

NATIONAL RESEARCH UNIVERSITY HIGHER SCHOOL OF ECONOMICS

Yury Lander, Timofey Arkhangelskiy

PRODUCING POLYSYNTHETIC VERB FORMS IN WEST CIRCASSIAN (ADYGHE): AN EXPERIMENTAL STUDY

BASIC RESEARCH PROGRAM

WORKING PAPERS

SERIES: LINGUISTICS

WP BRP 23/LNG/2015

Yury Lander¹, Timofey Arkhangelskiy²

PRODUCING POLYSYNTHETIC VERB FORMS IN WEST CIRCASSIAN (ADYGHE): AN EXPERIMENTAL STUDY³

This paper describes a pilot experiment which was conducted by the authors with speakers of the polysynthetic West Circassian (Adyghe) language and aimed at investigating their ability to use complex verb forms that cross-reference several arguments introduced by applicative morphology. The results of the experiment support the view that complex polysynthetic words can be constructed in the course of speech and do not necessarily belong to any common inventory of word forms. In addition, we make several conclusions which concern productivity of applicatives and their order within the West Circassian verb.

Keywords: West Circassian, Adyghe, polysynthesis, cross-reference, morphological complexity, morphological productivity, morpheme order, applicatives

JEL: Z

¹ National Research University Higher School of Economics, email: yulander@yandex.ru

² National Research University Higher School of Economics, email: timarkh@gmail.com

³ The research carried out in 2014-2015 by Yury Lander was supported by "The National Research University 'Higher School of Economics' Academic Fund Program" grant (No. 14-01-0083).

1. Introduction

Polysynthetic languages, with their tremendously complex morphology, pose a question of the status of complex forms: do such forms pre-exist or are they constructed in the course of speech? To be sure, any language presumably allows constructing words online, but it can be hypothesized that agglutinative polysynthetic languages may have specific morphological means that use this possibility much more actively and more easily than typical morphological means in non-polysynthetic language. We will call this the ad hoc morphology hypothesis.⁴ An alternative hypothesis is that even complex forms are just taken from some "storage" (be it a kind of mental lexicon or some ready paradigms), which may explain occasional defective paradigms and other irregularities, both formal and semantic, if they are found.⁵

In this paper, we provide experimental evidence for the ad hoc morphology hypothesis as applied to applicatives in West Circassian (Adyghe), a Northwest Caucasian language.

The structure of our paper is as follows. In Section 2 we provide the background information on West Circassian applicatives. In Section 3 we describe the experiment we conducted. Section 4 is devoted to the results of the experiment. The last section contains conclusions.

2. West Circassian argument structure morphology

West Circassian (also known as Adyghe) together with Kabardian constitutes the Circassian branch of the Northwest Caucasian family. In general, Northwest Caucasian languages are well-known for their complex morphology (cf. Hewitt 2005), which makes them similar to polysynthetic languages of the Americas and North Australia but at the same time makes them occupy a very specific place in Europe, where polysynthetic languages are not thought to be common.

While there are many aspects of grammar that motivate the morphological complexity of West Circassian (cf. Lander, to appear), here we will only focus on the morphological expression of the participants of the situation described in an utterance. For the sake of simplicity, we will use the term argument for all participants that get some morphological expression, even though Lander (2015) has argued that the argument/adjunct distinction can be found in morphology too, and hence the morphological expression does not entail the argument structure. Further, in this paper we will only consider arguments of the verb and do not discuss, for example, possessors and objects of postpositions.

In general, West Circassian allows cross-referencing for a bulk of arguments. Cross-reference prefixes together with applicatives (see below) constitute a separate argument structure

⁴ The term *ad hoc morphology* used for sporadic morphological processes can be found, for example, in Gensler 2002. Mattisen (2011) also speaks of ad hoc word formation as word formation in the course of speech.

