

## OTTOMAN CHRONICLER MUSTAFA NAIMA'S INDIVIDUAL STYLE IN CORPUS DIMENSION

Oles KULCHYNSKY\*

### ÖZ

“Tarih-i Naima” adlı Osmanlı düzyazı eseri artık uzun süre içinde üslubunun özellikleri sayesinde edebi eleřtiricilerin, tarihçilerin ve dilbilimcilerin ilgilerini geniş bir ölçüde çekmektedir. Kroniğin frekans leksik parametrelerinin diđer Osmanlı metinleriyle korpus yöntemleriyle genel karşılaştırılması, onun Osmanlı Türkçesi düzyazı kanunları ve normuna göre yaratılmış olmasını ispatlamaktadır. Ancak aynı zamanda “Tarih-i Naima” diđer Osmanlı Türkçesi metinleriyle belirsiz, ifadesiz bir şekilde deęişiklik göstermektedir. Bu deęişiklik ise Mustafa Naima'nın tarihine ünlü Evliya Çelebi'nin ‘Seyahatname’ adlı eserinin aksine Osmanlı Edebiyatında iki yüzyıl boyunca unutulmamış ve yeterince meşhur bile olmaya izin vermiştir.

**Anahtar Kelimeler:** Osmanlı Kroniği, Osmanlı Türkçesi Düzyazı, Mustafa Naima, Bireysel Üslup, Korpus Dilbilimi.

### ABSTRACT

An Ottoman prose work “Tarih-i Naima” (“Naima’s History”) for a long time has been the subject of special interest for literary critics, historians and linguists due to its stylistic features. The general comparison of the chronicle frequency lexical parameters with other Ottoman prose texts by corpus approach demonstrated that it had been created according to the canons of Ottoman Prose Literature and rules of its language norm. However, at the same time, “Tarih-i Naima” differs from other texts in some obscure, inexpressive form, which let the annalist work stay quite famous in Ottoman Literature for two centuries and not be forgotten as was the chronicle “Seyahatname” of the famous traveler Evliya Chelebi.

**Keywords:** Ottoman Chronic, Ottoman Prose, Mustafa Naima, Individual Style, Corpus Linguistic.

Corpus approach enables an exact idiolect analysis due to its broad comparison with other texts in synchronistic and diachronic perspective. The comparison is usually based on the data of texts frequency vocabularies and syntax. At the same time, the correlation analysis between frequency vo-

---

\* Dr., PhD, V. I. Vernadskyi Tavriya National University, Department of Foreign Languages and Literatures, Kyiv, Ukrain, ORCID: 0000-0003-4761-3843, kulchynsky@gmail.com

cabularies rank parameters can be used for the sample validation while compiling electronic documents collection (Prohorov 2003). Such collection (corpus) in total of 21 Ottoman prose texts was created for our research on idiostyle of the 18th century Turkish chronicle "Tarih-i Naima" ("Naima's History"), or in the full name: "The Garden of Husayn in the Summary of the Chronicles of East and West" (Ravdatü 'l-Hüseyn fi hulâsat-i aḥbâri 'l-ḥâfiqayn), written by the first vakânuvis (state chronicler) of Ottoman Empire Mustafa Naima (1655-1716). While researching his idiostyle, we proceeded from the linguistic view of this phenomenon as deviation from language norm and tried to fix it in empirical statistical way.

In our case, the calculation of the Pierson correlation coefficient (R) between rank parameters of electronic documents vocabularies verified the sampling adequacy for created corpus. However, the analogical computation between frequency vocabularies of researched text and documents collection demonstrated lexical deviation of "Tarih-i Naima" from other Ottoman chronicles. Meanwhile lexicon of selected documents in the view of their common frequency dictionary and according to high lexical closeness between them can be considered as hypothetical language norm of Ottoman chronicles.

For the compilation of Ottoman chronicles frequency dictionary as a model of Ottoman Prose language norm we initially sampled 22 chronicles electronic texts applying logical criteria. Most of them were transliterated in Latin letters by graduating students of Turkish Universities in their PhD and Master Theses<sup>1</sup>. Taking into account the recommendations on text typology and the metatext markup, proposed by Jon Sinclair (Sinclair 2004), 15 sampled chronicles could be classified in such way:

1) All texts are manuscripts created by Arabic letters of Ottoman Turkish alphabet and transliterated in Latin letters.

2) Text domain is Ottoman chronicles functional style (tarih) and its sub-styles, such as sefaretname, fetihname and sefername.

3) Language of corpus is the Ottoman Turkish of its Classical Period (16th-19th centuries).

4) Texts location is the former territory of Ottoman Empire, mostly – Anatolian peninsula, but two manuscripts are from Cairo and one – from Crimea.

5) Chronicles belong to the period from 1608 until 1760 y. ± 2 years, the time, in which Ottoman Turkish almost did not change, during its Classical Epoch that lasted from the beginning of XVI c. until the middle of XIX c. (Timurtaş 1997: 1).

As the result Ottoman chronicles corpus (OCC) was created on the base of the samples, taken from the next prose works<sup>2</sup>:

---

<sup>1</sup> The research was possible only due to these works; all of them are given in references.

<sup>2</sup> The transliteration given in the texts, which were used for creation of the korpus, is preserved in the list below though it does not coincide with different rules of transliteration

1. "Gazâvât-ı Murad Paşa" by Vâsîfî (XVII c.) (Göknur 2006).
2. "Tevârîh-i Âl-i Osman" by unknown author (XVII c.) (Çimen 2006).
3. "Peçevi Tarihi" (pages 80b-114a) by Ibrahim Peçevi (1572-1650) (Özbal 2005).
4. "Peçevi Tarihi" (pages 46b-80a) by Ibrahim Pechevi (1572-1650) (Gürışık 2005).
5. "Târihçe-i Fethi Revan ve Bagdad" finished by Kara Çelebi-Zâde Abdülaziz Efendi (1591(1592)-1658) (Yıldırım 2005).
6. "Târih-i Kamanîçe" by Hacı Ali Efendi (XVII c.) (Hacı 2007).
7. "Târih-i Mehmed Giray" by Mehmed Giray (XVII-XVIII c.) (Demir 2006).
8. "Târih-i Sülâle-i Köprülü" by Behcetî Seyyid İbrahim Efendi (XVII c. – 30-40 y.y. of XVIII c.) (Gökçek 2006).
9. "Çelebi-zâde Âsim Târîhi" by Çelebizâde Âsim Efendi (1685-1760 y.y.) (Aktaş 2008).
10. "Târih-i Sami" by Mustafa Sami (1680 y.y.-1734) (Karadayı 2008).
11. "Kıt'a-min-Tarih-i Sultan Mahmûd-ı Evvel" by Musaffa Mehmed Efendi (XVIII c.) (Çoruhlu 2005).
12. "Keyfiyet-i Rusiyye" by Hasan Kürdi (17 c.- 18 c.) (Tübençokrak 2007).
13. "Risale-i Teberdariyye Fi Ahval-i Darü's-Sa'ade" by Derviş Abdullah (18 c.) (Saka 2007).
14. "Sefâretnâme" by Nişli Mehmed Ağa (XVII ct. -1732 y.) (Mertayak 2005).
15. "Sefâretnâme" by Şehdi Osmanlı Efendi (1707-1769) (Polatçı 2003).

