

Izvorni naučni rad

UDK 811.124'373.7:373.5(497.5)

Matija ZORIĆ

Faculty of Humanities and Social Sciences,
Josip Juraj Strossmayer University in Osijek
trozub@gmail.com

**FREQUENCY OF LATIN PHRASEOLOGICAL UNITS IN THE
PERSEUS DIGITAL LIBRARY GREEK AND ROMAN (LATIN)
CORPUS AND THE CROATIAN LANGUAGE LIBRARY:
TEST PROBE A**

Education systems worldwide use a large number of Latin sayings and proverbs popularly called *Dicta et sententiae* (hereafter: D&S) which are taught randomly, without sub-categorizing. Usually, their importance to the educational area of a particular school is taken into account. This may suit law schools to some extent due to the formulaic nature of the legal discourse, while the choice of sayings in high school, medical, philosophical, and social-humanistic schools, in general, is left to the subjective choice of the teacher. The largest Croatian collection of Latin D&S *Latinum in aeternum* (LIA), Marević (2002) contains 18,632 references, and in secondary schools, 100–200 D&S are taught in two to four years of learning Latin. The catalog of the Croatian state graduation exams in Latin language (2016–2021) contains less than 200 sayings. It is therefore imperative to determine the frequency of Latin D&S in the largest possible corpus of Roman (and later Latin) literature and sub-categorize D&S, thus providing school systems around the world with a much-needed catalog based on D&S frequency as a tool for compiling teaching and examination catalogs. Manual entry of sayings into the corpus search engine and the extensiveness of the LIA collection require too much effort and time for a single scientific paper. Therefore, the corpus representation of the test probe of 1270 D&S from the LIA collection starting with the letter A in the Perseus Digital Library Greek and Roman Materials corpus is investigated here and also in the corpus of the Croatian language repository (HJR) of the Institute for the Croatian Language and Linguistics. The search results are useful for teaching Latin worldwide and for teaching Latin and Croatian in Croatian schools.

Keywords: frequency of Latin sayings, Latin phraseological units, dicta et sententiae, Latin phraseology, Latin in Croatian

INTRODUCTION

Since ancient times, Latin sayings and proverbs have not only been a matter of communication within certain registers but also a sign of prestige for the more educated classes of society, who like to flaunt their knowledge of the rich fund of Dicta et sententiae (D&S). Since Ancient Latin is linked exclusively to the written corpus, the pool of D&S from Ancient Latin should be quite limited. Considering the variety of types and subtypes of phraseological units, relying on the categorization proposed by Omazić (2015), the exact number of D&S cannot be determined, despite the limited nature of the corpus of Classical Latin because it is possible that a previously neglected quotation from Roman literature experiences a boom in popularity or that an expression undergoes the process of figuration or fossilization (Omazić 2015) and becomes frequent in the future. It should not be overlooked that Latin is not completely dead just because of the static syntax. Latin experienced the greatest expansion of its vocabulary after the classical period as evidenced by professional terminology in botany, zoology, pharmacopeia, etc. Numerous Latin D&S were created and remained in use after Latin ceased to be the native language, such as *Viribus unitis*, written by none other than the Austrian ruler Franz Joseph I. In numerous small or large collections of Latin D&S, no distinction was made between D&S originating from the classical corpus of antiquity and those that arose through the medieval or modern use of the Latin language. Nevertheless, the written tradition has preserved the association of some of the D&S with historical figures and attributes authorship of these D&S to them, which helps to determine their origin. The largest collection of Latin D&S in contemporary Croatian literary heritage is *Latinski zauvijek Latinum in aeternum* (hereafter: *Latinum in aeternum*) (Marević 2002), which indiscriminately contains more than 18,000 Latin expressions without categorizing them into semantic or syntactic subtypes. Due to the lack of categorization, it can be concluded that the selection criterion is the frequency of the use of D&S in different registers, or that the collection simply wanted to cover as many D&S as possible. Frequency could also be the most important criterion for the relevance of D&S in systems where hundreds of D&S examples are learned. Considering the results of some corpus studies on phraseological units (PU) in the English language, which have shown that the frequency of most PUs in written corpora is generally low (Moon 1998; Biber et al. 1999), as well as the shining glory of the tradition of

Latin D&S and their distribution, it is necessary to conduct a comprehensive corpus review of Roman and Latin literature that can provide relevant statistics for studies of the Latin language worldwide. Since the Croatian language is one of the many languages with a strong influence from the Latin tradition, it is quite appropriate to investigate the distribution of Latin D&S in the Croatian corpus as well, which can provide relevant statistics on the use of Latin D&S in the modern Croatian language. Since analyzing 18,000 PUs exceeds one research paper, a test sample of 1,270 PUs beginning with the letter A was taken from the *Latinum in aeternum* collection, while the corpus search in this paper focuses on the Latin corpus of the Perseus Digital Library Greek and Roman Materials (PDLGR) and the Croatian Language Treasury (Hrvatska jezična riznica, HJR) of the Institute of Croatian Language and Linguistics. The sample size is sufficient to provide a statistical cross-section from various aspects and to hypothesize about the nature, categories, and characteristics of the rest of the D&S corpus and to establish, at least roughly, numerical values that we can consider as thresholds for the frequent and rare use of Latin D&S in both Latin and Croatian.

I METHODS

I.1 Corpus

In order to analyze the frequency of Latin D&S in this paper, the corpus selected was the Perseus Digital Library; Greek and Roman Materials (PDLGR) and the selected search language was Latin. This digital collection contains original texts of Greek and Roman literature organized by author's name in a joint index for Greek and Roman authors. This collection also contains a search engine that can search individual words or multi-word phraseological units, making it a good candidate for a corpus. When using the search engine, it is possible to limit the text corpus to the Latin language by simply selecting the language of preference. The Latin part of the corpus comprises 10,525,338 words and contains the most important writers and works of the classical period as well as a significant corpus of Italian medieval (Renaissance) texts in Latin. In some cases, the corpus contains several editions of the same work according to its traditional popularity, which may increase the number of occurrences of some of the words or phrases searched. Multiple editions refer to more popular texts in educational institutions, which reinforces the legitimacy of the corpus.

1.2 Latin phraseological units

Most types of phraseological units according to modern categorization are found in the Latin D&S. For this paper, it suffices to present examples according to the taxonomy based on syntactic structure (Everaert et al. 1995:4) for idioms in the broader sense such as sayings (and proverbs) as presented in (1), routine formulae (2) and phrase like units, i.e. of compounds or units with personal names (3):

- (1) *Gutta cavat lapidem non vi, sed saepe cadendo.*
A drop does not penetrate a stone by force, but by falling frequently.
- (2) *Senatus decretum.*
Decision of the Senate.
- (3) *Domus Augusta.*
Domus Augusta (Sublime House).