⁵ A brief survey of the issue can be found, for example, in Aronoff and Fudeman 2011: 246-252.

zone of the word form. The language is ergative both in case marking and in verb morphology,⁶ and both the absolutive argument and the ergative argument are cross-referenced in distinctive slots of the word form. The absolutive marker begins the argument structure zone and the ergative cross-reference finalizes it. Between them, one can observe a directive prefix (sometimes functioning as a marker of inversion) and applicative complexes. The canonical applicative complex consists of an indirect object cross-reference prefix and an applicative prefix which often specifies its role, for instance, the benefactive, malefactive, comitative and numerous locative prefixes (for details, see Smeets 1984; Lander, to appear).

Example (1) (taken from our corpus of West Circassian texts) shows all of these parts of the argument structure zone (given in brackets). Here we find the absolutive reflexive prefix in the very beginning of the word and the ergative cross-reference prefix before the causative marker, at the very end of the argument structure zone. Between the absolutive prefix and the ergative prefix, one can see the directive prefix and an applicative complex 'with you', which consists of an indirect object cross-reference morpheme and the comitative applicative prefix.

(1) [zə-qə-b-d-jə]-ʁe-ze-š't

[RFL.ABS-DIR-2SG.IO-COM-3SG.ERG]-CAUS-turn-FUT

'it will turn (lit., make itself turn) together with you'

For a study of the complexity of West Circassian words, applicative complexes are of particular interest, because there may be several applicative complexes in a single word form. This is shown in another example from our corpus of West Circassian texts (2), where the second verb includes a locative applicative and a benefactive applicative. Curiously, the overt indirect object cross-reference marker is "introduced" by the benefactive prefix but is separated from it by the locative applicative, which presumably introduces the null 3rd person singular cross-reference prefix. Such "movement" of the cross-reference prefix is regularly attested, but its rules are not known.

(2) a-ве-wэс^wэ-ве q^waǯ'e-m meš'ət ЗРL.ERG-CAUS-stand-PST village-OBL mosque

_

⁶ Letuchiy (2012) also demonstrated that various morphological processes such as causativization also treat the language as ergative.

⁷ Note that null morphemes are not shown in the examples.

d-a-ş, 9-t-a-ş 9-ra-r

DIR-3PL.IO-LOC-BEN-3PL.ERG-make-PST-PST

'They made a mosque for them in the village founded by them.'

The principles ordering multiple applicative complexes are not fully understood. We contrast between two approaches to the order of applicative complexes. The first of them states that there is a template, i.e. some well-established order of applicatives, and any non-expected orders should be regarded as deviations. For example, Jakovlev & Ashkhamaf (1941: 357) in their first academic grammar of West Circassian noticed that the order of applicatives could vary, yet insisted that there were some orders that were used more frequently and described them with a template model. For example, they noted that the general locative prefix *š'a*-normally precedes the benefactive prefix fe-, and stated that they occupy different slots. Later Kumakhov (1964: 152-157; 1965) also represented the order of morphemes in Circassian languages with a template model, where, by the way, the benefactive applicative precedes the locative applicative. The second approach proposes that no template can be postulated and there are other principles at work here. Indeed, Caponigro & Polinsky (2011) and Lander (2015; to appear) suggested that the order of West Circassian applicative complexes is based on their scope. The arguments for the scope-based approach were based on quite specific contexts where the relative scope of applicative complexes matters. It is worth noting, however, that in most contexts the change of the scope of applicatives does not affect the resulting interpretation.

3. Experiment: basic information

The experiment which we conducted had two goals. First, we wanted to know whether there is any variation in the capability of constructing complex forms involving one or several applicative complexes. Second, we were interested in whether there was any variation in the structure of the forms our informants constructed.

Both questions could shed the light on the issue of whether complex polysynthetic forms should be listed or not. The second question also could support or refute the hypothesis according to which there exists a template governing the order of applicative complexes.

With this in mind, we asked our informants to translate 28 Russian sentences into West Circassian. These sentences had been constructed in such a way that they included a beneficiary, an accompanying participant and/or some location, and hence would require the use of the benefactive, comitative or general locative applicative complexes, or their combinations. In order to exclude the situation of direct calquing, we prohibited the appearance of words that could be used as substitutions for Russian prepositions, namely the West Circassian postposition *paje* 'for' and the verb $B^{w} \partial S^{-}$ 'to accompany'. All indirect objects were of the 3rd person and singular, which means that no overt cross-reference prefixes could appear. This prevented any variation related to the position of cross-reference prefixes (see above). An example including a Russian

sentence (in a simplified transcription) and its West Circassian equivalent is given in (3), where the Russian preposition *dlja* 'for' is translated with the applicative benefactive prefix.

