



HAL
open science

Recurrent Pattern Modelling in a Corpus of Armenian Manuscript Colophons

Emmanuel van Elverdinghe

► **To cite this version:**

Emmanuel van Elverdinghe. Recurrent Pattern Modelling in a Corpus of Armenian Manuscript Colophons. 2016. hal-01283638v1

HAL Id: hal-01283638

<https://hal.science/hal-01283638v1>

Preprint submitted on 5 Mar 2016 (v1), last revised 22 Dec 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Recurrent Pattern Modelling in a Corpus of Armenian Manuscript Colophons

Emmanuel Van Elverdinghe

Research Fellow, F.R.S.-FNRS – Université catholique de Louvain, Belgium

emmanuel.vanelverdinghe@uclouvain.be

Abstract

Colophons of Armenian manuscripts are replete with yet untapped riches. Formulae are not the least among them: these recurrent stereotypical patterns conceal many clues as to the schools and networks of production and diffusion of books in Armenian communities. This paper proposes a methodology for exploiting these sources, as elaborated in the framework of a PhD research project about Armenian colophon formulae. Firstly, the reader is briefly introduced to the corpus of Armenian colophons and then, to the purposes of our project. In the third place, we describe our methodology, relying on lemmatization and modelling of patterns into automata. Fourthly and finally, the whole process is illustrated by a basic case study, the occasion of which is taken to outline the kind of results that can be achieved by combining this methodology with a philologico-historical approach to colophons.

keywords

colophon formula; formulaic patterns; Armenian colophons; automata; finite state transducers; Unitex; codicology; lemmatization

INTRODUCTION

This paper puts forward a methodology for searching, defining, and modelling formulae, *i.e.* recurrent patterns in a corpus of Armenian manuscript colophons. This methodology was devised and is being experimented in the context of a doctoral research focusing on formulae in Armenian manuscript colophons and the history of copying centres, under the supervision of Prof. B. Coulie (Université catholique de Louvain, Louvain-la-Neuve) and funded by the Belgian National Fund for Scientific Research (F.R.S.-FNRS). The language processing side of the research forms part of the GREgORI project, about which see the contribution of [Kindt, 2016] in this issue. Sincere thanks are due to Dr. B. Kindt, whose support and advice were very helpful in the writing of this paper.

Strictly speaking, a colophon is a record of completion found at the end of a book. In the Armenian context however, “colophon” encompasses notes of various kinds left by authors, handlers and users of manuscripts, in margins, at the end of the book, or in whatever blank space seemed fit—as described in [Stone, 1995]. This corresponds to the broader notion of յիշատակարան *yišatakaran*, as the colophon is known in Armenian, which translates literally to “memorial”.

These texts are a prime source for historical, codicological, linguistic, and prosopographical research, because they are contemporary records, telling not only about how the book was written, but also about its fate, its owners, historical circumstances of the time, and ultimately about the world view of the society that produced it. Armenian colophons are especially frequent and rich when compared to colophons in other languages, and therefore constitute a vast and fertile field of study.

I THE CORPUS

The corpus consists of the text of approximately 13,500 Armenian manuscript colophons, ranging from the 5th century to 1500 AD and from 1601 to 1660 AD, that appeared in scholarly editions (the 16th century was prepared for publication but never went to print). Their length varies from one single word to several thousands, while the whole corpus totals 1,194,632 words at the time of writing this article.

These texts had originally been published by Armenian scholars between 1950 and 1988, in nine instalments [Xaç'ikyan, 1950, 1955, 1968, and 1967; Hakobyan and Hovhannisyan, 1974 and 1978; Hakobyan, 1984; Mat'evosyan, 1984 and 1988], and were digitized between 1994 and 1997 under the leadership of the late Prof. J.J.S. Weitenberg (Universiteit Leiden), with the help of an European INTAS grant [INTAS-94-2974]. Today, the corpus is in care of the Université catholique de Louvain (Louvain-la-Neuve), where it has been improved and refined, notably by bringing it in line with present-day norms, such as Unicode and TEI encoding.