Although the Latin phraseological inventory can provide an example of much more detailed categories of phraseological units, as elaborated by Omazić (2015), this paper remains focused on the frequency of the inventory of phraseological units and not on the qualitative analysis. The claim that unanalyzed parts of speech are stored in memory and retrieved as such, some of them being formulas of the language and others having an idiosyncratic character (Peters 1983), probably applies to Latin phraseological units more than to other phraseological thesauri, as the knowledge of Latin D&S far exceeds the knowledge of Latin due to its widespread use. Hence the necessity of traditional, analytical learning of Latin instead of the limitation of only learning D&S is often questioned without justification. When examining research dealing with the comprehension of idioms in L1, Panou (2017: 9) claims that the literal or figurative processing of words during online utterance interpretation is not an exclusive way of understanding language. But Latin is not an L1 for anyone in the modern world, and it is an L2 for only a few. The need to learn Latin as the L2 is particularly evident in D&S because many people who know Latin D&S are not familiar with the basics of Latin syntax and lexis, which is a strong limitation for possible modifications. YegAnehjoo (2012: 98) concludes that a unified representation of idiomatic expressions in the L2 is directly linked to its constituent lemmas. Such a finding explains two important phenomena: first, Latin PUs can be used by speakers who do not know Latin syntax or lexis at all.¹ Furt-

¹ YegAnehjoo (2012) states that the findings indicate that idiomatic expressions can be represented and stored in the L2 mental lexicon.

hermore, they can even use a large number of Latin PUs; and second, the use of Latin PUs is much easier for those who understand the constituents of the PUs, if not in detail, at least at the level of basic lemmas. A particular phenomenon of the Latin language is the use of Latin D&S by non-Latin speakers in third languages, where Latin D&S are transferred as indivisible expressions associated with an inseparable meaning so that their perception and comprehension could largely resemble the perception and comprehension of lexical items. Such an understanding of idiomatic expressions allows non-Latin speakers to easily use Latin D&S in any language they happen to speak. Omazić (2015) cites features of PUs but emphasizes they are relative, which can be extremely important when analyzing Latin D&S when it comes to the feature of multilexicality. Older views defining idiomatic expressions as multilexemic expressions (van der Linden 1992) should be understood as a general rather than an exclusive feature of the idiom category. By relativizing multilexicality as a feature of a PU, it is possible to include in the phraseological corpus of Latin expressions such as *alibi*, which has both metaphorical and metonymic meanings and is likely to be well represented in the corpora. This does not, in any case, mean that every individual word in the lexicon of a language is also an idiomatic expression.

Among the semantically transparent PUs (Everaert et al. 1995: 4) in Latin, it is easy to find examples of collocations (4), verbal collocations (5), irreversible binomials (6), comparative idioms (7) and specializations (8):

- (4) *Toga praetexta*.
Crimson-trimmed toga.
- (5) *Bellum gerere*.
To wage war (to fight in the war).
- (6) *Veneres Cupidinesque*.
Venuses and Cupids.
- (7) *Traiano melior*.
Better than Trajan.
- (8) *Spongia ex lacte*.
Eggs with honey (literally: spongy eggs with milk).

The analysis of the rich inventory of Latin D&S can be carried out according to more detailed or more general modern categorizations. A more general and simpler categorization of Latin D&S is proposed by Grudeva (2014: 93) as three lexical groups of Latin phrases: clichés, proverbial expressions, i.e. classical sayings and proverbs, and quotations from writers. It is interesting to note that Grudeva also includes monolexical expressions among the phraseological units, with *postfactum* and *alibi* being examples

of clichés. This paper aims to provide the total frequency of the expressions searched for and to allow teachers or learners of Latin phraseological units to look up the number of examples in each selected category. Thus, an individual may decide to teach all types of Latin D&S or only certain categories, but in either case, they will have the exact number of each example in each category. The ubiquity of Latin and its components in European and many world languages is beyond question and it ranges from phonology to lexicon and phraseology, since the human brain processes formulaic language units faster than non-formulaic ones (Tabossi, Fanari & Wolf 2009: 533-535), this finding should be all the more valuable for Latin, as the vast majority of speakers or listeners of Latin D&S are unable to analyze formulaic Latin due to their unfamiliarity with Latin syntax and lexis. Therefore, it can be concluded that the inability of users to perform a componential analysis of expressions has not prevented the widespread use of Latin D&S among those who do not know Latin. Moreover, the phenomenon of the Latin language not only surviving but also evolving despite the absence of native speakers provides perhaps the best evidence of the ultimate possibilities of using a formulaic language. The few who have mastered Latin syntax produce a PU, and those who have not mastered Latin syntax accept the received PU as a formula that is easy to pronounce and memorize. The meaning of such a PU is also accepted uncritically and akin to memorizing the meaning of a single word.

I.3 Corpus search

Since this paper aims to determine and compare the frequency of Latin D&S in the Latin and Croatian corpora, the search is carried out with or without adjustments, depending on the search engine in the individual corpus. The HJR corpus search engine accepts all entries and lists a clear numerical result. The frequency of an item per million words is simply calculated by dividing the number obtained by 90 since the size of the HJR is 90 million words. The Latin PDLGR corpus requires a much greater methodological adjustment due to the limitations of the search engine, as it is a large collection of texts that is primarily a philological and literary corpus and can also be used as a linguistic corpus with careful methodology. In the PDLGR search engine, it is necessary to select the Latin language as the search domain without additional restrictions so that the corpus contains not only ancient texts but also texts from the Italian Renaissance (it is assumed that part of the Latin D&S originates from the Italian Renaissance). The PDLGR corpus contains 10.5 million words, so the frequency of an item per million is calculated by dividing the number

obtained by 10.5. The following 9 points describe possible adjustments necessary for the PDLGR corpus search engine:

In the clear-cut case, the search engine returns a simple result (9):

(9) Aurea aetas (1 token)

which means that the corpus contains a total of 1 occurrence of this saying. However, it is necessary to thoroughly check and consider the following conditions:

1. The search engine lists numerous documents, thus, in the upper right corner, it shows the number of documents for which the results are displayed, and it is necessary to open the entire list.
2. Within a display sheet, the search engine lists the documents containing the item, while the number of tokens is in the upper right corner. If the number is not displayed, the number of tokens of an item is 1.
3. Sayings containing at least two meaningful, preferably consecutive words are the easiest to search because the search engine does not allow the search for Latin prepositions and verbs, as these are classified as so-called “stop words”. It is possible to request such two-part meaningful strings in the search engine as an inseparable required phrase; the search engine then returns an immediate and precise result.
4. It is more difficult to search for those D&S that contain prepositions (10) or verb to be in a multi-part expression, which must be omitted from the search:

(10) Argumentum (e) contrario.

because in such cases it is not possible to request a search for an inseparable required phrase, but a search for the required words of the phrase. The search engine then returns results for all texts in which both words occur, which do not have to be in the same phrase, sentence, or even part of the text. The search is then performed by opening each quoted part of the text to determine whether it is a searched phrase or a separate occurrence of the word(s). In this situation, there is a visual aid from the search engine, as the words searched for are always shaded blue, so it is easier to recognize PUs because their constituents are closely spaced and shaded blue. All results must be added up gradually, carefully, and manually, omitting the large excess of results related to the separate occurrence of the word(s). In the given example (10), the most accurate result is obtained by searching for the word *contrario* and browsing through each example individually (i.e., manually) to determine whether the word of interest belongs to the phrase searched for. The inability to search for prepositions is the major disadvantage of the search engine and a massive obstacle to searching the corpus.