(3) Ajdamyr mjas-o dlja a. reza-l brat-a. Aydamyr(NOM:SG) cut-PST meat-ACC:SG for brother-GEN:SG b. ajdemər f-jə-wəpč'eta-ĸ. lə-r ə-š Aydamyr meat-ABS 3SG.PR-brother BEN-3SG.ERG-cut-PST 'Aydamyr cut the meat for his brother.'

We had two sets of stems on which the verbs used in the experiment were expected to be based on. Each of the sets originally included two transitive stems and two intransitive stems, although for some intransitive stems, we had to replace them with causatives based on them for the sake of naturalness. The list of the stems is given in Table 1.

Set1	Set 2
də- 'sew' (transitive)	wəpç'ete- 'cut' (transitive)
λesə- 'bathe' (transitive)	gəč'ə- 'launder' (transitive)
ǯ'eg ^w ə- 'play' (intransitive)	zewe- 'fight' (intransitive)
к ^w ə- 'get dry' (intransitive), ке-к ^w ə- 'dry' (transitive)	febe- 'get warm' (intransitive), ве-febe- 'warm up' (transitive)

Table 1. Stems used in the experiment.

All participants of the experiment were native speakers of the Temirgoi dialect of West Circassian. All of them were female and the age of the informants varied from 31 to 58 at the time of the experiment. The participants were divided into two groups consisting of five and six persons, respectively. Each group operated with one set of stems and translated the sentences based on that set. No fillers were used. The sentences presupposing the use of the stems were certainly mixed, but the ordering of the sentences as regards their complexity was the same for both groups.

The Russian stimuli appeared as separate slides, and the informants had about 50 minutes for the task. If an informant had any difficulty with translation, some expected forms were suggested, and they were either accepted or considered infelicitous. Sometimes an alternative

variant of a form was suggested for a speaker even if some variant had already been produced, but this was by no means a rule.⁸

The results of each translation attempt were divided into three classes:

- (i) CONSTRUCTED: the forms constructed by an informant,
- (ii) APPROVED: the forms that had been constructed by us and approved by an informant,
- (iii) PROHIBITED: this class included the cases where any form for a given meaning had been considered infelicitous.

We further counted the tokens of each class. If there were several variants, we counted the variant that occupies the higher place in the hierarchy CONSTRUCTED > APPROVED > PROHIBITED. This means that, for instance, if there were two variants, one of which was suggested by us and approved by the informant and another was constructed by an informant, we only counted it as a member of the latter class.

According to the same principle, prohibitions were only counted if no alternative possible variants of a form existed. For the remaining prohibitions, we are not aware of their motivation. In fact, it could be that they resulted from incompatibility of specific applicatives with a specific stem. The case in point is related to the stem $\check{\mathbf{z}}'eg^{w}\partial$ - 'play', which showed low compatibility, first of all, with the benefactive prefix, but also with some other applicatives. It is likely that this was due to the fact that this word is more associated with a certain complex ritual (which probably does not give rise to the applicatives discussed here)⁹ than with a simple play.

It turned out that the results for Set 1 and Set 2 differed considerably; cf. Table 2, where it is shown that the percent for different sets differs more than as predicted by the standard statistical bias. To be sure, this may be due to the difference between our informants, but we hypothesize that this was due to the unsuccessful choice of $\check{z}'eg^{w}\partial$ -: in fact, this stem gave three approved cases for verbs with one applicative, five prohibited cases versus four approved cases for verbs with two applicatives, and two prohibited cases versus one approved case for a verb with three applicatives. Obviously, this verb appeared to affect the statistical results for Set 1 significantly. Hence, the results that were obtained for Set 2 seem more reliable to us.

7

⁸ After that, a few days later, each group got some word forms constructed or evaluated by another group and its members were asked to construct a sentence based on those verb forms. The presupposition was always that such forms had been provided earlier, even where the relevant forms had been prohibited. We do not discuss the results of that experiment here, though.