At the same time 7 more works were experimentally added to the corpus as far as they extended its timetable and sub-style variety. The experiment was performed with the aim of more broad text research in diachronic linguistics and its comparison with other sub-styles works.

The experimental part of the corpus can be represented in such chronological order:

16. "Tevârîh-i Âl-i Osmân" by Nişâncı Mehmet Paşa (? - 1571) (Yastı 2005).
17. "Gence Fetihnâmesi" by Rahîmîzâde İbrahim Çavuş (XVI c.) (Dündar 2006).
18. "Tebriziyye" by Talîkîzade Mehmed Subhî (?-1606) (Özkuzugüdenli 2005).
19. "Mukaddimetü's Sefer" by unknown author (18 c.) (Söylemez 2007).
20. "Tarih-i Sefer-i Rusya" by Said B. Halil İbrahim (1750 (?)-1820) (Altun 2006).

---

used in Turkey during the XX and the beginning of the XXI century for the Ottoman Turkish texts.

21. "Pehlivan İbrahim Paşa Vakayi'-nâmesi" by Abdüllatif Efendi (1776-1832) (Erol 2007).

22. "Tevarih-ı Tohta Bay qavli üzere ibtida Qırım ceziresine malik olan Hanları beyan ider", fragment of the «Seyahatname» by famous traveler Evliya Çelebi (17 c.) (Seytyagyayev 2002).

In the total amount, all 22 texts number almost 1 million word forms, which are divided among them almost in equal proportions. After the sampling by logical criteria, stemming and lemmatization the degree of the strength of a linear relationship (R) was calculated between frequency dictionaries of all chronicles. In quantitative linguistic various statistic correlations measures can be used as similarity measures between texts (Prohorov 2003). However, the Pearson correlation coefficient is one of the strongest among them and is often used in computational linguistic and corpus-based researches for this purpose (Salton 1989, Dinu & Popescu 2009, Prohorov 2003, Huang 2008, Mihalcea 2006, Guadalupe 2008). H.D. White and K.W. McCain especially put it to use in information since as predominant measure of similarity between author pairs in author co-citation analysis (White 2003; McCain), what is close to our field of study according to the Ottoman tradition to rewrite chronics and co-cite authors. Yet in modern stylometry method of Pearson correlation is often used for finding relationships between most frequent functional words or most frequent 50 words of the texts as their style markers (Dinu & Popescu 2009). The approach is based on ideas of van Halteren about the existing of human "stylome", a set of measurable characteristics of language products. Halteren defines stylome as "extensive enough to be able to distinguish between pairs of language users on the basis of their language use". So the set of language use quantitative characteristics –stylistic lexical and syntactic– form the human stylome (Dinu & Popescu 2009).

Taking into account such approaches, we used Pearson correlation not only as simply lexical but also stylistic similarity measure between whole texts of our OCC. Meanwhile in such approach we considered "Bradford's Law" which confirms that some "core texts" exist for each subject area, which is Ottoman prose in our research. Respectively creation of corpus as conjunction of "core" or relevant texts could be based on the method (Prohorov 2003) of using some statistical fitting criteria for comparison quantitative specifications of texts, which are their frequency vocabularies. Main idea that underlies this approach is the hypotheses about coincidence of words frequency functions in the texts, which belongs to one subject area. In other words, empirical functions of word distributions by their frequencies are close to each other. For comparison of ordered sequences of word frequency functions values of texts T1 and T2 can be used different correlation coefficients (Prohorov 2003).

Thus, we use Pierson correlation coefficient (R) as the method of texts similarity measurement in this research and it was calculated by the formula:

$$r_{XY} = \frac{\text{cov}_{XY}}{\sigma_X \sigma_Y} = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2 \sum(Y - \bar{Y})^2}}$$

The degree of closeness is represented in numbers from 0 to +1. The lower index R stands for the minor relationships between texts lexicon are and vice versa. Output data can be submitted in the form of correlation matrix as in the Table 1. (Ek Tablo: 1), where numbers in the first row and column of the table correspond with the order of chronicles names as they were introduced above.

The correlation degree (R) among frequency dictionaries of 21 out of 22 documents compounds a set of numbers from +0,709 to +0,981 ( $p < 0.001$ ). It validates our samples for creation of Ottoman chronicles language norm model, except one text that is the segment of the "Seyahatname", written by Evliya Çelebi. This text (№ 22) correlates with other chronicles on the level: from +0.663 to +0.734. Moreover, that confirms the thesis of many turkologists about especial stylistic uniqueness of that work.

As the result, the computation of correlation matrix made it possible to compile frequency dictionary of Ottoman chronicles, which approximately consists of 32,000-38,000 words considering the inaccuracy probability. Hypothetically, this dictionary represents the model of Ottoman prose language norm at its lexical level. In addition, the first 10 thousand lemma cover 99.39% of the whole hypertext, which after the separation of the segment of Evliya Çelebi's "Seyahatname" numbers 989,682 word forms.

Correlation results previously demonstrate:

1) The Ottoman Turkish prose genres have not been changed substantially at the lexical language level from the beginning of 16th century till 19th century.

2) The identical strict rules of prose creation and literary canon, which reveal themselves in monotonous lexical operations, equally functioned in all Ottoman prose sub-genres despite the variety of themes and plots.

Then analogical corpus and statistical operation were conducted with the researched text "Tarih-i Naima". Text samples were divided into three categories according to logical criteria based on extra linguistic information about the chronicle.

As Thomas Lewis convincingly showed it in his work "A Study of Naima", the largest part of "The Garden of Husain in the Summary of the Chronicles of East and West" rewritten from other chronicles, while the annalist himself created only several scores of pages (Lewis 1972: 130-131). Therefore, pages samples were divided into two, but in fact – three groups:

1) Chronicle pages, which were created by author himself and could be divided into two parts: a) prefaces for two volumes of history, both written at the beginning of 18th century, the first one was dedicated to sadrazam (grand vezir) Köprülü Amcazade Hacı Hüseyin Paşa, and the second –

to sadrazam Moralı Damat Hasan Paşa; b) segments of the chronicle itself composed by Mustafa Naima.

2) Segments of largest part of chronicle, which Mustafa Naima re-wrote from other annals.

The modern transliterated edition of "Naima's History" in six volumes numbers 2,247 pages (Naima 2007). Comparing it with edition of 1864-1866 years, published in Istanbul printing house Amire Press in the same six volumes (Naima 1866) we selected the majority of text segments compounded by the author himself. Their list is published in Tomas Lewis's work «A Study of Naima» (Lewis 1972: 130-131).

Therefore, another corpus for "Naima's History" (NHC) was created. The segments of researched text were distributed in such proportion: 1) 12,219 word forms that belong to prefaces of Naima's work; 2) 16,378 word forms, the largest part of those fragments which were compounded by Mustafa Naima and belong to the chronicle itself; 3) 69,334 word forms, or every tenth page of that part of the chronicle, which was rewritten by author from other annals.

