

# Features of Annotation of Verb Complements in Different Stages and Varieties of Armenian

**Anna Danielyan**

Yerevan State University  
Faculty of Oriental Studies  
Department of Arabic Studies

**Marat Yavrumyan**

Paris London University of Salzburg  
Centre for Research of the Christian East  
Armenian Studies

*Relevant UniDive working groups:* WG1

## 1 Introduction

This is a research project idea aimed at developing a combined annotation system and a tagset of linguistic universals for verb complements for all varieties of Armenian, utilizing the NLP-applicable universality of terminologies and methodologies (typologically oriented grammatical theories and dependency grammar). In this context it's an intra-language diversity research, in terms of differences exhibited within Armenian as Standard language in different historical stages on one hand, and Modern Standard Armenian vs. non-standard varieties on the other. The diachronic typology can raise specific universals that will improve cross-treebank consistency, and be more sensitive to the traditional grammar. The research will be built upon previous experiences of Universal Dependencies projects<sup>1</sup> that provides a community accepted and trusted concept for language modeling and processing, unified across many languages while preserving their diversity.

The purpose of the research is to promote the improvement of digital sets of tags and relations developed for Armenian, that are applicable in the fields of Natural language processing (NLP). The results can further be used in other similar researches or in projects with wider perspective like developing a syntactically annotated diachronic and/or parallel corpora for the Armenian language.

## 2 Verb complements in Armenian

In Armenian grammatical theory, the argument-adjunct distinction is basically semantic. Arguments are determined as nominals related to verbs, and adjuncts are circumstances related to verbs (for both eastern and western modern standards (Asatryan, 1987), (Yeghiaeanean, 2022)). Morphologically, there are prepositional, pronominal and noun-case complements that are generally distributed in

one of these subcategories: core-non-core; arguments of voice and nature; declensional and prepositional complements, nominal and adverbial arguments. Some authors distinguish the adjuncts into two groups - core adjuncts and adverbial arguments, where the adverbial arguments occur on the border between arguments and adjuncts. Syntactically these complements are expressed by a clause in the syntactic position of a complement. A special attention is paid on the governing capacity of verbal nouns, participles and adjectives. The key factor of the classification of verb complements (arguments and adjuncts) in different theories (governing, dependency and valency) is the grammatical category of voice (besides the languages where the subject-object roles are determined neither by verb morphology nor by the verb categories transitive-intransitive, but by the SVO positional devices (see Pasierbsky (2003))). In modern standards of Armenian, mainly 3 varieties of voice are distinguished: active, passive and middle, but sometimes there are 4 (including the causative, (Abrahamyan et al., 1974), (Arak'elian, 1967)) or even 6 (adding reflexive and reciprocal (Aghayan, 1967)) voices distinguished. Active and causative verbs are transitive, while passive (reflexive, reciprocal) and middle verbs are considered intransitive. Transitive verbs can have core arguments in accusative (when nonhuman) and in dative (when human), regardless of only one or both dependents are used in the sentence. At the same time a large number of non-core arguments (and some adjuncts) stand in dative, too. Due to the general meaning of verbs that arguments complete, there are several subgroups distinguished. As the semantic variety of these subgroups is very wide, sometimes these arguments are presented in one group called 'indirect objects in dative', with detailed explanation and determination inside the group (Abrahamyan et al., 1975). Adjuncts in dative are distinguished semantically or through questions.