⁹ See Bgazhnokov 1991 for anthropological discussion of the complex concept of $\check{g}'eg$ ¹¹ ϑ .

4. Results

4.1. Complexity

Table 2 shows that the complexity of a verb correlates with the difficulty of how this verb is constructed. In particular, we observed that the more complex a verb is, the more cases where the informant could not construct this verb themselves.

This result is by no means surprising. It should be noted, however, that to some extent this can be related to the fact that more complex forms are more probable to include less productive affixes (see Section 4.2). This may add noise to the results.

Interestingly, there is no correlation between the complexity of a verb and the (in)felicity of a word form. While the results for Set 1 may suggest the number of the negative evaluations is similar for verbs with two applicatives and verbs with three applicatives, the results for Set 2 demonstrate that this is not the case. It is more likely that the negative results for Set 1 resulted from the lexical incompatibility described above.

4.2. Different productivity

The concept of morphological productivity is very tricky (cf. Bauer 2001 for discussion), and it is impossible to measure the productivity of various affixes strictly with the small amount of data we have. Yet at least it can be said that different applicatives may differ in their productivity, if productivity is considered the ease with which an applicative can be added to the stem.

The results for different applicatives are shown in Tables 3 and 4.

The basic contrast we find is that between the comitative and benefactive on the one hand, and the general locative, on the other. The general locative applicative is added more easily than the two other applicatives. The non-zero prohibition rates for the forms with the general locative probably are due to the presence of other applicatives.

It is worth saying that the results we have may reflect not the productivity *per se* but the semantic restrictions that we are not aware of. However, the fact that there is variation in the acceptance of many forms, suggests that this is unlikely.¹⁰

¹⁰ Note, however, that the semantics of the benefactive applicative is described in detail by Letuchiy (2009).

		Set 1		Set 2	
Forms with one applicative	Problematic, incl.	8	13.33%	3	4.17%
	prohibited	0	0%	0	0%
	approved	8	13.33%	3	4.17%
	Constructed	52	86.67%	69	95.83%
	Total	60	100%	72	100%
Forms with two applicatives	Problematic, incl.	25	41.67%	17	23.61%
	prohibited	8	13.33%	2	2.78%
	approved	17	28.37%	15	20.83%
	Constructed	35	58.33%	55	76.39%
	Total	60	100%	72	100%
Forms with three applicatives	Problematic, incl.	14	70%	10	41.67%
	prohibited	3	15%	0	0%
	approved	11	55%	10	41.67%
	Constructed	6	30%	14	58.33%
	Total	20	100%	24	100%
Total	Problematic, incl.	47	33.57%	30	17.86%
	prohibited	11	7.86%	2	1.19%
	approved	36	25.71%	28	16.67%
	Constructed	93	66.43%	140	83.33%
	Total	140	100%	168	100%

Table 2. Total results.

		de- CO)M	fe- BE	N	š'ə- L	ос
Forms with one applicative	Problematic, incl.	5	25%	3	15%	0	0%
	prohibited	0	0%	0	0%	0	0%
	approved	5	25%	3	15%	0	0%
	Constructed	15	75%	17	85%	20	100%
	Total	20	100%	20	100%	20	100%
Forms with two applicatives	Problematic, incl.	19	47.5%	20	50%	11	27.5%
	prohibited	7	17.5%	7	17.5%	2	5%
	approved	12	30%	13	32.5%	9	22.5%
	Constructed	21	52.5%	20	50%	29	72.5%
	Total	40	100%	40	100%	40	100%
Forms with three applicatives	Problematic, incl.	14	70%	14	70%	14	70%
	prohibited	3	15%	3	15%	3	15%
	approved	11	55%	11	55%	11	55%
	Constructed	6	30%	6	30%	6	30%
	Total	20	100%	20	100%	20	100%
Total	Problematic, incl.	38	47.5%	37	46.25%	25	31.25%
	prohibited	10	12.5%	10	12.5%	5	6.25%
	approved	28	35%	27	33.75%	20	25%
	Constructed	42	42.5%	43	43.75%	55	68.75%
	Total	80	100%	80	100%	80	100%