II OBJECTIVES

Several past studies have shown how a corpus-based, computer-assisted approach to Armenian colophons yields stimulating results in the fields of dialectology [Ĵahukyan, 1997], historical linguistics [Hovsep'yan, 1997], cultural history [Mat'evosyan, 1998] and social history [Weitenberg, 2005]. We intend to extend this kind of approach to manuscript studies, and more specifically, to the study of scribal schools and traditions, as well as of transmission networks of books and knowledge.

With respect to their form, redundancy is the most striking feature of colophons, primarily manifested in the use of stereotypical patterns, called formulae [Sirinian, 2014:76-85 and 90-95]. These formulae may take different forms: an elementary syntagm, for instance prepositional or completive, a proposition or sentence, or even a complete paragraph or colophon. The assessment of their lifespan, frequency, structural and lexical variation, and geographical diffusion, leads to results relevant for various aspects of manuscript studies, as predicted by [Sirinian, 2014:84-85 and 90-95] and already proven by [Reynhout, 2008] with regard to Latin colophons. It must be stressed that colophon formulae developed very recently into a distinct research area: this movement, initiated by the seminal contributions of [Huglo, 1954] and [Garitte, 1962] about individual formulae, came to fruition when [Reynhout, 2008] endeavoured to study formulae on a large scale, relying on a statistics-oriented methodology influenced by quantitative codicology (see also in [Reynhout, 2001]).

III METHODOLOGY

In the framework of this paper, we will focus on the methodology we have developed for detecting and modelling a formula, as well as retrieving all its occurrences. This involves automata (finite state transducers), serving two purposes. On the one hand, an automaton allows to easily and automatically set up a concordance of all attestations of the formula, a most valuable tool for all subsequent research; on the other hand, it constitutes an abstract representation of the formula, both graphical and convenient for analysing its structure.

Our methodology depends on the NLP-software Unitex (about its role in the project, see the contribution of [Kindt, 2016] in the present issue); graphs (automata) that will be presented are therefore compliant to the formalism used by Unitex (described in [Paumier]'s user manual). This methodological presentation is followed by a short case study, putting into practice the procedures described below.

3.1 Lemmatization

Lemmatization, at least partial, of the texts is a prerequisite to any successful research on colophon formulae. It is all the more important given that colophons exhibit a high grade of linguistic variation, due to multiple peculiarities. A first factor of variation stems from the time frame of the corpus, spanning more than ten centuries, during which the Armenian language underwent considerable evolution. Various dialectal features also show up. Finally, fidelity to the original text means that spelling mistakes and oddities are not normalized. In this context, research on formulae cannot rely on word-forms alone without exposing itself to overlook attestations containing non-standard forms.

3.2 Lexicological tools

Another requirement to investigate formulae lies with lexicological tools, as produced within Unitex, as well as using the software suite developed in the frame of the GREgORI project. The most useful among these for our research are concordances, whether they be lemmatized or based on word-forms. Since it juxtaposes identical or close segments of text, a concordance, particularly a lemmatized concordance, is most useful towards locating formulae and already analysing them sketchily.

3.3 Pattern recognition

The first step is to locate potentially formulaic patterns. Here, the researcher is before three complementary approaches: 1) computing word collocations, 2) manually browsing concordances of the corpus, or 3) surveying the extant literature. With a corpus of about 1,200,000 words, browsing the concordance reveals itself very cumbersome, while literature on the subject of colophon formulae is, for the most part, yet to be written. On the other hand, in order to get relevant outcomes from collocation computing, one must restrict the search by advance by narrowing the lookup window or excluding some words deemed insignificant, which more often than not occults interesting facts. Moreover, collocation computing, even when performed on a lemmatized corpus, basically ignores lexical variation and cannot, therefore, be deemed sufficient to satisfactorily identify a formula: the replacement of its elements with synonyms, antonyms, derivatives, etc., ought rather to be assessed by browsing the concordance. Using the three methods in conjunction thus provides the best results.