<sup>1</sup>In the framework of UD syntactically annotated corpora were developed for both [Eastern](#) and [Western](#) Armenian

### 3 The frame of issues in brief.

The above-mentioned similarity of direct object (human) and indirect object in dative raises some difficulties concerning core-non-core distinction. On the other hand, in some varieties of Armenian human-denoting dependents of transitive verbs as well stand in accusative in the syntactic position of the direct object (this is widely used in Western Armenian and in some dialects) and the core - non-core distinction is mainly semantic (for this issue see (Przepiórkowski and Patejuk, 2018)). Besides it, there are some inconsistencies between definitions in grammatical theory and the language material. For instance, middle verbs are intransitive and can have arguments in all the cases except the accusative. However, in literature we find examples, where verbs marked in dictionaries as intransitive evidently have an argument in accusative, as the verb *lal* ‘to cry’ in the following:

*ir meġk'n ē lalis*  
s/he-GEN fault-ACC be-3SG cry-PTCP  
‘(s/he) cries his fault’

For such cases we suggest to edit the description of verbs like this in the dictionaries and mark them as having both middle and active voices.

There are some inconsistencies in Armenian concerning the arguments’ determination, generally in causative and passive constructions. This part of grammar is a bit complicated considering the morphological structure of verbs and definition of transitivity or intransitivity in Armenian grammar. For instance, the verb *nšanakel* ‘to appoint’ (as well as other similar verbs like *ëntrel* ‘to elect, to choose’, *arajadrel* ‘to nominate, to put forward’ etc.) is ditransitive and when transforming into passive, can’t lose its second direct object for not to have semantic deficit or incompleteness. Hence, the claim that passive (or passive-like<sup>2</sup>) verbs can’t be transitive is not convincing. In UD we use `xcomp` relation for the dependents of these verbs, considering them as a secondary predicate.

As for causative verbs, they can be formed from both middle and active verbs with the infix *<c'n>* or in combination with the verb *tal* (to make smb

do smth), or both. These verbs are transitive and vary in valency - middle-based causatives are transitive, while active-based causatives can be ditransitive and even tritransitive. The latter ones are described as real causatives, ditransitives or verbs of causative voice. On the other hand, there are a number of causative-like verbs that have lost their causative meaning, like *hatkac'nel* ‘to provide’, *korc'nel* ‘to lose’, *patkerac'nel* ‘to imagine’, and some active verbs, that have the same meaning with or without the infix *<c'n>*, like *halel - halec'nel* ‘to melt’, *pttel - pttec'nel* ‘to turn, to screw’, *mašel - mašec'nel* ‘to wear out, to use up’ etc. Obviously, morphological structure of these verbs isn’t a sufficient condition for them to be considered causative. The full correspondence of causative form and meaning have verbs used in combinations with the verb *tal* (*kardal tal* ‘to make read’, *berel tal* ‘to make bring’ etc.).

The main difficulty in sentences with real causative verbs concerns the semantic role of actants and core arguments distinction. Sometimes there is not a perfect match between the predicate in causative and the subject, which isn’t the real doer of the action, but rather the reason or impulse for another actant (the syntactic object and semantic subject) to do it, like:

*na vordown t'ey ē*  
*xmec'nowm*  
s/he son-DAT tea-ACC be-3SG  
drink-PTCP.CAUS  
‘S/he gives the son tea to drink’  
(lit. s/he makes the son drink tea)

The infix *<c'n>* modifies the lexical meaning of the verb establishing the semantic frame of the sentence in the same way the passive infix *<v>* does. For real passive constructions there is a separate type of argument called ‘active argument’<sup>3</sup>, that mentions the semantic subject of the sentence. So, if the semantic role of this syntactic position is taken into account for passive predicates, it’s not clear why a similar approach can’t be used for real causative ones. Abrahamyan (1962) has paid attention to this issue and offered to name this argument a ‘mediatory argument’<sup>4</sup>.

As for the core argument distinction, in sentences like this, when the syntactic position of

<sup>2</sup>Reflective and reciprocal verbs, that have the morphological marker of passive (*<v>*) are called passive-like middles, as they are intransitive (or monovalent). Some authors, as Gyulzatyan (2015) argue, that these verbs are semantically transitive, as the action is directed to and affects the subject itself.