In this case, Pierson coefficient was calculated among the frequency dictionaries of mentioned three sample groups, but afterward – among them and frequency dictionary of 21 Ottoman chronicles. At the end of the operation vocabularies of three sample groups were united into one, the R coefficient was calculated between it, and frequency dictionary of Ottoman chronicles corpus.

Output can be represented in the correlation matrix form as in Table 2 (Ek Tablo: 2). Signs of the table mean: T1 – text segment rewritten by the author; T2 – segment compounded by the author in the chronicle itself so it must express his idiostyle features as much as it's possible; T3 – chronicle prefaces, which author wrote by himself, yet following some officialese so far as they are dedicated to the rulers of Ottoman Empire; ALLT – three segments together; Corpus – texts of above represented 21 Ottoman chronicles electronic collection.

As Table 2 shows, different segments of "Naima's History" are at the very high level of lexical similarity. The R coefficient degree between them is so high that we can conclude: pages of chronicle written by Mustafa Naima himself do not differ by their vocabulary from the rest of the text rewritten from other sources.

Meanwhile, the correlation coefficient  $R + 0,771$  ( $p < 0.001$ ) among the frequency dictionaries of the whole "Naima's History" and 21 Ottoman chronicles collection can demonstrate that author's individual style exists in obscure features which radically don't contradict the literary canon and language norm of 16th-18th century Ottoman prose. As it was shown in correlation matrix of Table 1, such language norm can be characterized as a numbers set from  $+0.709$  to  $+0.981$  ( $p < 0.001$ ), the shot interval which shows high lexical similarity of sampled prose segments despite of their chronology and genre variety. In such lexical similarity measure coordi-

nates "Naima's History" place ( $R + 0.771$  ( $p < 0.001$ )) is somewhere at their periphery but not beyond the scope.

As far as the high degree of closeness characterizes all three segments of NHC, we can conclude that obscure individual style features exist in the whole Naima's chronicle. In other words, chronicle segments created by Mustafa Naima himself do not differ from the rest of rewritten text. However, it can mean that the whole chronicle is not simply rewriting, but the author's vision of used historical sources represented in his inexpressible individual style which not cardinally but distinguishes the chronicle from the rest of Ottoman prose. Can such vision influence the meaning of text, especially when we analyze important historical source like "Naima's History"? The question is up to historians.

Anyway, according to the data represented in correlation matrix of Table 2 we can admit that author's individual style reveals itself at the lexical level of language as a low deviation from the Ottoman prose language norm. The author implicitly challenged literary canon at the same time trying not to go beyond its bounds, using own tools for text creating. Such thesis is applied concerning statistical comparison of "Naima's History" not only with texts of his contemporaries but also in diachronic linguistics throughout the whole Classical Period of Ottoman Turkish Language.

The calculation of Pierson coefficient between "Naima's History" vocabulary and dictionary of an Ottoman prose at least demonstrates lack of high lexical-morphemic similarity between them and high probability of Naima's chronicle individual style features existence. In order to prove the thesis we tried to fix deviation of the chronicle from language norm, finding those exact lexemes, morphemes and some syntactic peculiarities that define "Naima's History" as literary work.

The author vocabulary (9,500 lemma in total sum) has its own parameters of text covering and they quite differ from OCC parameters. If the first 225 lemma of frequency dictionary in "Naima's History" cover 51,66% of the text then in our collection of documents analogical coefficient is 49.97%. If the most frequent 1,000 lemma in "Naima's History" cover 73.8% of the whole text, in the rest of Ottoman prose they cover only 71%. In addition, such disparity grows up until the 98.74% of text covering for first 9,000 lemma in researched text and 94.5% for 21 Ottoman document. It means that the author's vocabulary is much simpler in comparison with other works of analogical period or even does not correspond to the demands of Ottoman prose canon.

Above-mentioned individual style peculiarities exist on different text levels. On the lexical level, they obviously occur within the main chronicles concepts, one of which is the figure of Ottoman sultan or even sovereign as an abstract notion. Usually this concept is designated by words: *hümâyün* (august, imperial), *pâdişah* (padishah), *sultan*, *han*. They form the core of lexical-semantic field, which include less frequent words with synonymous sense. For example, these are *cîhângîr* (conqueror of the world), *cîhân-*

penâh (holder of the world), cîhânşâh (shah of the world), hâkan (a sovereign), halife (Caliph), hânedan (dynasty), melîk (sultan), melîkane (royal), pâdişahane (imperial), pâdişahî (the imperial dignity or function) etc.

In our Ottoman chronicles corpus and "Naima's History" the core of lexical-semantic field that characterize the notion of sovereign are almost equal. Such words as hümeyun, padişah, sultan, han cover about 1 % of texts. In total sum, these lexemes have a frequency 9,900 per million words in "Naima's History" and 10,482 in selected Ottoman prose works. However, many other synonyms of them like mentioned above cîhândar, cîhânpenah, cîhânşâh etc. has much less frequency in researched text as compared with context or even simply absent in it.

In addition, other parallel tendency has place in "Naima's History". Lexemes, which denote different classes and titles of Ottoman Empire are used in it two times more frequently than in OCC. They cover 3% of text while the same coefficient for the rest of Ottoman prose is only 1.5%. The most frequent of such terms are: vezîr (vizier), asker (soldier), vâlî (governor-general), yeniçeri (janissary), beylerbeyi (governor), şeyh (sheikh), kâdî (qadi), kethüdâ (kethuda), ulemâ (ulama), çavuş (chavush), sipâhi (sipahi), kapıcı (doorkeeper), bosstancı (bosstanji), me'mûr (civil servant), müftî (mufti), silâhdar (silahdar). Two tendencies, domination of different titles and lack of glorification of Ottoman sovereign in "Naima's History", as contrasted with other Ottoman texts in fact mean that deep stylistic code of the text opposites the rules of literary canon. We observe not only the lack of words that denote sultan or senses connected with its figure but also epithets and metaphors dedicated to its person. However, some other lexicon scheme connected with new heroes of chronicle in "Naima's History" appears.

For revealing other lexical peculiarities of researched text, we used popular TFIDF method of searching keywords (Ramos 2003). According to P. Guiraud, keywords are words whose frequency of occurrence in text is usually much higher than the frequency of occurrence of a content word (Guiraud 1953: 155). But as far as we got the data about lexical dissimilarity of "Naima's History" and 21 Ottoman chronicle (hypertext) we used TFIDF method for seeking keywords in OCC as well to know which of them as typical to Ottoman prose are less represented in "Naima's History". Keywords were calculated by the formula:

$$tf(t, d) = 0,5 + \frac{0,5 \times f(t, d)}{\max \{f(w, d) : w \in d\}}$$

$$idf(t, D) = \log \frac{|D|}{|\{d \in D : t \in d\}|}$$



- $|D|$ : cardinality of  $D$ , or the total number of documents in the corpus
- $|\{d \in D : t \in d\}|$  : number of documents where the term  $t$  appears
- $Tf(t, d) \neq 0$ : If the term is not in the corpus, this will lead to a division-by-zero. It is therefore common to adjust the formula to  $1+|\{d \in D : t \in d\}|$ .