The sketch of the formula obtained at the end of this recognition process constitutes its “canonical” or “prototypical” form, corresponding to its most representative configuration. This representativity is defined as the best compromise between elementariness, frequency, and respect of the norm, and does not necessarily correspond to the earliest attestation. On this base, an abstract representation accounting for all particular occurrences can now be devised.

3.4 Modelling

Firstly, all constitutive elements of the formula have to be defined in terms of either lemma or grammatical category. It may also be necessary to summon flexional constraints in order to exclude irrelevant correspondences, and even to take the context into account.

Then begins the sequencing of the formula’s constitutive elements. This phase is meant to chart the path that will be followed by Unitex when searching the text with the automaton. For complex formulae, the difficulty lies in organizing the nodes efficiently to avoid redundancies and keep the graph readable and manageable. In such a situation, it is generally suitable to store recurring information in sub-graphs that are called from the main graph, *e.g.* subsets of lemmata, or elements susceptible of reduplication. Another common situation where resorting to sub-graphs is useful when dealing with words consisting of several concatenated lemmata: to process such sequences of lemmata, Unitex needs a specific concatenation marker to precede the first lemma, and this marker has to be accounted for in graphs. Prepositional prefixes

are a typical instance: being always preceded by a concatenation marker, it is more convenient to create a simple sub-graph with the sequence {(concatenation marker) + (preposition)}, which allows for better readability of the main graph, because only one node will be needed, linking to a separately stored sub-graph.

In most cases, and typically in the event of a markedly intricate, versatile, or frequent formula, the first attempt will not suffice to correctly model the pattern. Successful description of the formula can then only be achieved through progressive refining, helped by constant feedback from corpus observations.

3.5 Occurrence extraction and analysis

Once the formula has been modelled into a graph, the resulting automaton can be applied to the corpus. This enables extracting from it all sequences matching the formula's structure and displaying them in a concordance. In that way, the researcher is able to verify that the formula has been correctly plotted and, if it is indeed the case, he now has at his disposal an exhaustive listing of all its attestations. As stated earlier, it may be necessary to repeat the whole operation several times before achieving fully satisfactory results.

After this final stage, the philological and historical analysis begins by matching the references of data from the concordance with a database covering all metadata associated with the texts (manuscript call number, contents, date, author, place of origin, etc.); but this is a whole new chapter that exceeds the purposes of the present paper.

IV AN EXEMPLATIVE CASE STUDY

To exemplify this process, we now propose a brief case study illustrating the setting up of a formula graph as well as demonstrating its usefulness. For the purpose of this paper, we have picked a pretty simple and infrequent formula describing the mindset of the sponsor of a manuscript. Its canonical form is իմ կամաւ եւ յաւժարութեամբ *im kamaw ew yawzarut'eamb* "with my will and with readiness".

Queries on isolated forms reveal a wider range of possibilities within this formulaic frame: amongst others, the pattern may lack the possessive pronoun իմ *im* "my", or replace it with its plural equivalent մեր *mer* "our"; the substantives may appear in reverse order (իմ յաւժարութեամբ եւ կամաւ *yawzarut'eamb ew kamaw* "with readiness and will") or in the plural; modifier nouns may be introduced next to them, such as սրտի *srti* "heart" (genitive), or they may be determined by suffixed deictic particles. Once every possibility has been taken into account, the formula can be modelled by the graph below:

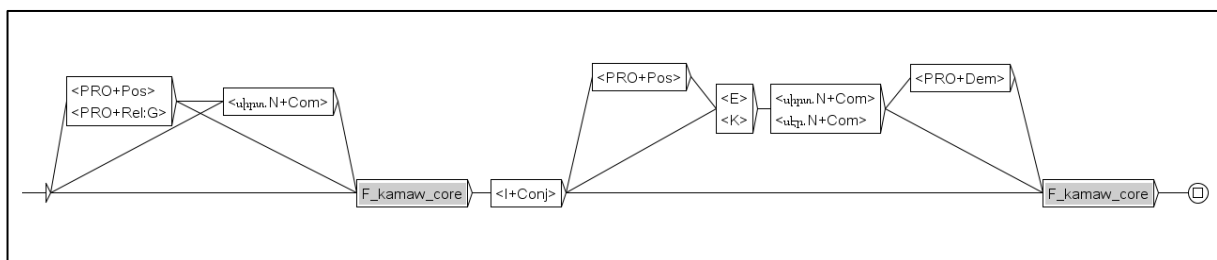


Figure 1. Formula graph for իմ կամաւ եւ յաւժարութեամբ *im kamaw ew yawzarut'eamb*.