5. It is most difficult to search for D&S that contain only two words, one of which is a preposition (11) because the search engine does not allow the search for prepositions:
(11) (Ad) *nauseam*.
In such a situation, the only possible search for the word *nauseam* is the accusative singular. The search engine returns results that are not all PUs and need to be checked thoroughly, as in the case in point 4. This case is possibly even worse because one-word searched item occurs much more frequently in the texts than two-word items searched for. The only rescue from this painstaking labor is found with words such as *nauseam*, which are not particularly common, especially not in the accusative case.
6. It is necessary to take into account the reversible binomials and enter such an option in the search engine, especially when searching for inseparable PUs. For example, the introductory example (9) *Aurea aetas* should also be entered as *Aetas aurea*. A reliable result is obtained only by combining both searches.
7. It is necessary to take into account the modifications that have long been fossilized as variants of the PU (12). Thus, all occurrences of the variant *Argumentum a contrario* must be added to all examples found from point 4 for (10) *Argumentum e contrario*. Ofcourse, a more thorough knowledge of the Latin language is required for such a step, since many known or possible variants do not occur in collections of sayings or can be found in other places, such as the ellipsis, *A contrario* (12).
(12) *Argumentum a contrario/Argumentum e contrario/A contrario*
8. The problem of single-membered expressions is solved by accepting the relativization of the characteristics of PUs (Omazić 2015), according to which a PU can contain only one word. It is taken into account that some words have become PUs, such as *alibi*, due to their deeper metaphorical meaning, context, and register of use, and they are also included in the search. For this paper, the term monolexical PU is used. The issue with the monolexical PUs lies in the search result containing both the ordinary use of the word and the phraseological use of the word as a monolexical PU, which can only be separated by examining each token individually. It is reassuring that such examples are extremely few and far apart in the Latin D&S.
9. Searching for one-letter abbreviations such as (13) is generally not possible according to the search engine settings. The same applies to better-known examples, that can only be found as whole words (14). However, there are also rare examples that the search engine accepts

and shows the result, as in the case of (15). The results for the abbreviation (if it could be determined) and the whole term in this paper are listed both individually and in total.

(13) A.A. (for Aulus Agerius)

(14) A.a (Ad acta)

(15) A.U.C/AUC (Ab Urbe condita)

When using the search engine, a warning was accepted (Benko&Rausova 2020) that in the case of Latin D&S, search engines of some corpora can distinguish between lowercase and uppercase and return faulty results. The search engine of the PDLGR corpus did not show lower- and uppercase sensitivity even in a more aggressive check. However, the above-mentioned paper by Benko and Rausova (2020) presents other findings that are important for this paper. It is found that some Latin phrases are numerically more represented (more popular) in the Russian corpus than in the Latin, of which *alma mater* is an excellent example, and a similar finding is expected in this paper when comparing the Latin and Croatian corpora. Furthermore, Benko and Rausova (2020) have stated “proof of feasibility” as the aim of their paper, which is evident from their technical manipulation of the corpus search engine, ruling out a manual corpus search as impossible. The PDLGR corpus does not have such capabilities (and it can be said that it also has numerous shortcomings) in its search engine, and the approach described in points 1-9 is often condemned to manual search with partial help from the search engine. In this respect, this approach in points 1-9 should provide proof of feasibility not only for the selected 1270 D&S but also for the remaining 17,000 Latin D&S from the *Latinitum in aeternum* collection. Benko and Rausova (2020) warn of the major labor of recognizing Latin tokens in the corpus (so that other types of tokens do not sneak into the results, especially those that are not PUs), which is prevented here by selecting the largest available ready-made collection of traditional D&S and entering ready-made D&S into the search engine, rather than the search engine itself looking for potential PUs by searching the corpus for everything that is Latin. Benko and Rausova (2020) also warn of the problem of searching for Latin PUs in English corpora, where a large amount of Latin lexical material can confuse the search engine when automatically searching for Latin expressions. It is possible that choosing a ready-made collection as a source of D&S for input into search engines would be a solution for searching for Latin D&S in English or Romance languages, but then it is necessary to enter the D&S into the search engine manually one by one, as in this paper. Finally, when searching the HJR, points 6 and 7 must be applied, which refer to entering reversible binomials and examining familiar modifications when entering the required saying.

II RESULTS

In the Latin corpus analyzed in this paper, the most frequent saying is *Alter ego/Another self/Close friend* with 1767 tokens (168.29 per million words). This number of tokens is also the highest reference value found in this paper (see Table 1, Appendix). Since a large number of sayings in such a large corpus do not have a single token and their existence cannot be denied, 0 tokens are accepted as the lowest reference value. Apart from *Alter ego*, there is only one other saying with more than a thousand tokens, i.e., *Ad hoc/For this occasion/For this purpose*, with 1040 tokens. The sayings *Alibi* (876), *Ad modum* (719), *Ad oculos* (557), and *Ad curiam* (524) have between 500 and 1000 tokens, and a further 23 sayings in Table 1 have more than 100 tokens. The next 67 sayings have more than 10 and less than 100 tokens, and 406 sayings have at least one and no more than 10 tokens. Of the 1270 sayings beginning with the letter A that were examined, tokens were found for a total of 502 sayings, while no tokens were found for 768 sayings. The absence of at least one token for as many as 61% of the sayings beginning with the letter A from the *Latinum in aeternum* collection upholds the comprehensiveness of this collection.

If a saying is possible or known in several modifications, such possibilities were also analyzed and such subdivisions are statistically indicated in parentheses. For example, the original saying *Ad curiam* has 329 tokens, the modification *In curiam* has 175 tokens and the modification *Ante curiam* has 20 tokens. In the “Latin saying” column in Table 1 (Appendix), all three modifications are listed. Namely, the starting one begins with the letter A or is listed first in the *Latinum in aeternum* collection and it is followed by the other modifications found. For each modification, the number of tokens found is listed in parentheses, while the “The number of Latin tokens in the PDLGR” column indicates the sum of the tokens of all modifications. The example of *Ad curiam* has 524 tokens including the modifications mentioned above.

If the search engine returns many results, it is, unfortunately, necessary to carry out an individual review of each token. For example, *Agenda/What needs to be done*, although it can be an example of a monolexical PU (16), is not always used as a PU. It is also used as a simple gerundive (17) or a nominal predication in the nominal predicate of the passive periphrastic conjugation (18) without idiomatic meaning:

(16) *Eius agenda*/His plan.

(17) *Opera agenda*/Work to be done.

(18) *Multa cras agenda sunt*/Many things need to be done tomorrow.