Table 3. Set 1: Results for applicatives.

		de- COM		fe- BEN		š'ə- LOC	
Forms with one applicative	Problematic, incl.	2	8.33%	1	4.17%	0	0%
	prohibited	0	0%	0	0%	0	0%
	approved	2	8.33%	1	4.17%	0	0%
	Constructed	22	91.67%	23	95.83%	24	100%
	Total	24	100%	24	100%	24	100%
Forms with two applicatives	Problematic, incl.	14	29.17%	13	27.08%	7	14.58%
	prohibited	2	4.17%	2	4.17%	0	0%
	approved	12	25%	11	22.92%	7	14.58%
	Constructed	34	70.83%	35	79.92%	41	85.42%
	Total	48	100%	48	100%	48	100%
Forms with three applicatives	Problematic, incl.	10	41.67%	10	41.67%	10	41.67%
	prohibited	0	0%	0	0%	0	0%
	approved	10	41.67%	10	41.67%	10	41.67%
	Constructed	14	58.33%	14	58.33%	14	58.33%
	Total	24	100%	24	100%	24	100%
Total	Problematic, incl.	26	27.08%	24	25%	17	17.71%
	prohibited	2	2.08%	2	2.08%	0	0%
	approved	24	25%	22	22.92%	17	17.71%
	Constructed	70	72.92%	72	75%	79	82.29%
	Total	96	100%	96	100%	96	100%

Table 4. Set 2: Results for applicatives.

4.3. Order of applicatives

It is not possible to give any definite statistical results concerning the choice of the order of applicatives, since we did not always check all possible orders in word forms with several applicatives. However, if we only consider first answers, it is easy to find some tendencies. Cf. Table 5, which provides data for both sets. Here the first number counts examples with the first applicative given in the line and the second applicative given in the column. The second number refers to the whole number of self-constructing cases for a given combination.

First applicative	š'ə-	fe-	de-
↓	LOC	BEN	СОМ
š'ə- LOC		14/14, 17/20	14/15, 19/20
fe- BEN	0/14, 3/20		3/5, 12/14
de- COM	1/15, 1/20	2/5, 2/14	

Table 5. Forms with two applicatives: the order of applicatives (first answers).

In the absence of context, in forms with two applicatives it is almost always the case that the locative applicative comes first and the benefactive applicative appears before the comitative applicative. This is confirmed by examples with three applicatives where the first answers gave the order LOC-BEN-COM- in 15 cases out of 19 (in sum, for both sets).

Nonetheless, these orders are not obligatory. In most cases, speakers allowed alternative orders, as in (4). In one case a speaker even considered an example provided by most other speakers infelicitous and insisted on the alternative order, as in (5) (note that the order LOC-COM-, which is said to be prohibited here, had been given independently for the same base by three informants and approved by two more informants).

(5) č'etə-lə-m kartweške-r kwəxne-m qə-də-š'ə-feba-ʁ /
chicken-meat-OBL potato-ABS kitchen-OBL DIR-COM-LOC-warm-PST
*qə-š'ə-de-feba-ʁ
DIR-LOC-COM-warm-PST
'The potatoes got warm in the kitchen together with the chicken.'

5. Conclusion

The results of our experiment presented above seemingly support the view that complex polysynthetic words can be constructed in the course of speech and do not necessarily belong to any common inventory of word forms. If this view were not correct, we would not expect difficulties in constructing word forms motivated by their complexity. Another piece of evidence for the ad hoc morphology hypothesis comes from the variation in constructing and accepting various forms that is observed between the participants of the experiment. Indeed, we may expect personal preference in the way of building complex constructs but not in the acceptance of words that already exist. It could be hypothesized that speakers may differ in the realizational rules they have, yet the fact that quite often the participants accepted the presence of several variants of a form contradicts this. We thus conclude that the West Circassian applicative morphology indeed can be actively used in constructing forms online. In this respect, it clearly resembles syntactic means in non-polysynthetic languages.

However, the experiment also showed that there are some properties of the West Circassian applicative morphology that make it more similar to word formation in non-polysynthetic languages. In particular, it turns out that applicatives differ in productivity, which is not expected for compositional syntactic (or syntactic-like) components, but is certainly natural for word formation.