Nodes in this graph mostly represent either lemmata or grammatical categories. Codes for these categories are pretty straightforward: PRO+Pos stands for possessive pronoun, PRO+Rel for relative pronoun, N+Com for common noun, and I+Conj for conjunction (see the GREgORI *Tagset* for a comprehensive listing of tags used for the processing of Greek—this tagset is provisional and not exhaustive for Armenian). As an inflectional code, G (in the first box from the left on fig. 1) restricts the matches to genitive forms. K (in the fourth box

from the right) is the concatenation marker, while E is the empty string. Lastly, information about the core substantives կամաւ *kamaw* and յաւժարութեամբ *yawzarut 'eamb* is contained in a sub-graph, called from the grey filled boxes. Because these substantives behave in the same way in both halves of the formula, it is more advisable to put the information in the following sub-graph that will be called twice, allowing for greater clarity, too.

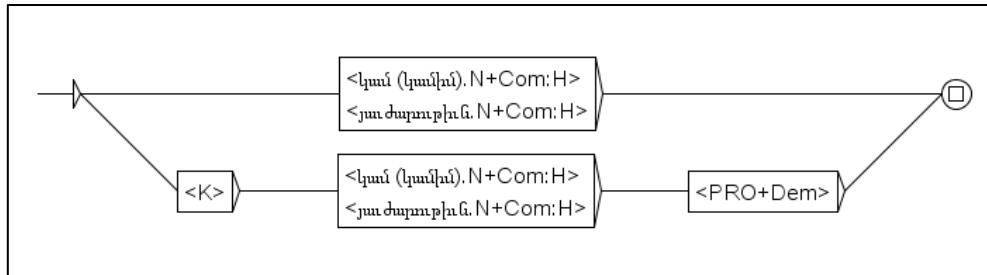


Figure 2. Sub-graph “F_kamaw_core”. The inflectional code H stands for the instrumental case.

Using this automaton as a search argument matches all occurrences of the formula, and nothing else. The search results are then ordered at the convenience of the researcher in concordances such as the one below, sorted on left context.