As the result, we got 269 keywords. For validating the difference of keyword frequencies in collection of Ottoman chronicles and Naima's History we used Pearson's chi-square test ( $\chi^2$ ), which is common for such researches. Results could be demonstrated in the form of the Table 3 (Ek Tablo: 3), where frequencies are given per million (fpm). And if  $P \chi^2$  is  $<0,5$  it means for linguistic that probability values are assumed to be significant, so we can therefore say with quite a high degree of certainty that the difference between frequencies is a true reflection of variations in text and hypertext and is no due to chance (MacEnery 2001: 85).

Afterwards keywords of OCC and NHC were divided into three groups:

1) 71 keywords of Naima's chronicle, whose *frequency of occurrence* in his text is obviously much *higher* than in 21 Ottoman chronicles hypertext ( $P<0,05$ ;  $P<0,01$ ;  $P<0,001$ ). These are âdetli (customary), altmış (sixty), ancak (only), asâkir (soldiers), batac (morass), baş (head, chief), belki (perhaps), biri (one, person), bizzât (personally), boyun (neck), bugün (today), değil (not), deli (mad), delil (proof), dere (stream), derviş (dervish), duvar (wall), fesad (disorder), gece (night), geçit (pass), gizlice (secretly), gulüvv (assault), güç (strength), güft (saying), gyureş (wrestling), halk (folk), hamle (attack), herif (man, fellow), hiç (none), husûsı (particular), hüccet (argument), ihânet (treachery), ihrâc (exportation), ihtilâl (insurrection), iş (job, occupation), karpuz (watermelon), katil (murderer), kazak (cossack), kendi (self), keyse (purse), kırmızı (red), kimse (person), konak (mansion), kulluk (slavery), kurb (nearness), musâhib (interlocutor), mütemekkin (established in any place), nâs (men, mankind), nice (many, much), oğlan (boy), oturak (seat), parmak (finger), söz (word), şahs/şahıs (person, individual), şehirli (urban), şikayet (complaint), şimdi (now), tabya (redoubt, bastion), tahrib (laying waste), taşra (country, outback), töhmet (fault), ulûfe (pay, salary), umûm (being general or universal), üftâde (fallen), voyvoda (voivode), vürûd (arrival, coming), yalnız (alone, only), yakın (close, near), yer (place, location), yok (no, not), ziyade (increase, more).

2) 99 keywords of 21 Ottoman chronicle whose *frequency of occurrence* is obviously much *higher* than in Naima's text, but they define the genre features ( $P<0.05$ ;  $P<0.01$ ;  $P<0.001$  or even many words from the list below are absent in "Naima's History") . These are âsafî (vizirial), avatîf (kindness), azîmet (departure), bâ'de (after, then), bâdehû (afterward), bâis (cause), berkuk (peach), binbaşî (major), boşluk (space, emptiness), bûd (was), celîl (great), cemazi (name of two lunar months), cemî (all), cenâh (wing of an army), cihânban (sovereign), çaharşenbih (wednesday), çocuk (child, boy), dâvet (invitation), dilâver (brave), dikkat (attention), ecnad (soldiers), emâret (emirate), emrillah (order of Allah), fâzîl (virtuous), fir-

kateyn (frigate), fûnûn (arts), güşâ (-opener), hâcegân (hocas), hadîs (hadith), halledallah (religion formula), hâm (unprepared), hândeک (ditch), haziran (June), hengâm (time), humbara (bombshell), ibtidâr (setting about any work), imâm (imam), imrâr (to cause to pass), inâm (favours), irâde (will), İslâmbol (lots of Islam; Istanbul), itaât (obedience), ka'de (repose by sitting down), kahraman (hero), kaim (existing), kal' (pulling up or out), kalga (qalga), kâzı (kadi), kemâyenbagî (as is proper), mâ'adâ (besides), mâdâ (what is past), mamur (prosperous), meâb (resort), mârifet (ingenuity), ma'zûl (superseded), mehâbetlü (awful), merâsim (ceremony), meşhûd (seen), meyân (middle), mihmândâr (mihmandar), muhâfiz (guard), mu-rahhas (having permission), mûtad (accustomed), mübâderet (beginning), müderris (teacher, professor), mükemmel (excellent), müteveccih (pointed), nâil (who receives) nâire (fire), nebî (messenger), nezâret (supervision), nümâ (which shows), pâd (keeper), pehlivan (pehlivan), rahş (beautiful horse), reîs (chief), sâir (diğer), sallallah (may God look with favour), seferiyye (related to campaign), sel (flood), semmûr (sable), sÛdde (gates), sÛvâri (cavalry, cavalryman), şehinşâh (king of kings), şemşir (sword), şenbih (day), şerefyâb (to be honoured), şÛd (going), tâhir (clean), Tanrı (God), Teâlâ (May his name be glorified), tebâreke (consecration), tebcil (reverence), tesyîr (sending), velîk (yet), vÛrud (arrival), zâbitân (officers), zindîk (apostate).

3) 99 key words, which frequencies are equal for "Naima's History" and other works of the genre. These are âliye (superior), altı (six), aslan (lion), ateş (fire), avrat (women), azametlü (lordly, pompous), başarde (big galley), bilâd (cities, countries), birâder (brother), câriye (bondmaid), ce-behâne (arsenal), cisr (bridge), cuma (friday), cÛlÛs (enthronement), çekdiri (rowboat), dost/dostluk (friend/friendship), dÛş (shoulder, dream), dÛvel/dÛveli (states/*adj.* state), firistâde (sent), ganâyim (spoils, plunder), gene (again), giriftâr (seized), hâberdâr (informed), hacı (hajji), hadim (servant), harbiye (pertaining to war), harem (harem), harik (fire), hatun (hatun), her (each, every), hesab (account), hususâ (especially), HÛsrev (Khosrov), hÛsrevâne (princely), ibâdullah (God's servants), idâm (killing), ilâm (declaring), inâyetlü (favourable), inşallah (hopefully), istirahat (rest), iyi (good), kâimmakam (kaymakam), kalyon (galleon), kerametlü (noble), kethüdâlık (function of kethüda), kız (girl), kub (who bits), kudretlü (puissant), levâzım (supplies), levendât (levends), mahsur (surrounded), mahzar (place of audience), malûm (known), ma'rûf (favorably known), medrese (madrasa), memnun (satisfied), menâkıb (praiseworthy traits of character), merkume (written), mesmû (heard), meteris (trenches), mızrak (spear), mÛceb (requiring), muhabbet (affection), mukabele (mutuality), mumaileyh (the aforementioned), mübaşir (bailiff), mÛlÛkâne (royal), mÛkerreme (honoured), mÛsâ'ade (permission), mÛsalaha (peace), mÛslÛman (muslim), mÛte'ayyin (distinguished), neferât (persons), nef'î (advantegous), nevâziş (careness), nitâk (belt), oğul (son), ömr (life), pişgâh (the space in front of anything), radiyallah (may God be pleased!), rahman (God of mer-

cy), sadâkat (loyalty), sâhil (seashore), sâhir (magician), sefine (ship), sefir (ambassador), şehbâz (man of courage), şey (thing), tâd (thing), tahlis (saving), tamâm (complete, done), tavâif (nations), temşiyet (causing to move on), te'lifât (writings), türk (Turk), umûmen (generally), vakiyye (okka), vâlâ (high), varoş (suburb), vasiyyet (making a testamentary bequest), vilâdet (nativity), yalan (lie), yekşenbih (Sunday).