This meticulous approach to searching for the frequency of Latin D&S also brings solutions to some age-old issues about the relevance of alternations, i.e., modifications of a single saying. For example, the corpus search engine resolves the age-old dispute over the saying *Ante portas/Ad portas*. The origin of the saying is traditionally associated with Hannibal's threat to Rome in the Second Punic War from the pen of Titus Livius and Marcus Tullius Cicero (*Hannibal ante portas* or *Hannibal ad portas*), and in the collection *Latinum in aeternum*, only *Ante portas* is mentioned under the letter A and it has 37 tokens in the corpus. Although it is not the subject of this study, it is worth looking under the letter H in *Latinum in aeternum*, where the entire saying should be found. Both modifications are listed under the letter H: *Hannibal ad portas* and *Hannibal ante portas*. This finding justifies the addition of *Ad portas* when examining the saying *Ante portas*, for which 161 tokens were found. The name Hannibal appears in 5 tokens exclusively in the form *Hannibal ad portas*. Finally, more than two modifications are included in the search (19). A closer examination reveals that Cicero uses the term *Hannibal ad portas* in the *Philippics* and then also in *De finibus bonorum et malorum*. *Hannibal ad portas* was also used by Titus Livius in *Ab Urbe condita*.

(19) (Hannibal) ante portas/(Hannibal) ad portas.

Of course, the dominance of one modification or its originality does not detract from the existence of another accepted and used modification (*Hannibal*) *ante portas*. It should also be noted that both *Ad portas* and *Ante portas* could exist in Classical Latin as permanent or ordinary prepositional phrases independently of the metaphorical association with imminent danger embodied by *Hannibal ad portas*.

When searching the Croatian corpus of the HJR, the Latin expression *Ad referendum/ On an opinion (to give an opinion!)/To consult* is by far the most represented with 1807 tokens (20.08 per million words), which is the highest reference value for D&S in the HJR. Even in the Croatian corpus, for the largest number of Latin D&S, not a single token was found, so the lowest reference value is 0 tokens. *Alibi/In another place* has more than 500 tokens (547). Only two sayings have between 100 and 500 tokens, i.e., *Ad hoc/For this/For this occasion/For this purpose* (290) and *A priori/A priore/In advance* (193). The following sayings have between 10 and 100 tokens: *Ad acta/In the files* (93), *Anno Domini/A.D./Year of the Lord* (93), *Agenda/What needs to be done* (60), *Alter ego/Another self/Close friend* (35), *Ad rem/Straight to the point. No hesitation.* (30), *(Argumentum) ad hominem/In a human way./According to human understanding/(Against man?)* (29), *Amicus curiae/Friend of the court* (20), *Apage, Satanas!/Get lost, Satan!/Go away, devil!* (17), *Ad astra/To the stars/Toward the light/To the heights* (15), *(Audiatur) et altera*

pars/Let the other side be heard (13), *Artes liberales/Liberal arts* (11) and *Antemurale Christianitatis/Bulwark of Christendom* (11). It is interesting to note that the famous *Ante portas/Ad portas/At the door* has only 10 tokens in the HJR, and belongs to the smallest group of sayings, those having no more than 10 tokens. To summarize, in the HJR there is a single saying with more than 1000 tokens, 1 saying in the group from 500 to 1000 tokens, 2 sayings in the group from 100 to 500 tokens, and 12 sayings occur with more than 10 and less than 100 tokens. In total, at least one token was found for 96 (7.56%) of the sayings searched for, while no tokens were found in the HJR for 1174 (92.44%) sayings. The results obtained refer to the original Latin D&S and not to translations or equivalents in the Croatian language. It turned out that some Latin D&S are more represented in the Croatian corpus than in the Latin one, in terms of the total number of tokens, but not necessarily in terms of frequency per million words, as the HJR corpus is almost 9 times larger (regarding the wordcount). Due to the much lower representation of Latin sayings in the HJR compared to the PDLGR, a specific count of individual modifications of the same saying was considered less relevant, and in Table 1, of all 1270 sayings, the number of tokens of individual modifications is shown only for the PDLGR. Translations and Croatian equivalents of sayings are not the subject of this paper but are cited alongside the original Latin sayings to understand their meaning.

A particularly interesting finding is that Latin sayings that are significantly more represented in the HJR than in the PDLGR, such as *Ad referendum*, or those that occur more often in the HJR but not at all in the PDLGR, such as *Ad acta*, *Anno Domini*, (*Argumentum*) *ad hominem*, *Amicus curiae* or *Apagē*, *Satanas*, and *Audiatur et altera pars*, which indicates that they were coined quite late in literary history, at least from a Eurocentric perspective, and are probably not older than a few centuries.

III DISCUSSION

In large corpora of modern languages, it is possible to run an automatic search for Latin phrases, even largely with the help of a search engine, and to separate ordinary Latin phrases or individual words from PUs or probable PUs, as shown in the methodology of Benko and Rausova (2020). The search engine of the PDLGR does not have such capabilities. In this paper, the classification of PUs according to the most detailed modern categorizations was not applied. The ready-made collection *Latinum in aeternum* is used as a source of Latin D&S, which may not contain some well-known or frequent Latin sayings.

On the other hand, the automatic search for Latin D&S can only return results if Latin PUs are searched for in the corpus of another language. Such a result holds for the frequency of Latin D&S in the target or recipient language, not in the Latin language. In other words, one object of the search in this paper is the corpus of the Latin language as the source language, and Latin PUs can only be searched for in it based on individual entries. The results of this paper do not diminish the importance of the frequency of Latin D&S in any modern recipient language for the selection of Latin phrases to be included in an educational program. Thus, this paper gives an initial value for each Latin phrase rather than a final value for all individual modern languages or individual registers, except for the Croatian language thanks to a parallel search of the HJR corpus. The detailed division of PUs into subtypes is not the focus of this paper, as it can be done later, as well as the ad hoc elimination of certain subtypes of phrases from the list, all according to the user's needs. The *Latinum in aeternum* collection is extremely extensive, and according to the results of this paper, it can be assumed that at least half of its occurrences are not represented in the PDLGR, indicating that the said collection is very thorough. The main objection to the methodology of the paper is certainly at times laborious manual search of the PDLGR collection, which is used as the largest digital corpus of Latin texts, which does not diminish the value of the findings, but increases the effort and time required to obtain the findings.

IV CONCLUSION

An intensive search of the PDLGR with Latin as the selected language domain and of the HJR has confirmed the presence of 1270 examples of Latin D&S. The examples are Latin phrases qualified as D&S in the collection *Latinum in aeternum*, all beginning with the letter A. It was found that of these 1270 Latin D&S, according to the number of tokens in the PDLGR corpus, the following 27 stand out: *Alter ego* (1767), *Ad hoc* (1040), *Alibi* (876), *Ad modum* (719), *Ad oculos/Ante oculos* (557), *Ad curiam/In curiam/Ante curiam* (524), *Ad populum* (399), *Ab Urbe condita/A.U.C* (395), *Ad arma* (331), *Ad rem* (308), *Argumentum e contrario/e contrario/(ex) contrario* (269), *Ad astra* (238), *A patre* (225), *Ad extremum* (217), *Ab initio* (203), *Aequo animo* (203), *Ante portas/Ad portas* (198), *A contrario/E contrario* (150), *Agenda* (143), *Ad summam* (137), *A parte* (134), *Aes alienum* (130), *Ad ultimum* (122), *Ante diem* (121), *Ad finem* (118), *A principio* (116), *Ad usum* (111), *A latere/Ab latere* (103). On the other hand, the following 16 Latin D&S stand out in the HJR: *Ad referendum* (1807), *Alibi* (547), *Ad hoc* (290), *A priori* (193), *Ad acta* (93), *Anno Domini/A.D.* (93), *Agenda* (60), *Alter ego* (35), *Ad rem*