[XIII-624-3-22]	անգամ հոգոյ ակտակար եւ շահաւետ գործ:	<u>յաւժարութեամբ եւ կամաւ</u>	շարժելով եւ գործելով: Իսկ ոմն ճշմարիտ և իսկապէս
[XV_A-407-0-2]	(Եղեգեցի), վերադիտող այսմ նահանգի,	<u>իմ կամաւ եւ յաւժարութեամբ</u>	զիմ անետարանս, որ բերի յերուսաղէմս, ստացուած
[XIV-711-0-12]	զայս ի մայրաքաղաքս եզնկայ, եւ եղի	<u>իմ կամաւ եւ յոժարութեամբ</u>	յիշատակ ինձ, եւ այս վերոյգրեալ ննջեցելոցս, ի տաճարս
[XIV-737-0-22]	զայս ի մայրաքաղաքս երզնկայ, եւ եղի	<u>իմ կամաւ եւ սիրոյ յաւժարութեամբ</u>	յիշատակ ինձ եւ այս վերոյգրեալ ննջեցելոցս, ի տաճարս
[XVII_C-1146-4-14]	ամեն: Ես դարձեալ ծաներուք սիրելիք, զի	<u>իմ կամաւ եւ յոժարութեամբ</u>	արարի զսուրբ Ստեփաննոսի վախմ. մի՛ ոք իշխէ
[XIII-403-2-6]	առաքելոցն Պաւղոսի եւ Պետրոսի եւ Անդրէի,	<u>կամաւ եւ յաւժարութեամբ</u>	եւ մեծաւ հոգաբարձութեամբ իմոյ անսխտանութեանս՝
[XIII-287-1-44]	և զսուրբ Պաւլոսի՝ բուերորդի տեանն Գրիգորի,	<u>որոց կամաւ եւ յաւժարութեամբ</u>	յանկ ելեալ կատարեցաւ Պատճառ նամակս տաւնից,
[XV_A-25-0-1]	եւ Լիպիկ որդի Ասկանտարայ,	<u>իմ կամաւ եւ իմ սրտի յաւժարութեամբ.</u>	ծախեցի զսուրբ անետարանս Ասագտիլիս և առի
[XVII_C-678-0-10]	ի գրչաց եղիագար տրուպ ծառայ նորին,	<u>կամաւ եւ յաւժարութեամբ</u>	յանձն եւ կատարեցի զհրամայեալսն ի նմանէ ըստ իմում
[XVII_A-321-1-143]	եւ արդ ես՝ Պետրոսս և Եղբարս իմ՝ Աղազարն,	<u>մեր կամանս եւ մեր սրտի յոժարութեամբն</u>	ուսաբ զայս հրաշալի փոքրիկ սուրբ Անետարանս
[XVII_C-1399-1-164]	մարդիկը, վկայք լինիք առաջի Աստուծոյ, որ ես	<u>իմ կամօքս եւ իմ սրտիս յոժարութեամբն</u>	և իմ հիւսն ոգովս, զիմ գոյքս ամենայն ետու
[XVII_B-919-0-4]	մարն և կողակցին: Արդ՝ ես խաճառ Պալիս	<u>իմ կամաւ եւ յաւժարութեամբ</u>	առաջի Աստուծոյ շնորհեցի զսուրբ անետարանս
[XVII_B-1061-0-2]	և Նորա վկայութեամբն ես՝ Դշխունս	<u>իմ կամաւ եւ յոժարութեամբ</u>	իմ ճաշու գիրքն ծախեցի, տեր Մարտիրոսն Կանանց
[XVII_C-1435-0-1]	եւ՝ Պետրոս կաթողիկոս	<u>կամաւ եւ յոժարութեամբ</u>	զաւետարանս տուի Աշուղի որդի տեր Յովակիմին:
[XV_A-198-0-1]	մեք՝ նուաստ ոգի Սարգիս եպիսկոպոս,	<u>իմ սրտի կամաւ եւ յաւժարութեամբ</u>	տուաբ զգիրքս զայս ըստ հոգոյ որդոյն իմոյ և
[XV_C-224-0-4]	Ետզ զԱնետարանս: Ես՝ Սարգիս եպիսկոպոս	<u>իմ կամաւ եւ յաւժարութեամբ</u>	ծախեցի զսուրբ անետարանս Մարտիրոսայ
[XIV-400-0-31]	զիւստացեալս անուան իւրոյ, ամեն: Եւ որ	<u>սրտի կամաւ եւ յաւժարութեամբ</u>	արժանացնեն զմեզ միոյ տեր-որդոյնսի. Քրիստոս
[XIII-624-3-18]	կայեալ և ամենայն ջանից եւ աշխատութեանց	<u>յաւժարութեամբ եւ կամաւ</u>	զիւքեանս տուեալ սկիզբն առնեն հոգեւոր եւ փրկական
[XV_C-247-0-7]	յամենայն բարեաց Անգղեսս անարժան երիցու,	<u>կամաւ եւ յաւժարութեամբ</u>	խաճայ Ամիրտովաբ բժշկին, թանգի հմուտ գրով
[XV_A-189-0-6]	եւ՝ Յիսես և իմ բոյր Մարիամս, որ տվաք	<u>մեր սրտի կամանսն եւ յաւժարութեամբ</u>	զարծաթապատ անետարանս յիշատակ մեզ և
[XVII_B-629-3-2]	եպիսկոպոս և եղբայրն իմ՝ Ամիրխանս. մեք	<u>մեր կամաւ եւ յաւժարութեամբ</u>	մեր հոգոյս վախմ եւ յիշատակ որբաբ զհոբրիկ

Figure 3. Concordance, sorted on left context. The matching sequences are shown in underlined, blue font.