The analysis of three groups of keywords showed that simpler and close to modern Turkish language lexicon is prevailing in "Naima's History". We see this not only in keywords that differ the text from the rest of Ottoman prose like: ancak, bugün, boyun, duvar, geçit, gıyureş, derviş, herif, oğlan, kırmızı, parmak etc, but even in keywords that are common for this chronicle and the rest of Ottoman prose: aslan, çekdiri, gene, hesap, iyi, kız, kudretlü, memnun, türk, varoş, yalan etc. Meanwhile many difficult for acoustical perception of Turk listener, considering the tradition of reading chronicles in the yards of Ottoman aristocrats, Arab and Persian loans are quite rare among key words of "Naima's History". They are present in a second group of keywords and represent a genre of Ottoman chronicles as well.

The data of Turkic verb distribution in 21 Ottoman prose works and "Naima's History" even more supplements these conclusions. Chi-square calculated between the frequencies of the most used verbs of Turkic origin in NHC and OCC shows that they obviously prevail in "Naima's History" as represents the Table 4 (Ek Tablo: 4). Verbs in the Ek Tablo: 4 also imply masdars as far as we united these different forms in the process of lemmatization. The statistical output for only three of selected lemma indicates that they are more frequent in 21 text of OCC. Meanwhile quite a number of originally Turkic verbs prevail in "Naima's History". As it is shown below these are: almak (to get, take); bilmek (to know); bulmak (to find); bulunmak (to be present, situated); çekmek (to attract); çıkmak (to exit); etmek (to do); geçme (to pass); gitmek (to go); göndermek (to send); göstermek (to show); kalmak (to stay); varma (to arrive); verilme (to be given); vermek (to give).

Such domination of Turkic verb in "Naima's History" naturally is the evidence of other individual stylistic tendency, which is prevailing of Turkic affixes at morphemic-syntactical level of the text. As the Table 5 (Ek Tablo: 5) shows, all affixes of the main tenses of Turkic verb obviously prevail in "Naima's History".

The same we can say about affixes of imperative and conditional mood of the verb and affixes of the most spread in Ottoman Turkish adverbial participle -ıp/-ip/-up/-üp-ib/-ib/-ub/-üb as it is shown in the Table 6 (Ek Tablo: 6). While distribution of other important verbal participles of Turkish origin like iken/ken (1766 in "Naima's History"; 1957,19 in Ottoman prose corpus) and ince/inca (1613,38 and 1374,18) are equal ( $P \chi^2 > 0,05$ ).

On the base of Table 6 (Ek Tablo: 6) it should be accentuated as well that some other important and most frequent Turkic morphemes obviously

dominate in Naima's History as compared with 21 Ottoman prose works. These are affixes that combine with verbal morphemes like: -ız/-iz/-uz/-üz/-k (first person of plural number; 818,44 and 1347,89; <0,001); -n/-ın/-sın (second person of singular number; 486,01 and 1552,11; <0,001); -lar/-ler/ (third person of plural number; 9522,25 and 16 521,84; <0,001). Also affix -lar\ -ler if to take it as universal without dividing its grammatical functions that can be indicating of third person of plurality or plural form of the nouns takes place with frequency 50167 in Naima's History and 48 154 in OCC ( $P \chi^2 < 0,01$ ). Affix -dir/-dir/-dur/-dür/-tır/tır/-tur/-tür also dominates in Mustafa Naima's work with frequency 10 527, 82 against 9282, 78 in OCC ( $P \chi^2 < 0,001$ ). Calculation of the quantitative parameters of the rest of grammatical forms showed that even cases where Turkic affixes less represented in "Naima's History" in comparison with other Ottoman works are not so much important because of their low frequencies.

Yet according to the given data more Turkic phonation and unadorned individual style of Naima's work in comparison with many other Ottoman prose works is obvious. At the same time we can't say that quantitative traits of Naima's "stylome" cardinally differentiate it from the rest of Ottoman prose.

As one of the first editors of Mustafa Naima's work Ibrahim Müteferrika maintained, this chronicle, properly, was obliged for its popularity among different classes of Ottoman society do to the Naima's style. It was not as "overly-sweet" and pompous as in the rest of Ottoman narratives (Lewis 1972: 123; Müteferrika 1967: 20-26). In the same way, such researchers of the chronicle as Lewis V. Thomas (Lewis 1972: 36, 82), Z. Danişman (Danişman 1967: 7-10), A. Coşkun (Coşkun 2004: 33-35) and others accentuated that "Tarih-i Naima" differs from other Ottoman annals by more simple language. Naima's quotes of direct speech also provided numerous examples of "vulgar Ottoman" for Erich Prokosch «Studien zur Grammatik des Osmanisch-Tuerkischen unter besonderer Beruecksichtigung des Vulgarosmanischen» (Hagen 2006). Statistical analysis of Naima's work confirms above-mentioned conclusions by empirical fixation of Mustafa Naima's individual style, which according to matrices given at the beginning of our paper visualized in the form of Diagram №1 (Grafik: 1).

At this graphic one straight line, which denotes the Naima's text lies at a periphery of a curves accumulation that reflects all matrix texts, represented in Table 1. Actually, that accumulation visualizes the language norm of Ottoman chronicles functional style. At the same time the straight line (Naima's text and vocabulary) does not exceed the bounds of curves accumulation as the curve line №22 (Evliya Çelebi) does it. Properly line № 22 marks the prose work "Seyahatname" of the famous Ottoman XVII century traveler Evliya Çelebi. That work was forgotten in Ottoman Empire for many years because of its deviation from that time prose canons (Bahrevskiy 2008: 8-9). However, "Naima's History", as Grafik: 1 shows, even lying out of the center of genre canon, yet does not oppose it fully as "Seyahat-

name” does. Thus, its stylistic features exist implicitly. Still, as we may suppose, even such objection to the strict rules is a quite uncommon phenomenon in the Ottoman Turkish and Middle East Literature. It’s fixation by computational methods makes impossible any speculations in modern Turkish literary criticism and linguistics on the topic: is Naima one of the pillars of pre-modern Turkish Language and Literature, or does he belong only to the Ottoman Prose, which is too complicated and unintelligible for modern reader, has the chronicler his own style or is he simply a compiler of other Ottoman annalists works. The tentative answer is clear: Naima objected the canons of Ottoman prose and language norm in implicit form. The chronicler had his own inexpressible style, and presumably, it reflects at least his vision of historical processes and events he wrote about, or at most served for him as a tool for changing or proving the historical realities in the imagination of his contemporaries. Anyway, the implicit style must be an instrument for him to influence his contemporary reader in a way, which let the annalist work stay quite famous in the Ottoman Literature for two centuries and not repeat the fate of Evliya Çelebi’s “Seyahatname”.