<sup>3</sup>See `obl:agent` in UD

<sup>4</sup>In UD, for defining real causatives in Armenian the relations `iobj:agent` and `aux:caus` were used.

the non-human dependent (*t'ey* in the example) is missed, there is a consensus to consider the human-denoting dependent (*vordown* in the example) as the direct object of the verb. But when *t'ey* is added, some authors stand by the opinion that the syntactic position *vordown* is an indirect object, and some continue considering it as a direct object. The verb *xmec'nel* (to make drink) is derived from the active verb *xmel* (to drink), so they assume it becomes ditransitive when becoming causative. Thus, both arguments in this sentence are thought to be direct objects - *t'ey* stands in accusative (as it is non-human), and *vordown* stands in dative (as it is human). One of the ways to determine the real direct object, is to transform the sentence into a passive construction and see whether the word in the syntactic position of the direct object can become the subject of the passive verb. But in this case there isn't any consensus about the human-denoting dependent to take this position, and in real causative constructions like this, other grammatical layers should be taken into account and the issue should be discussed in the frame of having or not a causative voice in Armenian.

Issues like these are often the case with traditional view, which confuses different levels of linguistic representation, namely, semantic roles and (syntactic) grammatical functions. A detailed description of verbs in the varieties of Armenian is necessary for the annotation of complements within the scope of core - non-core concept in the framework of dependency grammar. The precise scope of particular grammatical functions can be defined intra-linguistically (also considering the historical development of these phenomena in Armenian), on the basis of case marking, agreement and word order, as well as on the basis of diathesis alternations and actant derivation (Plungian, 2003). Traces of many issues go back to older stages of the language and the study of verb morphology and semantics in Classical and Middle Armenian can offer solutions and help to make a more precise set of tags and relations in annotation system of complements.

## References

Ashot Abrahamyan. 1962. *Bayë žamanakacic' hayerenowm [The Verb in Modern Aarmenian]*, volume A. NAS, Yerevan.

Sergey Abrahamyan, Varag Arakelyan, and Vagharshak Qosyan. 1975. *Hayoc' lezow, mas 2,*

*Šarahyowsowt'yown [Armenian Language. Part 2. Syntax]*. NAS, Yerevan.

Sergey Abrahamyan, Nvard Parnasyan, and Hmayak Ohanyan. 1974. *Žamanakacic' hayoc' lezow [Modern Aarmenian Language]*, volume 2. NAS, Yerevan.

Eduard Aghayan. 1967. *Žamanakacic' hayereni holovowmë ew xonarhowmë. karowc'vačk'ayin verlowçowt'yown [Inflection and conjunction in Modern Armenian. Structural Analysis]*. NAS, Yerevan.

Garō Arak'elean. 1967. *Arti hayereni k'eraganut'iwn [Grammar of Modern Armenian]*, volume B. Khalifé Printing Est., Beirut.

Manvel Asatryan. 1987. *Žamanakacic' hayoc' lezow, Šarahyowsowt'yown [Modern Armenian Language, Syntax]*. YSU, Yerevan.

David Gyulzatyan. 2015. Anc'oğakanowt'yan ew bayaseři k'erakanakan kargerı tipabanowt'yownë [typology of grammatical categories of the transitivness and the voice of the verb]. *Bulletin of Yerevan University*, 3(18):37–44.

Fritz Pasierbsky. 2003. Toward a classification of complements. In *Dependency and valency, An International Handbook of Contemporary Research*, volume 1, pages 803–814. Walter de Gruyter, Berlin.

Vladimir Plungian. 2003. *Obščaja morfologija: vvedenie v problematiku. [General morphology: introduction to the problematic]*. Èditorial URSS, Moscow.

Adam Przepiörkowski and Agnieszka Patejuk. 2018. Arguments and adjuncts in universal dependencies. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 3838–3852.

Armenak Yeghiaean. 2022. *Arewmdahayereni ughghakragan, ughghakhosagan, vojapanagan ughets'oyts' [A Manual for Western Armenian: An Orthographic, Orthoepic and Stylistic Guide]*. ARI literature foundation, Beirut - Lebanon.