As a synthesis of the information assembled in the concordance, the table below lists every single configuration found in the corpus, along with its frequency and the year when it is first attested.

Frq.	1 st	Poss.	Modifier ₁	Substantive ₁	Conj.	Poss.	Modifier ₂	Substantive ₂
1	1267	<i>oroc</i> ՝ (Pro+REL:Gp)		<i>kamawk</i> ՝ (will:Hp)	<i>ew</i> (and)			<i>yawzarut 'eamb</i> (readiness:Hs)
2	1279			<i>kamawk</i> ՝	<i>ew</i>			<i>yawzarut 'eamb</i>
2	1296			<i>yawzarut 'eamb</i>	<i>ew</i>			<i>kamawk</i> ՝
1	1340		<i>srti</i> (heart:Gs)	<i>kamaw</i> (will:Hs)	<i>ew</i>			<i>yawzarut 'eamb</i>
1	1391	<i>im</i> (PRO+Pos1s)		<i>kamaw</i>	<i>ew</i>		<i>siroy</i> (love:Gs)	<i>yawzarut 'eamb</i>
1	1402	<i>im</i>		<i>kamaw</i>	<i>ew</i>	<i>im</i>	<i>srti</i>	<i>yawzarut 'eamb</i>
1	1415	<i>mer</i> (PRO+Pos1p)	<i>srti</i>	<i>kamawk 'n</i>	<i>ew</i>			<i>yawzarut 'eamb</i>
1	1416	<i>im</i>	<i>srti</i>	<i>kamaw</i>	<i>ew</i>			<i>yawzarut 'eamb</i>

5	1428	<i>im</i>	<i>kamaw</i>	<i>ew</i>			<i>yawžarut'eamb</i>
1	1607	<i>mer</i>	<i>kamawn</i>	<i>ew</i>	<i>mer</i>	<i>srti</i>	<i>yawžarut'eambn</i>
1	1631	<i>mer</i>	<i>kamaw</i>	<i>ew</i>			<i>yawžarut'eamb</i>
2	1651		<i>kamaw</i>	<i>ew</i>			<i>yawžarut'eamb</i>
1	1660	<i>im</i>	<i>kamawk's</i>	<i>ew</i>	<i>im</i>	<i>srtis</i>	<i>yawžarut'eambn</i>

Table 1. Formula types, with frequency and date of first attestation. N.B. As a reminder, the 16th century is not covered in the corpus.

Armed with the graph and with concordances, the researcher is now able to start analysing the formula, beginning with structural analysis. Subsequently, metadata about each colophon are retrieved from the database and combined to the text and context of the formula, allowing undertaking a historical analysis. Although this part of our methodology falls beyond the scope of the present paper, we would like to bestow some space to the results that it leads to, in order to demonstrate the usefulness of graph modelling towards achieving them.

Historical and structural analysis of the occurrences reveals the classical, three-level development scenario of many formulae: 1) experimentation; 2) settlement; 3) dissemination. The first phase is represented by two manuscripts dated from 1267 and 1279, respectively. They were produced in different regions and use the formula in a rather rudimentary form, without the personal pronoun իմ *im* “my” but with կամաւք *kamawk* “with will” in the plural—in conformity with the Classical Armenian practice. At this time, the formula is probably still nothing more than an isolated, spontaneous combination of words.

In the settlement phase, this combination will become standardized by way of its repeated usage in a definite environment. For the formula at stake in this paper, this happens in Erzinka (Երզնկա), a city of Upper Armenia (mod. Erzincan, Turkey) between the years 1296 and 1391, as evidenced by three different colophons (one of them even using the formula twice).

The third and last period of our formula’s development starts in the 15th century, when it spreads outside of Erzinka. However, most early fifteenth-century attestations come from un-localized deeds of sale or gift, recorded as colophons in manuscripts of various origins. It may thus not be said with certainty whether they were connected in some way to Erzinka or not. Anyway, the formula did not enjoy considerable success, with only six occurrences during the whole 15th century. In the 17th century, it followed the same path, with eight occurrences between 1601 and 1660, which is almost proportional to the increase in preserved manuscripts between the 15th and the 17th century as noted in [Kouymjian, 1983: 433].