### **Acknowledgments**

I’m very grateful to Ms. Yaroslava Kulchynska, who created the structural software in Microsoft Office Access database format for this research and provided all the calculations in SPSS Program. I would also like to thank Dr. Iryna Dryga for her useful comments on an earlier draft of this paper.

### **REFERENCES**

- AKTAŞ Ali (2008). *Çelebizâde Âsım Târîhi*, <https://archive.org/stream/CelebizadeAsimTarihi>, [www.yazoku.net/CelebizadeAsimTarihi\\_metni\\_djvu.txt](http://www.yazoku.net/CelebizadeAsimTarihi_metni_djvu.txt) (04.02.2018)
- ALTUN Abdullah (2006). *Said. b. Halil İbrahim'in «Tarih-i Sefer-i Rusya» Adlı Eseri (Transkripsiyon ve Değerlendirme)*, Erciyes Üniversitesi, Sosyal Bilimler Enstitüsü, Kayseri (Yayımlanmamış Yüksek Lisans Tezi).
- BAHREVSKİY E.V. (2008). "Evliya Chelebi i ego "Kniga Putishestviya". In: *Evliya Chelebi. Kniga Putishestviya. Krym i Sopridelnyye Oblasti*. Simferopol: Dolya.
- COŞKUN Ali (2004). *Osmanlı'da Din Sosyolojisi. Naima Örneği*, İstanbul: İz Yayıncılık.
- HACI Ali Efendi (2007). *Tarih-e Kamanîçe [hazırlayan Ayşe Hande Can]*, Ankara: T.C. Kültür ve Turizm Bakanlığı Kütüphaneler ve Yayımlar Genel Müdürlüğü. <http://ekitap.kulturturizm.gov.tr/Eklenti/10726,girispdf.pdf?0> (04.02.2018).
- ÇİMEN Cihan (2006). *Anonim Tevârih-i Âl-i Osman (Kuruluştan 897/1487'e kadar)*, Marmara Üniversitesi, Türkiyat Araştırmaları Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).
- ÇORUHLU Melek (2005). *Musaffa Mehemmed Efendi «Kıt'a-min-Tarih-i Sultan Mahmud-ı Evvel» (Tahlil ve Metin)*, Mimar Sinan Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi)
- DANIŞMAN Zuhuri (1969). "Önsöz". In: *Naima Tarihi: 6 Cilt [çeviren Huzuri Danışman], C.1*, İstanbul: Z. Danışman Yayınevi.
- DEMİR Uğur (2006). *Târîh-i Mehmed Giray (Değerlendirme-Çeviri Metin)*, Marmara Üniversitesi, Türkiyat Araştırmaları Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).
- DİNU L. P. – Popescu M. (2009). "Ordinal measures in authorship identification", *Overview of the 1-st International Competition on Plagiarism Detection: 1-9*.
- DÜNDAR Hasan (2006). *Rahîmîzâde İbrahim (Harîmî) Çavus'un Gence Fetihnâmesi Adlı Eseri'nin Transkripsiyonu ve Kiritizasyonu*, Kocatepe Üniversitesi, Sosyal Bilimler Enstitüsü. Afyon (Yayımlanmamış Yüksek Lisans Tezi).
- EROL Salih (2007). *Pehlivan İbrahim Paşa Vakayı'Namesi (Baba Paşa Tarihi)*, Anadolu Üniversitesi, Sosyal Bilimler Enstitüsü. Eskişehir (Yayımlanmamış Yüksek Lisans Tezi).
- GÖKÇEK Mehmet Fatih (2006). *Behceti Seyyid İbrahim Efendi "Tarih-i Sülale-i Köprülü" (transkripsiyon ve tahlil)*, Marmara Üniversitesi, Türkiyat Araştırmaları Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).
- GÖKNUR Çelik (2006). *Vâsıtî'nin "Gazâvât-ı Murad Paşa" Adlı Eserinin İncelenmesi*, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).

- GUADALUPE J.T. vd. (2008) "Similarity Measure for Clustering and its Applications", *Proceedings of World Academy of Science, Engineering and Technology*, 31: 490-496.
- GUIRAUD P. (1953). *Langage et versification*, Paris: Klincksieck.
- GÜRIŞİK Bihter (2005). *Peçevi tarihi (46b/80a metin, dizin, özel adlar sözlüğü)*, Marmara Üniversitesi, Türkiyat Araştırmaları Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).
- HAGEN Gottfried (2006). *Osman II and the Cultural History of Ottoman Historiography*.  
<http://www.h-net.org/reviews/showrev.php?id=11651> (04.02.2018).
- HUANG A. (2008). "Similarity measures for text document clustering", *Proceedings of NZCSRSC 2008*: 49-56.
- KARADAYI Aynur (2008). *Osmanlı tarih yazıcılığının gölgede kalmış bir eseri "Tarih-i Sâmî" (metnin transkripsiyonu)*, Adnan Menderes Üniversitesi, Sosyal Bilimler Enstitüsü, Aydın (Yayımlanmamış Yüksek Lisans Tezi).
- MacENERY T. (2001). *Corpus Linguistics: An introduction*, Edinburg University Press.
- MERTAYAK Aydın (2005). *Nişli Mehmet Aga'nın Rusya Sefareti ve Sefaretnamesi (1722-1723)*, Gaziosmanpaşa Üniversitesi, Sosyal Bilimler Enstitüsü, Tokat (Yayımlanmamış Yüksek Lisans Tezi).
- Mihalcea R. vd. (2006). "Corpus-based and Knowledge-based Measures of Text Semantic Similarity", *Proceedings of the 21st national conference on artificial intelligence (AAAI'06)*: 775-780.
- MÜTEFERRİKA İbrahim (1967). "Naima'nın Hal Tercümesi ve Kitabı Hakkında", *Naima Tarihi: 6 Cilt [çeviren Huzuri Danışman]*, İstanbul: Zuhuri Danışman Yayınevi.
- NAİMA Mustafa (1281-1283/1864-1866). *Tarih-i Naima: 6 Cilt*. Constantinople: Amire Press.
- NAİMA Mustafa (2007). *Tarih-i Naima: 6 Cilt [hazırlayan Mehmet İpsirli]*, Ankara : Türk Tarih Kurumu Yayınları.
- ÖZBAL Mustafa (2005). *Peçevi tarihi (80b/114a metin, edisyon kritik, dizin, özel adlar sözlüğü)*, Marmara Üniversitesi, Türkiyat Araştırmaları Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).
- ÖZKUZUGÜDENLİ Bülent (2005). *Talîkîzade Mehmed Subhî "Tebriziyye" (Metin Transkripsiyonu, Eser ve Bilgilerin Değerlendirmesi, Yazar Hakkında İnceleme)*, Marmara Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).
- POLATCI Türkan (2003). *Şehdi Osman-Efendi'nin Rusya Sefareti ve Sefaretnamesi (1757-1758)*, Gaziosmanpaşa Üniversitesi, Sosyal Bilimler Enstitüsü, Tokat (Yayımlanmamış Yüksek Lisans Tezi).
- PROHOROV A.V. (2002). Metod analiza YEYA opisaniya predmetnyh oblastey. In: A. V. Prohorov & Y. N. Phillipovich. *Semantika informatsyonnyh tehnologiy: opyty slovarno-tezaurusnogo opisaniya*. Moskva: MGUP. [http://it-claim.ru/Library/Books/Semantics\\_IT/gl1\\_2/glava12.htm](http://it-claim.ru/Library/Books/Semantics_IT/gl1_2/glava12.htm) (04.02.2018).