(30), (*Argumentum*) *ad hominem* (29), *Amicus curiae* (20), *Apage, Satanas* (17), *Ad astra* (15), (*Audiatur*) *et altera pars* (13), *Artes liberales* (11) and *Antemurale Christianitatis* (11). An important finding of the paper is the necessary adaptation of the search methodology of the Latin corpus, which is a necessary combination of automatic machine search and manual search. Of the 1270 examples searched in the PDLGR, at least one token was found for 502 D&S, and no tokens were found for 768 D&S in the corpus with 10.5 million words. The statistics obtained are not definitive for some D&S, which may occur in additional modifications with a different initial letter, such as the example *Ad portas/Ante portas*, which must be added to the examples *Hannibal ad portas/Hannibal ante portas* in further research when the corpus search is carried out, and for other D&S from the collection *Latinum in aeternum*, which begin with other initial letters. This search in the Latin corpus provides a tool for the relevance of Latin D&S in Roman and Latin literature, allowing participants in educational processes to further filter D&S according to their subtypes, always taking into account their numerical presence in the corpus. The number of Latin sayings beginning with the letter A is significantly lower in the HJR than in the PDLGR, and even lower in the calculation per million words. A large number of sayings unattested in the corpus of Roman Latin texts strongly suggests their later medieval and modern origin and, with further investigation, could provide evidence for the vitality of the Latin language even after the extinction of the original speakers.

Keywords: frequency of Latin sayings, Latin phraseological units, dicta et sententiae, Latin phraseology, Latin in Croatian

APPENDIX

Comparative tabular presentation of the number of occurrences of Latin D&S starting with the letter A from the collection *Latinum in aeternum* in the Perseus Digital Library Greek and Roman Materials: Latin (PDLGR) corpus and in the corpus of the Croatian language repository of the Institute for Croatian Language and Linguistics (HJR) (Table 1):

Table 1: 203 most frequent S&D starting with the letter A according to the Perseus Digital Library Greek and Roman Materials: Latin (PDLGR) corpus

Frequency of latin phraseological units in the perseus digital library...

Rank	Latin saying	The number of Latin tokens in the PDLGR	Frequency of Latin tokens per million words in the PDLGR	The number of Latin tokens in the HJR	Frequency of Latin tokens per million words in the HJR
1	Alter ego/Another self/Close friend.	1767	168.29	35	0.39
2	Ad hoc/For this purpose.	1040	99.05	290	3.22
3	Alibi/In other place./Elsewhere.	876	83.43	547	6.08
4	Ad modum/Modeled after.	719	68.48	0	0.00
5	Ad oculos (25)/Ante oculos (532)/Before/To the eyes.	557	53.05	0	0.00
6	Ad curiam (329)/In curiam (175)/Ante curiam (20)/To the court.	524	49.90	0	0.00
7	Ad populum/To the people.	399	38.00	2	0.02
8	Ab Urbe condita (369)/Anno ab Urbe condita (26)/From the city having been founded.	395	37.62	0	0.00
9	A.U.C. (see Ab Urbe condita).	0	0	6	0.07
10	Ad arma/To arms.	331	31.52	0	0.00
11	Ad rem/Straight to the point. No hesitation.	308	29.33	30	0.33
12	Argumentum e contrario/e contrario (92)/(ex) contrario (177)/a contrario (4)/Proof by contradiction.	273	26.00	2	0.02
13	Ad astra/To the stars/To the light/To the heights.	238	22.67	15	0.17
14	A patre/From father/From the father's side.	225	21.43	0	0.00
15	Ad extremum/At the end/Until the end.	217	20.67	0	0.00

16	Ab initio/From the begining./ Again.	203	19.33	1	0.01
17	Aequo animo/With clear conscience.	203	19.33	0	0.00
18	Ante portas (37)/Ad portas (161)/At the door.	198	18.86	10	0.11
19	A contrario (3)/E contrario (147)/ On the contrary/From another point of view.	150	14.29	2	0.02
20	Agenda/What needs to be done.	143	13.62	60	0.67
21	Ad summam/Total/On the whole.	137	13.05	0	0.00
22	A parte/From the side/on the side/from one party.	134	12.76	1	0.01
23	Aes alienum/Someone else's money.	130	12.38	0	0.00
24	Ad ultimum/To the end.	122	11.62	0	0.00
25	Ante diem/Days before. Before the time.	121	11.52	0	0.00
26	Ad finem/Until the end.	118	11.24	0	0.00
27	A principio/From the start/ From the beginning/From the first steps.	116	11.05	0	0.00
28	Ad usum/ For use./According to custom.	111	10.57	6	0.07
29	A latere (44)/Ab latere (59)/ From the side.	103	9.81	4	0.04
30	A primo/ Beginning with the first/First.	92	8.76	0	0.00
31	Ab officio (24)/Ex officio (6)/De officio (48)/From the office.	78	7.43	2	0.02
32	Ad arbitrium/As you please.	76	7.24	0	0.00
33	Ad exemplum (61)/In exem- plum (13)/For example.	74	7.05	1	0.01
34	Ad unum omnes (67)/In unum omnes (1)/All to one./All without exception.	68	6.48	0	0.00

Frequency of latin phraseological units in the perseus digital library...

35	A matre/Mother's side/ From the mother's side/Mother's kinship.	60	5.71	0	0.00
36	Ad praesens/Today./At present.	54	5.14	0	0.00
37	Ad inferos/To the other world./To the dead.	53	5.05		0.00
38	Ad honores/In honor./Out of honor.	50	4.76	1	0.01
39	Ad nutum/At free will.	48	4.57	0	0.00
40	Ad speciem/For appearance.	48	4.57	0	0.00
41	Ab aerario/From the treasurer.	47	4.48	0	0.00
42	Ad gloriam/For glory.	47	4.48	4	0.04
43	Auctoritate sua/By his own authority.	46	4.38	0	0.00
44	Ad loca/To your seats./Sit down.	44	4.19	1	0.01
45	Ad propositum/To the purpose./For the purpose of.	42	4.00	0	0.00
46	Arma movere (37)/bellum (3) movere/Start war.	40	3.81	0	0.00
47	Ager publicus/Public property.	38	3.62	0	0.00
48	Altera pars/Other side/Other party/Opposite party.	36	3.43	13	0.14
49	Ad latus/To the side.	35	3.33	0	0.00
50	Alma parens/Mother provider.	35	3.33	0	0.00
51	Ad colloquium (20)/In colloquium (14)/For discussion./To communicate/To converse.	34	3.24	0	0.00
52	A proposito/According to the purpose/In accordance with the purpose.	33	3.14	0	0.00