The formula we have discussed here is thus a rather anecdotal one. With more interesting patterns, it is possible, in addition to what has been described above, to map a transmission network and to go into far greater detail, even though the fact that many manuscripts have disappeared in the course of time may hamper positive conclusions.

CONCLUSION

In the frame of our research project on formulae in Armenian colophons, the traditional method of historical, critical, and philological analysis, assisted on occasion by statistical calculations, is combined with modern techniques of computer-aided processing, namely, lemmatization, production of concordances, and designing of automata.

Such an integrated approach to colophon formulae leads to interesting and innovative results. The most conspicuous impact of this research concerns traditions and networks of circulation of books and the people who produced or handled them. In certain cases, such a study also allows to put forward more or less definite opinions as to the provenance or authorship of some manuscript otherwise unattributed. On a larger scale, we get an original insight into the life and activity of copyists, spanning several centuries of text and knowledge transmission.

The input of pattern modeling through automata, assisted by other lexicological tools, must not be neglected: not only do graphs make researching formulae more convenient, more effi-

cient and more rigorous, but they also have a role of their own in assessing and communicating the configuration of the formula. For example, the existence of different types of a given formula is often reflected in structural variations; therefore, part of the subsequent analysis also relies upon its graph.

References

- Garitte G. Sur une formule des colophons de manuscrits grecs (ἡ μὲν χεῖρ ἢ γράψασα). *Collectanea Vaticana in honorem Anselmi M. card. Albareda a Bibliotheca Apostolica edita*, vol. 1. Studi e Testi, 219. Biblioteca Apostolica Vaticana (Vatican City), 1962:359-390.
- GREgORI *Tagset* https://alfresco.uclouvain.be/alfresco/service/guest/streamDownload/workspace/SpacesStore/57e2c121-ff0f-41f5-a2f7-3515fb33a3e6/GREgORI_tagset.pdf?guest=true
- Hakobyan V. *Hayeren jeṛagreri ŽĒ dari hišatarakanner*, t. 3: 1641—1660 t't'. Nyut'er hay žolovrdi patmut'yan, 17. Haykakan SSH GA Hratarakč'ut'yun (Yerevan), 1984.
- Hakobyan V. and Hovhannisyanyan A. *Hayeren jeṛagreri ŽĒ dari hišatarakanner*, t. 1: 1601—1620 t't'. Nyut'er hay žolovrdi patmut'yan, 14. Haykakan SSH Gitut'yunneri Akademiayi Hratarakč'ut'yun (Yerevan), 1974.
- Hakobyan V. and Hovhannisyanyan A. *Hayeren jeṛagreri ŽĒ dari hišatarakanner*, t. 2: 1621—1640 t't'. Nyut'er hay žolovrdi patmut'yan, 15. Haykakan SSH Gitut'yunneri Akademiayi Hratarakč'ut'yun (Yerevan), 1978.
- Hovsep'yan L.S. *ŽG dari hayeren jeṛagreri hišatarakanneri lezun*. Accessing Armenian Colophons, 2. Van Aryan (Yerevan), 1997.
- Huglo M. Christe fave votis. *Scriptorium*. 1954;8:108-111.
- INTAS-94-2974 http://cordis.europa.eu/project/rcn/32718_en.html
- Ĵahukyan G.B. *Barbaṛayin erevuyt'ner haykakan hišatarakannerum*. Accessing Armenian Colophons, 1. Van Aryan (Yerevan), 1997.