- SAKA Pınar (2007). *Derviş Abdullah «Risale-i Teberdariyye Fi Ahval-i Darü's-Sa'ade»*. Değerlendirme – Çeviri – Metin, Marmara Üniversitesi, Türkiyat Araştırmaları Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).
- SALTON G. (1989). *Automatic Text Processing*, New-York: Addison-Wesley.
- SETYAGYAYEV N (2002). “Evlıya Çelebi Tevarih-ı Tohta Bay qavli üzere ibtida Qırım ceziresine malik olan hanları beyan ider. Transcription and translation”, *Nauchnyi bulletin*, 2: 3–8.
- SINCLAIR John (2004). *Developing Linguistic Corpora: A Guide to Good Practice. Corpus and Text – Basic Principles*. [http://icar.univ-lyon2.fr/ecole\\_thematique/contaci/documents/Baude/wynne.pdf](http://icar.univ-lyon2.fr/ecole_thematique/contaci/documents/Baude/wynne.pdf) (04.02.2018)
- SÖYLEMEZ Hatice (2007). *Mukaddimetü's Sefer (1736 – 1739 Seferi Hakkında Bir Eser)*. Metin, Değerlendirme, Marmara Üniversitesi, Türkiyat Araştırmaları Enstitüsü, İstanbul (Yayımlanmamış Yüksek Lisans Tezi).
- TİMURTAŞ Faruk (1997). *Osmanlı Türkçesi Grameri – III*, İstanbul: Alfa.
- LEWIS V. Thomas (1972). *A Study of Naima*, New York: New York University Press.
- McCain K.W. (1990). “Mapping Authors in Intellectual Space: A Technical Overview”, *Journal of the American Society for Information Science*, 41: 433-443.
- RAMOS J. (2003). Using TF-IDF to Determine Word Relevance in Document Queries. <http://www.cs.rutgers.edu/~mlittman/courses/ml03/iCML03/papers/ramos.pdf> (04.02.2018).
- TÜBENÇOKRAK Öznur (2007). *Keyfiyet-i Rusiyye'nin Transkripsiyon ve Tahlili (H. 1206 / M. 1791-1792)*, Osmangazi Üniversitesi, Sosyal Bilimler Enstitüsü, Eskişehir (Yayımlanmamış Yüksek Lisans Tezi).
- WHITE H.D. (2003). “Author Cocitation Analysis and Pearson's r”, *Journal of the American Society for Information Science and Technology*, 54 (13): 1250-1259.
- YASTI Mehmet (2005). *Nişancı Mehmet Paşa «Tevarih-i Al-i Osman» (1b-120a)*, Selçuk Üniversitesi, Sosyal Bilimler Enstitüsü, Konya (Yayımlanmamış Yüksek Lisans Tezi).
- YILDIRIM Nermin (2005). *Kara Çelebi-Zade Abdülaziz Efendi'nin Zafername adlı eseri (Tarihçe-i Feth-i Revan ve Bağdad). Tahlil ve metin*, Mimar Sinan Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul: (Yayımlanmamış Yüksek Lisans Tezi).



**ADDENTA**

Table 1: Correlation matrix of Ottoman chronicles lexicon

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
1																							
2	0,858																						
3	0,862	0,874																					
4	0,863	0,866	0,976																				
5	0,9	0,891	0,868	0,857																			
6	0,892	0,868	0,875	0,906	0,885																		
7	0,891	0,897	0,882	0,889	0,885	0,889																	
8	0,918	0,906	0,926	0,936	0,916	0,889	0,914																
9	0,917	0,858	0,84	0,864	0,876	0,881	0,926	0,926															
10	0,927	0,85	0,833	0,859	0,874	0,888	0,927	0,981	0,981														
11	0,925	0,848	0,873	0,874	0,864	0,907	0,932	0,953	0,972	0,972													
12	0,865	0,823	0,927	0,892	0,896	0,896	0,875	0,842	0,844	0,893	0,893												
13	0,735	0,771	0,837	0,863	0,891	0,896	0,881	0,806	0,797	0,742	0,797	0,797											
14	0,862	0,874	0,837	0,866	0,897	0,885	0,803	0,816	0,729	0,787	0,819	0,819	0,787										
15	0,862	0,874	0,89	0,976	0,868	0,882	0,892	0,935	0,936	0,944	0,893	0,893	0,944	0,936									
16	0,863	0,866	0,841	0,857	0,885	0,899	0,838	0,817	0,835	0,836	0,755	0,755	0,836	0,835	0,845								
17	0,9	0,891	0,847	0,857	0,906	0,885	0,915	0,926	0,954	0,957	0,845	0,845	0,957	0,972	0,827	0,827							
18	0,892	0,868	0,839	0,856	0,885	0,902	0,902	0,911	0,953	0,947	0,827	0,827	0,947	0,968	0,917	0,827	0,89						
19	0,891	0,888	0,902	0,922	0,884	0,917	0,89	0,941	0,93	0,945	0,917	0,917	0,959	0,938	0,941	0,945	0,945	0,938					
20	0,862	0,874	0,879	0,915	0,855	0,89	0,852	0,848	0,733	0,783	0,824	0,824	0,783	0,734	0,734	0,783	0,783	0,734	0,733				
21	0,862	0,874	0,844	0,824	0,829	0,696	0,696	0,7	0,684	0,708	0,709	0,709	0,708	0,678	0,678	0,684	0,684	0,678	0,684	0,684			
22	0,862	0,874	0,717	0,73	0,672	0,696	0,696	0,7	0,684	0,708	0,709	0,709	0,708	0,678	0,678	0,684	0,684	0,678	0,684	0,684	0,684		