53	Animus aequus optimum est aerumnae condimentum/ Indi- fference is the best condiment for worries.	32	3.05	0	0.00
54	Animo suo/By own intention./ By own desire.	30	2.86	0	0.00
55	Ab antiquo/From ancient times./From long ago.	28	2.67	0	0.00
56	Aliud quid/Something else./A completely different matter.	28	2.67	0	0.00
57	A die/From the day.	26	2.48	0	0.00
58	Ad verbum/Literally./ Word for word.	26	2.48	0	0.00
59	Anno ab Urbe condita/(see Ab Urbe condita)/Years from the city having been founded.	26	2.48	6	0.07
60	Amabo te/Please./I will be grateful.	25	2.38	0	0.00
61	Amor patriae/Love for the homeland./Patriotism.	24	2.29	0	0.00
62	Ab aeterno/Forever./ From the earliest times./For as long as can be remembered.	23	2.19	0	0.00
63	Alia ratione/In other way.	23	2.19	0	0.00
64	Ad voluntatem/As one wis- hes.	22	2.10	0	0.00
65	Arbitrium liberum/Free will./ Free choice.	22	2.10	0	0.00
66	Ad vocem/Regarding somet- hing.	21	2.00	4	0.04
67	Animo et corpore/With body and soul./With all strength.	21	2.00	0	0.00
68	Ad deliberandum/On reflecti- on./On decision.	20	1.90	0	0.00
69	Ad faciendum/Due to doing./ On execution.	20	1.90	0	0.00
70	Ad partem/With every part.	20	1.90	0	0.00

Frequency of latin phraseological units in the perseus digital library...

71	Apud iudicem/In front of the (appointed) judge.	20	1.90	0	0.00
72	Ab ovo (6)/Ex ovo (13)/ From the beginning./From the embryo.	19	1.81	6	0.07
73	Ad hominem/In a human way./According to human understanding/(Against man?)	19	1.81	29	0.32
74	Ab imo pectore/From the bottom of the heart./From the depth of the soul.	18	1.71	0	0.00
75	Ab cunabulis (4)/Ab incunabulis (14)/From the cradle.	18	1.71	0	0.00
76	Ad persequendum/Due to persecution./Due to conducting court proceedings.	18	1.71	0	0.00
77	Ad instar/After the manner of.	17	1.62	1	0.01
78	Ante meridiem/Before noon.	17	1.62	0	0.00
79	A priori (0)/A priore (17)/In advance.	17	1.62	193	2.14
80	Ars poetica/The art of poetry.	17	1.62	2	0.02
81	Ad arbitrium suum/Independently.	16	1.52	0	0.00
82	Ad mensuram/As needed./As measured.	16	1.52	0	0.00
83	Ad nihilum redigere (8)/In nihilum redigere (7)/To turn into nothing.	15	1.43	0	0.00
84	Apparatus belli/Means for waging war.	15	1.43	0	0.00
85	Ab incunabulis (see Ab cunabulis)/From the cradle.	14	1.33	0	0.00
86	A limine/From the threshold.	14	1.33	4	0.04
87	Ab illo tempore/From that time./From before./About the old days.	13	1.24	0	0.00
88	Ad bestias/Before the beasts.	13	1.24	0	0.00

89	Ad defendendum/For the sake of defense./For representation./For the purpose of protection.	13	1.24	0	0.00
90	Aurea sunt vere nunc saecula: plurimus auro venit Honos, auro conciliatur amor/Now is truly a golden age: the greatest honor goes to gold, love is gained with gold.	13	1.24	0	0.00
91	Ad fidem/On honesty/On word of honor.	12	1.14	0	0.00
92	Ab invito/Against the will./Unwillingly./Against the wishes./By force.	11	1.05	0	0.00
93	Ad legendum/For reading.	11	1.05	0	0.00
94	Alterum tantum/Twice as much.	11	1.05	0	0.00
95	A parvis/From the childhood.	11	1.05	0	0.00
96	A prima aetate/From the earliest youth./From a young age.	11	1.05	0	0.00
97	Abi in crucem/Go away./Scram.	10	0.95	0	0.00
98	Ad fontes/At the source./At sources.	10	0.95	1	0.01
99	Ad ostentationem (8)/In ostentationem (2)/In order to show off/In order to stand out.	10	0.95	0	0.00
100	Ad pacem inclinare/Incline to peace.	10	0.95	0	0.00
101	Amor sceleratus habendi/Notorious greed for acquisition.	10	0.95	0	0.00
102	Auri sacra fames/Accursed hunger for gold.	10	0.95	0	0.00
103	Ab oriente ad occidentem/From east to west.	9	0.86	0	0.00

Frequency of latin phraseological units in the perseus digital library...

104	Ad dies vitae/For life.	9	0.86	0	0.00
105	Ad personam/Only for one person./Only for the case in question.	9	0.86	9	0.10
106	Adversa fortuna/Misfortune	9	0.86	0	0.00
107	Alta mente repostum/Implanted deep into the soul./Cut deep into the mind.	9	0.86	0	0.00
108	Amor omnia vincit (5)/Omnia vincit amor (4)/Love conquers all.	9	0.86	0	0.00
109	Ad alium diem (0)/In alium diem (8)/On another day.	8	0.76	0	0.00
110	Ad altiora/Aim higher.	8	0.76	0	0.00
111	Ad communem utilitatem/For the common good.	8	0.76	0	0.00
112	Ad normam/According to the norm.	8	0.76	0	0.00
113	Ad sensum/According to...	8	0.76	0	0.00
114	Aequo animo (or: facile) pati/ Calmly (easily) put up,	8	0.76	0	0.00
115	Alia tempora (1)/Aliis temporibus (7)/At other times./On other occasions.	8	0.76	0	0.00
116	Ad minimum/At least./At the very least	7	0.67	0	0.00
117	Ad unguem/To the nail./Exactly./Precisely.	7	0.67	0	0.00
118	Aequum et bonum/Right and good.	7	0.67	0	0.00
119	A tenero/From childhood./From the early days.	7	0.67	0	0.00
120	Aurea aetas/Golden age.	7	0.67	4	0.04
121	Ab abrupto. (see Ex abrupto.)/ All of a sudden.	6	0.57	0	0.00
122	Abi in malam rem/Go to Hell.	6	0.57	0	0.00

123	Ad comparandum/To compare.	6	0.57	0	0.00
124	Ad concludendum/To conclude./For (final) conclusion./In order to make a definitive decision.	6	0.57	0	0.00
125	Ad disputandum/For discussion.	6	0.57	0	0.00
126	Adnuit et totum nutu tremefecit Olympum/(Thus spoke Jupiter) and nodded...mercifully; his nod shook Olympus.	6	0.57	0	0.00
127	Ad omnem eventum paratus/Ready for anything.	6	0.57	0	0.00
128	Alieno nomine/In someone else's name/Instead of another.	6	0.57	0	0.00
129	Altum mare/Wide sea.	6	0.57	0	0.00
130	Arrectis auribus/Ears pricked up.	6	0.57	0	0.00
131	Ars oratoria/Oratory skill.	6	0.57	0	0.00
132	Ab irato/In anger./Angrily.	5	0.48	0	0.00
133	Ad hunc locum/To this place./Here.	5	0.48	0	0.00
134	Ad litem/For a court hearing.	5	0.48	1	0.01
135	Ad oculum/Apparently.	5	0.48	1	0.01
136	Ad respondendum/To answer.	5	0.48	0	0.00
137	Ad satisfaciendum/For satisfaction./To satisfy.	5	0.48	0	0.00
138	Ad ultimum inopiae adducere/To force to the last resort.	5	0.48	0	0.00
139	Adversa secundis pensare/Repair troubles with successes.	5	0.48	0	0.00
140	Adverso rumore esse/Unpopular.	5	0.48	0	0.00

Frequency of latin phraseological units in the perseus digital library...

