another source?

Newmeyer 2003: no relation view

Bresnan et al. 2001: grammar is of stochastic nature and statistical information must be inherent in grammar, as in stochastic implementations of optimality theory.

Arguments for no relation

We still do not have (?) studies showing whether/how frequency of relativization is related to syntactic alignment

- (i) We have much evidence that Keenan and Comrie's Accessibility Hierarchy works in many languages, i.e. that subject relatives are more frequent than object relatives.
- (ii) However, most studies do not break down subject counts into S and A, so we do not know whether the subject preference is due to S or A, or whether A-relatives outnumber P-relatives in accusative languages and vice versa in ergative languages.

Arguments for no relation: English

Frequency of relativization on particular GFs depends on genre/register

	A	P
Fox 1987 (spoken)	10	46
Fox&Thompson 1990 (spoken)	46	151
Gordon&Hendrick 2005 (Switchboard corpus, spoken)	220	393
Gordon&Hendrick 2005 (Brown corpus, written)	1016	841

Usage frequencies mirror hard constraints?

Morphological ergativity: only coding properties grouping S and P together to the exclusion of A (case marking, agreement)

Syntactic ergativity: behavioral properties

A'-movement constructions (relativization, focus, wh-questions)

Morphological ergativity > Syntactic ergativity

Usage frequencies mirror hard constraints?

- Lezgian is a morphologically ergative language
- S and P are in the absolutive, A is in the ergative

Dide at:a-na.
mother(abs) come-pst
'Mother came.'

Dide-di lam gata-na.
mother-erg donkey(abs) beat-pst
'Mother beat the donkey.'

Usage frequencies mirror hard constraints?

- Lezgian is not a syntactically ergative language, relativization on A is grammatical

— fu t'ür gada
(erg) bread(abs) eat:perf.part boy(abs)
'the boy who ate the bread'

It is tempting to account for the dispreference of A-gap relatives in Lezgian by the same mechanism that bans relativization on A in syntactically ergative languages

Syntax?

Usage dispreferences in relativization result from a grammatical asymmetry which makes production/processing of A-gap relatives more difficult as compared to the production/ processing of P-gap relatives

Probably, yes.

⇒ a similar asymmetry between ergative and dative subjects in complex anaphors

Complex reflexives

Structure of complex anaphors:

REFL-CASE(antecedent) + REFL-CASE(reflexivized)

Syntax of complex anaphors:

Normally, the overt DP is in the case of a higher argument.

Complex reflexives

‘look’ (ABS-DAT)

Ajal-ar za-z kilig-na.

Kid-pl(abs) 1sg-dat look-pst

‘The kids looked at me.’

Ajal-ar čeb čp:i-z kilig-na.

kid-pl(abs) refl:pl(abs) refl:pl-dat look-pst

‘The kids looked at one another.’

*ajal-ri-z čp:i-z čeb kiligna

Complex reflexives

‘say, tell’ (ERG-DAT)

A-da za-z laha-na.

he-erg 1sg-dat tell-pst

‘He told me...’

A-da wič-i wiči-z laha-na.

he-erg refl-erg refl-dat lahana

‘He told himself...’

*ada-z wiči-z wič-i lahana

Complex reflexives

‘kill’ (ERG-ABS)

i-bur-u sew req’i-da.

3-pl-erg bear(abs) kill-fut

‘They will kill a bear.’

i-bur-u čp:-i čeb req’i-da.

3-pl-erg refl:pl-erg refl:pl(abs) kill-fut

‘They will kill one another.’

*i-bur čeb čp:-i req’i-da.

Complex reflexives

'love, want' (DAT-ABS)

Za-z a ruš k'an-zawa.

1sg-dat 3 girl love-prs

'I love that girl.'

A-bur-uz čp:i-z čeb k'an-zawa

3-pl-dat refl:pl-dat refl:pl(abs) love-prs

'They love each other.'

OK! A-bur čeb čp:iz k'anzawa

Complex reflexives: summary

Normally, the overt DP in the case of a higher argument, and the first part of the complex reflexive is in the same case.

ERG-ABS reflexives behave as expected, the overt DP is in the ergative

DAT-ABS reflexives behave exceptionally, since the overt DP may be in the dative or in the absolutive

Case assignment

All case assignment seems to happen low in the structure, presumably, vP-internally (evidence from do-support constructions, infinitival complementation, and nominalizations), cf. Gagliardi et al. (2014), Polinsky (in press)

Dative subjects are clearly theta-related, and the dative is thus most probably an inherent case

Ergative subjects are (at least partly) dissociated from the agent theta-role, as in constructions like
'Fire caught the house' = 'The house caught on fire.'

Case assignment

With transitive verbs, the ergative is a structural case licensed in spec,vP above the absolutive DP.

With subject experiencer verbs, the dative is an inherent case licensed below the absolutive DP. For some reason, the dative DP can optionally move past the absolutive (to an outer spec,vP?)

Transitive verbs vs. SE verbs

Transitive verbs

- ergative is unambiguously higher than absolutive
- ergative is structural (?)
- dispreference for relativization on A

Subject experiencer verbs

- at some point of derivation dative is lower than absolutive
- dative is inherent
- no dispreference for relativization on A

A soft version of syntactic ergativity?

- Dispreference for A seems to stem from the mere fact that the ergative is licensed above the absolutive
- Exactly like in the case of syntactic ergativity
- Still no good theory of syntactic ergativity, see Polinsky (in press)
- In particular, no theory that would disprefer an argument located outside some local domain

Summary

- With transitive verbs, ergative subject gaps are dispreferred as compared to absolutive object gaps
- No dispreference of dative subjects relative to absolutive objects is found with dative subject verbs
- This asymmetry has a parallel in narrow syntactic behavior, namely, binding of complex reflexives

Summary

- Binding of complex reflexives shows that dative subjects apparently have a different derivational history than ergative subjects
- Ergatives are licensed above the absolutive object, datives are licensed below it
- This fact seems to be responsible for the asymmetry in relativization frequencies

Summary

- Binding of complex reflexives shows that dative subjects apparently have a different derivational history than ergative subjects
- Ergatives are licensed above the absolutive object, datives are licensed below it
- This fact seems to be responsible for the asymmetry in relativization frequencies
- After all, Lezgian is a discourse ergative language

THANK YOU!