Ottoman Chronicler Mustafa Naima's Individual Style in Corpus Dimension

22	0,69	0,694	0,717	0,73	0,672	0,696	0,696	0,7	0,684	0,678	0,708	0,709	0,674	0,734	0,695	0,706	0,685	0,73	0,663	0,663
14	0,76	0,754	0,81	0,816	0,77	0,783	0,803	0,816	0,721	0,729	0,787	0,819	0,798			0,841	0,698	0,741	0,722	0,81
15	0,913	0,866	0,89	0,912	0,877	0,907	0,892	0,935	0,929	0,936	0,944	0,893	0,827	0,841			0,839	0,931	0,924	0,81
16	0,798	0,841	0,785	0,809	0,804	0,81	0,838	0,817	0,84	0,835	0,836	0,755	0,73	0,698	0,839	0,868	0,867	0,924	0,863	0,81
17	0,931	0,878	0,847	0,869	0,885	0,9	0,915	0,926	0,954	0,972	0,957	0,845	0,723	0,741	0,868		0,975	0,949	0,951	0,792
18	0,922	0,868	0,839	0,856	0,865	0,885	0,902	0,911	0,953	0,968	0,947	0,827	0,719	0,722	0,867	0,975	0,945	0,916	0,818	0,844
19	0,918	0,888	0,902	0,922	0,884	0,907	0,917	0,941	0,93	0,938	0,945	0,89	0,827	0,81	0,863	0,949	0,945	0,938	0,951	0,844
20	0,917	0,842	0,879	0,915	0,855	0,906	0,89	0,913	0,927	0,941	0,959	0,917	0,762	0,792	0,818	0,929	0,916		0,785	0,844
21	0,8	0,796	0,844	0,824	0,829	0,805	0,852	0,848	0,733	0,734	0,783	0,824	0,818	0,844	0,732	0,776	0,764	0,831	0,785	0,844
14	0,76	0,754	0,81	0,816	0,77	0,783	0,803	0,816	0,721	0,729	0,787	0,819	0,798			0,841	0,698	0,741	0,722	0,81
15	0,913	0,866	0,89	0,912	0,877	0,907	0,892	0,935	0,929	0,936	0,944	0,893	0,827	0,841			0,839	0,931	0,924	0,81
16	0,798	0,841	0,785	0,809	0,804	0,81	0,838	0,817	0,84	0,835	0,836	0,755	0,73	0,698	0,839	0,868	0,867	0,924	0,863	0,81
17	0,931	0,878	0,847	0,869	0,885	0,9	0,915	0,926	0,954	0,972	0,957	0,845	0,723	0,741	0,868		0,975	0,949	0,951	0,792
18	0,922	0,868	0,839	0,856	0,865	0,885	0,902	0,911	0,953	0,968	0,947	0,827	0,719	0,722	0,867	0,975	0,945	0,916	0,818	0,844
19	0,918	0,888	0,902	0,922	0,884	0,907	0,917	0,941	0,93	0,938	0,945	0,89	0,827	0,81	0,863	0,949	0,945	0,938	0,951	0,844
20	0,917	0,842	0,879	0,915	0,855	0,906	0,89	0,913	0,927	0,941	0,959	0,917	0,762	0,792	0,818	0,929	0,916		0,785	0,844
21	0,8	0,796	0,844	0,824	0,829	0,805	0,852	0,848	0,733	0,734	0,783	0,824	0,818	0,844	0,732	0,776	0,764	0,831	0,785	0,844
22	0,69	0,694	0,717	0,73	0,672	0,696	0,696	0,7	0,684	0,678	0,708	0,709	0,674	0,734	0,695	0,706	0,685	0,73	0,663	0,663

Table 2: Vocabulary correlation of Naima's History and Ottoman prose

	T1	T2	T3	ALLT	Corpus
T1		0,980	0,908	0,995	0,753
T2	0,980		0,949	0,992	0,761
T3	0,908	0,949		0,943	0,792
ALLT	0,995	0,992	0,943		0,771
Korpus	0,753	0,761	0,792	0,771	

Table 3: Keywords: Naima's History vs. Ottoman Prose Canon. Example

Word	Translation	Fpm (OCC)	Fpm (NHC)	P Xi <sup>2</sup>	Word	Translation	Fpm (OCC)	Fpm (NHC)	P Xi <sup>2</sup>
Avatif	Kindness	47.5			Inâyetlü	Favour-able	54.6	10.2	
Birâder	Brother	129.3	194.0		Kerametlü	Noble	45.5	40.9	
Dostluk	Friendship	155.6	40.9	<0,01	Muhabbet	Affection	111.2	153.2	
Fâzıl	Virtuous	144.5	61.3	<0,05	Sadâkat	Loyalty	160.7	91.9	
Inâm	Favours	32.3			Şerefyâb	Be honoured	32.3		

Table 4: Turkic verb frequencies: Naima's History VS Ottoman Prose canon

Word	Translation	Fpm (OCC)	Fpm (NHC)	P Xi <sup>2</sup>	Word	Translation	Fpm (OCC)
Olmak	22315,3	22587			Kılmak	748,7	275,70
Olunmak	8778,6	6576,1	<0,001		Kalmak	737,6	1255,9
Eylemek	7388,2	5217,9	<0,001		Bulunmak	736,6	919,01
Varmak	1264	2889,8	<0,001		Verilmek	735,6	1531,7
Gitmek	1139,8	2348,6	<0,001		Göndermek	733,6	1929,9
Bulmak	1012,5	1347,9	<0,01		Çıkmak	728,5	1868,7
Vermek	1765,2	3379,9	<0,001		Girmek	669,9	796,5
Etmek	16300	25569	<0,001		Geçmek	576,9	1276,4
Görmek	1311,5	1450			Bilmek	543,6	939,4
Almak	1285,3	2123,9	<0,001		Göstermek	321,3	714,8
Çekmek	289,9	612,7	<0,001		Atmak	319,3	428,9

Table 5: Distribution of Turkic verbal morphemes in NHC & OCC

Affix	Tense/Aspect	Fpm (OCC)	Fpm (NHC)	P Xi <sup>2</sup>
-ir/-ir/-ur/-ür	Aorist	6120,15	9139,09	<0,001
-dı/-di/-du/-dü	Praeterite	25177,8	33115,15	<0,001
-miş/-miş/-muş/-müştü	Dubitative (Indirective Past Tense)	4855,09	6913,03	<0,001
ırdı/-irdi/-urdu/-ürdü	Aorist Past	601,20	1878,87	<0,001
-makta/-mekte	Imperfective	17, 18	183, 8	<0,001
-mişti/-mişti/-muştu/-müştü	Plusperfect	731,55	1245,78	<0,001
-irmiş/-irmiş/-urmuş/-ürmüş	Simple Present (Indefinite)	24,25	153,169	<0,001
-yormuş	Present Continuous (Indefinite)	1,01	10,21	<0,05

Table 6: Distribution of Turkic verbal morphemes in NHC & OCC

Affix	Category	Fpm (OCC)	Fpm (NHC)	P Xi <sup>2</sup>
-alım/-elim/-sınlar/-sinler	Imperative of 1 p. pl.	962,9356	1245,77	<0,01
-sa/-se	Desideratif (Conditional Tense)	1983,47	3124,65	<0,001
-dikça/-dikçe	Gerundium (converb)	298,08	418,66	<0,05
-ıp/-ip/-up/-üp-ib/-ib/-ub/-üb	Gerundium (converb)	30232,94	46890,16	<0,001
-dik/-dik/-duk/-dük/-	Subordinate Clause	10693,33	13958,81	<0,001

Graphic 1:

*Diagram 1*  
Ottoman chronicles lexical norm and Mustafa Naima's idiosyle