141	Aequa lanx/Equal weight./ Balance./Impartiality.	5	0.48	0	0.00
142	Age, quid agis/Finish what you've started.	5	0.48	0	0.00
143	Age, si quid agis/Don't get distracted.	5	0.48	0	0.00
144	Agitata res/Agitating matter.	5	0.48	0	0.00
145	Aliter dictum/Expressed in another way./In other words.	5	0.48	0	0.00
146	Alternatim/One after another./ Alternating.	5	0.48	0	0.00
147	Arrectae (sunt) horrore comae/The hair stood up in horror.	5	0.48	0	0.00
148	A(b) parvulis/From the young age.	4	0.38	0	0.00
149	Ad amussim/Perfectly./With precision.	4	0.38	0	0.00
150	Ad inquirendum/For exami- nation./For the purpose of investigation./To determine./ To investigate.	4	0.38	0	0.00
151	Ad marginem/On the margin of a book.	4	0.38	0	0.00
152	Ad notam/To the knowledge.	4	0.38	3	0.03
153	Ad omnia paratus/Ready for anything.	4	0.38	0	0.00
154	Ad suum arbitrium referre/ Judge according to yourself.	4	0.38	1	0.01
155	Ad ultimum supplicium progredi/To proceed to the last execution./To take one's own life.	4	0.38	0	0.00
156	Adversus hominem/Against man./To the detriment of people.	4	0.38	0	0.00

157	Ah, Corydon, Corydon, quae te dementia cepit?/Ah, Corydon, Corydon, what madness has befallen you.	4	0.38	0	0.00
158	Alea iacta est/The die has been cast.	4	0.38	1	0.01
159	Aliena appetens, suorum bonorum profusus (0)/Alieni appetens (adpetens) (4)/Longing for someone else's, squander one's own goods	4	0.38	0	0.00
160	Alieni appetens, sui profusus/ (see Aliena appetens)	4	0.38	0	0.00
161	Alieni iuris/Under someone else's dominion./Under someone else's rule.	4	0.38	0	0.00
162	Aliud clausum in pectore, aliud promptum in lingua habere/To have one thing locked up in the chest, another ready in the tongue.	4	0.38	0	0.00
163	Animus meminisse horret/The mind shudders to remember.	4	0.38	0	0.00
164	Aperto capite/With an open head.	4	0.38	0	0.00
165	Argumentum a contrario/ (added: a contrario). (see Argumentum e contrario)	4	0.38	2	0.02
166	Armata manu/Armed force./Power of arms.	4	0.38	0	0.00
167	At pulchrum est digito monstrari et dici hic est/It's nice when people point to you and say: That's him.	4	0.38	0	0.00
168	Aura popularis/Popular wind./Whim of the masses.	4	0.38	0	0.00
169	Abi atque abstine manum/Go and quit your job.	3	0.29	0	0.00

Frequency of latin phraseological units in the perseus digital library...

170	Abi, nuntia, inquit, Romanis coelestes ita velle, ut mea Roma caput orbis terrarum sit/Go, he said, and tell the Romans that the gods want my Rome to be the center of the world. (often quoted only: Caput orbis).	3	0.29	0	0.00
171	Ab invisio (invisis)/Without seeing./Without looking./To sign something believing it is good.	3	0.29	0	0.00
172	Acutae nares/Fine(refined) taste.	3	0.29	0	0.00
173	Adhuc sub iudice lis est/ The dispute is still before the judge./The dispute has not yet been resolved./ the case lies still unsolved in the hands of the judge.	3	0.29	0	0.00
174	Admodum tenui filo suspensum esse/To be hanging by a thread.	3	0.29	0	0.00
175	Ad pericula subeunda parati/ Prepared to take risks./Ready to take all the perils.	3	0.29	0	0.00
176	Ad referendum/On an opinion (to give an opinion!)/To consult.	3	0.29	1807	20.08
177	Ad respectum suarum quemque rerum vertere/To turn everything to the defense of one's own benefit.	3	0.29	0	0.00
178	Ad sanitatem redire/Come to your senses.	3	0.29	0	0.00
179	Ad vitam aeternam/For the eternal life.	3	0.29	1	0.01
180	Aere alieno premi/Be pressed by debts.	3	0.29	0	0.00
181	Aeterna nox/Eternal night (darkness).	3	0.29	0	0.00

182	Agentem te ratio ducat, non fortuna/Let reason, not luck, guide you in your actions.	3	0.29	0	0.00
183	Agnus Dei/Lamb of God.	3	0.29	9	0.10
184	Aio te, Aeacida, Romanos vincere posse/I say that you, Aeacides, can defeat the Romans./I tell you, Aeacides, Romans can defeat you.	3	0.29	0	0.00
185	Albus an ater sit, nescio/Whether he is white or black, I don't know.	3	0.29	0	0.00
186	Aliorum vitia cernere, oblivisci suorum/To see the faults of others, to forget one's own.	3	0.29	0	0.00
187	Aliquid capitalis noxae habere/Considering something as a mortal sin.	3	0.29	0	0.00
188	Aliquid in pectore claudere/To lock something in your heart.	3	0.29	0	0.00
189	Aliquid utilitate, non honestate ponderare/To judge something with regard to utility, not honesty.	3	0.29	0	0.00
190	Alitur vitium vivitque tegendo/Vice feeds and lives on concealment.	3	0.29	0	0.00
191	Alpha et omega/Alfa i omega)/The beginning and the end.	3	0.29	0	0.00
192	Amicitia nisi inter bonos esse non potest/Friendship cannot exist except between good people.	3	0.29	0	0.00
193	Amico amicus/Friend to friend.	3	0.29	0	0.00
194	Amor nummi/Love of money.	3	0.29	0	0.00

Frequency of latin phraseological units in the perseus digital library...