Another interesting result of the experiments concerns the order of applicatives. In particular, it turned out that there are some clear preferences in ordering the three applicatives we explored. These preferences do not, however, constitute any strict template, as alternative orders are possible as well.

Still, we regard our experiment only as a preliminary step towards a broad psycholinguistic investigation of the peculiarities shown by polysynthetic languages in general and West Circassian in particular. Obviously, much more contexts should be investigated in order to make it clearer how the polysynthetic morphology functions, in what aspects it resembles syntax and in what aspects it nonetheless remains morphology.

Abbreviations used in glosses

ABS - absolutive; ACC - accusative; BEN - benefactive; CAUS - causative; COM - comitative; DIR - directive; ERG - ergative; FUT - future; GEN - genitive; IO - indirec object; LOC - locative preverb; NOM - nominative; OBL - oblique case; PL - plural; PR - possessor; PST - past; RFL - reflexive; SG - singular.

References

- Aronoff, Mark and Kirsten Fudeman. 2011. *What is Morphology?* 2nd edition. Wiley-Blackwell. Malden, Oxford: Wiley-Blackwell.
- Bgazhnokov, Barasbi Kh. 1991. Cherkesskoe igrische. Sjuzhet, semantika, mantika. Nalchik.
- Caponigro, Ivano and Maria Polinsky. 2011. Relative embeddings: a Circassian puzzle for syntax/semantics interface. *Natural Language and Linguistic Theory* 29 (1): 71—122.
- Gensler, Orin D. 2002. Why should a demonstrative turn into a preposition? The evolution of Welsh predicative *yn. Language* 78 (4): 710-764.
- Hewitt, George. 2005. North West Caucasian. Lingua 115: 91-145.
- Jakovlev, Nikolay and Daud Ashkhamaf. 1941. *Grammatika adygejskogo literaturnogo jazyka*. Moscow, Leningrad: Izd. AN SSSR.
- Kumakhov, Mukhadin A. 1964. *Morfologija adygskix jazykov. Sinxronno-diaxronnaja xarakteristika*. I. *Vvedenie, struktura slova, slovoobrazovanie chastej rechi*. Nalchik: Kabardino-Balkarskoe knizhnoe izd.
- Kumakhov, Mukhadin A. 1965. Distributivnyj analiz polisinteticheskogo kompleksa. *Voprosy jazykoznanija* 5: 112-117.
- Kumakhov, Mukhadin and Karina Vamling. 2009. *Circassian Clause Structure*. Malmö: Malmö University.
- Lander, Yury. 2015. Aktanty i sirkonstanty v morfologii i v sintaksise adygejskogo jazyka. *Vestnik RGGU. Ser. Istorija. Filologija. Kul'turologija. Vostokovedenie* 1: 7-31.
- Lander, Yury. To appear. Adyghe. In: Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds), *Word Formation, An International Handbook of the Languages of Europe*. Vol. 5. Berlin: Mouton de Gruyter.
- Letuchiy, Alexander. 2009. Affiksy benefaktiva i malefaktiva: sintaksicheskie osobennosti i krug upotreblenij. In: Yakov G. Testelets (ed.), *Aspekty polisintetizma: ocherki po grammatike adygejskogo jazyka*. Moscow: Izd. RGGU, 329-371.

Letuchiy, Alexander. 2012. Ergativity in the Adyghe system of valency-changing derivations. In: Gilles Authier and Katharina Haude (eds), *Ergativity, Valency and Voice*. Berlin: Mouton de Gruyter, 323-354.

Mattisen, Johanna. 2011. Complexity in nouns. Rivista di Linguistica 23 (1): 79-103.

Smeets, Rieks. 1984. Studies in West Circassian Phonology and Morphology. Leiden: The Hakuchi Press.

Yury Lander

National Research University Higher School of Economics, email: yulander@yandex.ru

Timofey Arkhangelskiy

National Research University Higher School of Economics, email: timarkh@gmail.com

Any opinions or claims contained in this Working Paper do not necessarily reflect the views of HSE.

© Lander, Arkhangelskiy, 2015