- Kindt B. Tools and data of the GREgORI project for the study of Ancient Greek and the languages of the Christian East. To be published in *Journal of Data Mining and Digital Humanities*. 2016;3.
- Kouymjian D. Dated Armenian Manuscripts as a Statistical Tool for Armenian History. Samuelian T.J. and Stone M.E. (eds.) *Medieval Armenian Culture: Proceedings of the Third Dr. H. Markarian Conference on Armenian Culture*. University of Pennsylvania Armenian Texts and Studies, 6. Scholars Press (Chico, CA), 1983: 425-438.
- Mat'evosyan A.S. *Hayeren jeṛagreri hišatarakanner: ŽG dar*. Nyut'er hay žolovrdi patmut'yan, 20. Haykakan SSH Gitut'yunneri Akademiayi Hratarakč'ut'yun (Yerevan), 1984.
- Mat'evosyan A.S. *Hayeren jeṛagreri hišatarakanner: E—ŽB dd*. Nyut'er hay žolovrdi patmut'yan, 21. Haykakan SSH GA Hratarakč'ut'yun (Yerevan), 1988.
- Mat'evosyan A. *Hayeren jeṛagreri hišatarakannerə miṅnadaryan hay mšakuyt'i usumnasirut'yan skzbnalbyur*. Accessing Armenian Colophons, 3. Van Aryan (Yerevan), 1998.
- Paumier S. *Unitex 3.0. User Manual* (<http://igm.univ-mlv.fr/~unitex/UnitexManual3.0.pdf>).
- Reynhout L. Codicologie quantitative et paradigmes scientifiques: Une typologie des formules latines des colophons de manuscrits occidentaux. *Gazette du livre médiéval*. 2001;39:1-9.
- Reynhout L. *Formules latines de colophons*, 2 vol. Bibliologia, 25 A-B. Brepols (Turnhout), 2008.
- Sirinian A. On the Historical and Literary Value of the Colophons in Armenian Manuscripts. Calzolari V., with the collaboration of Stone M.E. (eds.). *Armenian Philology in the Modern Era: from Manuscript to Digital Text*. Handbuch der Orientalistik, VIII, 23/1. Brill (Leyden and Boston), 2014:65-100.
- Stone M.E. Colophons in Armenian Manuscripts. Condello E. and De Gregorio G. (eds.) *Scribi e colofoni: Le sottoscrizioni di copisti dalle origini all'avvento della stampa: Atti del seminario di Erice, X Colloquio del Comité international de paléographie latine (23-28 octobre 1993)*. Biblioteca del «Centro per il collegamento degli studi medievali e umanistici in Umbria», 14. Centro italiano di studi sull'alto medioevo (Spoleto), 1995:463-471.
- Weitenberg J.J.S. Cultural Interaction in the Middle East as Reflected in the Anthroponomy of Armenian 12th – 14th Century Colophons. van Ginkel J.J., Murre - van den Berg H.L. and van Lint Th.M. (eds.). *Redefining Christian Identity. Cultural interaction in the Middle East since the Rise of Islam*. Orientalia Lovaniensia Analecta, 134. Peeters and Departement Oosterse Studies (Leuven, Paris and Dudley, MA), 2005:261-273.
- Xaç'ikyan L.S. *ŽD dari hayeren jeṛagreri hišatarakanner* (Nyut'er hay žolovrdi patmut'yan, 2). Haykakan SSR Gitut'yunneri Akademiayi Hratarakč'ut'yun (Yerevan), 1950.
- Xaç'ikyan L.S. *ŽE dari hayeren jeṛagreri hišatarakanner*, t. 1: 1401—1450 t't'. Nyut'er hay žolovrdi patmut'yan, 6. Haykakan SSR Gitut'yunneri Akademiayi Hratarakč'ut'yun (Yerevan), 1955.
- Xaç'ikyan L.S. *ŽE dari hayeren jeṛagreri hišatarakanner*, t. 2: 1451—1480 t't'. Nyut'er hay žolovrdi patmut'yan, 8. Haykakan SSR Gitut'yunneri Akademiayi Hratarakč'ut'yun (Yerevan), 1958.
- Xaç'ikyan L.S. *ŽE dari hayeren jeṛagreri hišatarakanner*, t. 3: 1481—1500 t't'. Haykakan SSH Gitut'yunneri Akademiayi Hratarakč'ut'yun (Yerevan), 1967.