195	A nativitate/(left out:from birth etc.)	3	0.29	1	0.01
196	A negotiis publicis se remove- re/To retire from government service.	3	0.29	0	0.00
197	Animus purus/Pure heart./ Innocent heart.	3	0.29	0	0.00
198	Anno aetatis suae/ Years of my life.	3	0.29	0	0.00
199	A planta pedis usque ad verti- cem/(see from head to toe in other D&S)	3	0.29	0	0.00
200	Arcades ambo/Both Arca- dians.	3	0.29	0	0.00
201	Artes liberales/Liberal arts.	3	0.29	11	0.12
202	Audire per nebulam/To hear through the fog.	3	0.29	0	0.00
203	Aures habent et non audient/ They have ears but will not hear.	3	0.29	0	0.00

SOURCES AND DICTIONARIES

- Cicero, Marcus Tullius. *De finibus bonorum et malorum*. Perseus Digital Library Greek and Roman Materials. 2021. M. Tullius Cicero, de Finibus Bonorum et Malorum, Th. Schiche, Ed. Tufts University. <http://www.perseus.tufts.edu/hopper/text?doc=Perseus%3atext%3a2007.01.0036> (accessed: 15 January 2022)
- Cicero, Marcus Tullius. *In Marcum Antonium orationes Philippicae I,5*. Perseus Digital Library Greek and Roman Materials. 2021. M. Tullius Cicero, Philippics, Albert Clark, Albert Curtis Clark, Ed. Tufts University. <http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.02.0011:text=P-hil.:speech=1:chapter=5&highlight=hannibal%2Cportas> (accessed: 15 January 2022)
- Ispitni katalozi za državnu maturu 2015/2016. Ispitni katalog iz latinskog jezika – viša razina str. 30-33. Nacionalni centar za vanjsko vrednovanje obrazovanja. http://dokumenti.ncvvo.hr/Ispitni_katalozi_15-16/Hrvatski/LATA_IK_16.pdf (accessed: 28 October 2021)

- Ispitni katalozi za državnu maturu 2020/2021. Ispitni katalog iz latinskog jezika – viša razina str. 37-39. Nacionalni centar za vanjsko vrednovanje obrazovanja. <https://www.ncvvo.hr/wp-content/uploads/2020/09/LATIN-SKI-A-2021.pdf> (accessed: 28 October 2021)
- Ispitni katalozi za državnu maturu 2021/2022. Ispitni katalog iz latinskog jezika – viša razina str. 18-21. Nacionalni centar za vanjsko vrednovanje obrazovanja. <https://www.ncvvo.hr/wp-content/uploads/2021/09/LAT-2022.pdf> (accessed 9 April 2022)
- Korpus PDLGR: *Perseus Digital Library Greek and Roman Materials*. All search options. Ur. Crane, Gregory R. Tufts University. <http://www.perseus.tufts.edu/hopper/search> (accessed: 2 May 2021 – 26 February 2022)
- Livius, Titus. *Ab Urbe condita XXIII, 16*. Perseus Digital Library Greek and Roman Materials. 2021. Titus Livius (Livy), The History of Rome, Book 23. Robert Seymour Conway, Charles Flamstead Walters, Ed. Tufts University.
- <http://www.perseus.tufts.edu/hopper/text?doc=Perseus:text:1999.02.0161:book=23:chapter=16&highlight=hannibal%2Cportas> (accessed 15 January 2022)
- Marević, Jozo. 2000. *Latinsko-hrvatski enciklopedijski rječnik*. Marka. Velika Gorica
- Marević, Jozo. 2002. *Latinski zauvijek Latinum in aeternum*. Marka. Velika Gorica

LITERATURE

- Benko, V., & Rausova, K. 2020. Data-Driven Approach to Identification of Latin Phrases in Russian Web-Crawled Corpora. In IMS (CLCO) (pp. 11-20).
- Biber, Douglas, et al. 2000. "Longman grammar of spoken and written English."
- Everaert, Martin; Erik Jan van der Linden, Andre Schenk (eds.). 1995. *Idioms: Structural and Psychological Perspectives*. Hillsdale-New Jersey-Hove: Lawrence Erlbaum Associates Publishers.
- Грудева Е.В. 2014. Кодовые переключения с русского языка на латынь в русских текстах разного типа (на материале Национального корпуса русского языка) // Вестник Череповецкого государственного университета. № 4 (57). — С. 93—97.
Grudeva Ye.V. Kodovyye pereklyucheniya s russkogo yazyka na latyn' v russkikh tekstakh raznogo tipa (na materiale Natsional'nogo korpusa russkogo yazyka) // Vestnik Cherepovetskogo gosudarstvennogo universiteta. — 2014. — № 4 (57). — S. 93—97.

- Grudeva E.V. Prelazak koda s ruskog na latinski u ruskim tekstovima raznih vrsta (na materijalu Nacionalnog korpusa ruskog jezika) // Bilten Državnog sveučilišta Čerepovec. - 2014. - Broj 4 (57). - S. 93-97.
- Moon, R. 1998. Fixed Expressions and Idioms in English: A Corpus-based Approach. Oxford University Press.
 - Omazić, Marija. 2015. "Phraseology through the looking glass." Osijek: Josip Juraj Strossmayer University, Faculty of Humanities and Social Sciences.
 - Panou, D. 2017. A Critical Assessment of L1 Idiom Comprehension and Processing Models. Journal of Applied Linguistics and Language Research 4(3), 1-12.
 - Peters, Ann M. 1983. The Units of Language Acquisition. Cambridge University Press, Cambridge.
 - Tabossi, P., Fanari, R., & Wolf, K. 2009. Why are idioms recognized fast?. Memory & cognition 37(4), 529-540.
 - van der Linden, H. J. B. M. 1992. Idioms, non-literal language and knowledge representation. (ITK Research Memo). Institute for Language Technology and Artificial Intelligence, Tilburg University.
 - YegAnehjoo, M., & Thai, Y. N. 2012. Lexical access in production of idioms by proficient L2 learners. 3L, Language, Linguistics, Literature 18(1).

**FREKVENCIJA LATINSKIH FRAZEOLošKIH
JEDINICA U LATINSKOM I HRVATSKOM
PROBNA SONDA A**

Svjetski obrazovni sustavi koriste velik broj latinskih izreka i poslovice popularno zvanih *Dicta et sententiae* (nadalje: D&S) koje se uče nasumično, bez podjele na podvrste, a obično se vodi računa da su važne obrazovnom području određene škole. To može donekle odgovarati pravnim školama zbog formulaičnosti pravnog diskursa, dok je izbor izreka u gimnazijskim, medicinskim, filozofskim i općenito društveno-humanističkim školama prepušten subjektivnom odabiru učitelja. Najveća hrvatska zbirka latinskih D&S *Latium in aeternum* (LIA), Marević (2002) sadrži 18632 navoda, a u srednjim školama se u dvije do četiri godine učenja latinskog jezika uči 100-200 D&S. Katalog hrvatske državne mature iz latinskog jezika (2016-2021.) sadrži manje od 200 izreka. Zato je prijeko potrebno utvrditi učestalost latinskih D&S u što većem korpusu Rimske (i kasnije latinske) književnosti i izvršiti podjelu na podvrste dajući tako školskim sustavima diljem svijeta prijeko potreban katalog temeljen na frekvenciji D&S kao alat za sastavljanje nastavnih i ispitnih kataloga. Ručni unos izreka u korpusnu tražilicu i opsežnost zbirke LIA zahtijeva prevelik napor i vrijeme za jedan znanstveni rad. Zato se ovdje istražuje korpusna zastupljenost probne sonde od 1270 D&S iz zbirke LIA koje počinju slovom A u korpusu Perseus Digital Library Greek and Roman Materials i u korpusu Hrvatska jezična riznica Instituta za hrvatski jezik i jezikoslovlje. Rezultati pretrage korisni su za nastavu latinskog u svijetu te za nastavu latinskog i hrvatskog jezika u hrvatskim školama.

Ključne riječi: *učestalost latinskih izreka, latinske poslovice, dicta et sententiae, latinska frazeologija, latinski u hrvatskom*