DIGITAL CLASSICS ONLINE

Hypotheseis, a Database of Named Entities Surrounding Greek Rhetorical Exercises

Camillo Carlo Pellizzari di San Girolamo

Abstract: The rhetorical exercises written in Greek from the Hellenistic to the Byzantine age constitute a wide corpus of texts that in the last two centuries have been the subject of few studies in comparison with many other areas of ancient Greek literature. As a result, there is no comprehensive list of Greek *progymnasmata* and declamations nor an overall study of the topics they cover and of the Named Entities they mention (mythological characters, historical persons, places, events, etc.) has already been published. The relational database *Hypotheseis*[1], founded in March 2024, is a *Wikibase* instance that aims to fill this gap, providing an easy way to describe the Greek rhetorical exercises and the Named Entities to which they are connected as structured data, available in CC0 license; a SPARQL endpoint allows making sophisticated analyses on the collected data. This paper presents the data model of the database, the corpus of texts it intends to catalogue, the work done so far and what could be done in the future.

Progymnasmata and Declamations: an Overview

Definitions and Classification

Progymnasmata and declamations are two types of rhetorical exercises used in Greek-language rhetorical teaching and as self-standing literary works from the first centuries AD to the Palaeologan age (15th century).

A *progymnasma*[2] (προγύμνασμα, 'preliminary exercise') is an exercise that prepares to the declamation. Four ancient manuals about *progymnasmata* survive (Theon, Pseudo-Hermogenes, Aphthonius, Nicholas of Myra);[3] the last three agree in defining 14 types of *progymnasmata*, arranged in a standard order of increasing difficulty: fable, narration, anecdote (*chreia*),[4] maxim, refutation, confirmation, common-place, praise, invective, comparison, *ethopoeia*,[5] description, thesis, introduction of a law.[6]

1 https://hypotheseis.wikibase.cloud/ (last access 11.07.2025).

2 Cf. Berardi (2017); for an extended bibliography, Chiron (2017).

3 Editions of reference: for Theon, Patillon / Bolognesi (1997) with French translation; for Pseudo-Hermogenes and Aphthonius, Patillon (2008) with French translation; for Nicholas of Myra, Felten (1913). All the manuals are translated in English by Kennedy (2003), the first three are translated in Spanish by Reche Martínez (1991).

4 Cf. Hock / O'Neil (1986); Hock / O'Neil (2002); Hock (2012).

5 Cf. Amato / Schamp (2005).

6 The SPARQL query on *Hypotheseis* https://tinyurl.com/2xjd9wg3 (last access 11.07.2025) shows the 14 types with Aphthonius' definitions.

A declamation[7] (μελέτη, 'exercise') is a speech that is imagined to be pronounced by a fictional speaker in a given context. It is the last and most important part of the rhetorical education. Declamations are usually classified by scholars according to their topic: mythological declamations, historical declamations, and *plasmata* (i.e. declamations with stock characters set in an undetermined classicizing city).[8]

Lists

The main lists of surviving *progymnasmata* and declamations by epoch are the following:

- Hunger (1978), 92–120 for the *progymnasmata* and the declamations from the Late Antiquity to the Byzantine age;
- Russell (1983), 3–6 for the surviving declamations up to the Late Antiquity;
- Hock / O'Neil (2002) for all the known anecdotes (*chreiai*), from the Imperial age to the Byzantine age;
- Amato / Ventrella (2005) for all the known *ethopoeiae*, both in Greek and Latin, from the Hellenistic age to the Byzantine age;
- Guast (2023), 2–9 for the declamations of the first three centuries AD.

There are also lists of *progymnasmata* and declamations organized thematically:

- Jacobs (1899) for mythological themes in *progymnasmata*;
- Kohl (1915) for historical themes in Greek and Latin declamations;
- Ureña Bracero (1999), 324–329 – with additions in Ureña Bracero (2005), 95 n. 4 – for Homeric themes in *ethopoeiae*; see also Fernández Delgado (2025), 112–116;
- Gibson (2004) for historical themes in the manuals about *progymnasmata*.

Classification

Progymnasmata and declamations constitute a very diverse corpus, which can be sorted according to various parameters:

- form: whilst extant *progymnasmata* and declamations are mostly in prose, a certain number of *ethopoeiae* in verses survive, both as independent short works and as parts of longer poems, as well as other types of *progymnasmata*, like praises;[9]
- literariness: extant *progymnasmata* and declamations are mostly literary works; nonetheless, there are also some *progymnasmata* composed as school exercises, mainly thanks to papyrus discoveries in Egypt;[10]

7 Cf. Russell (1983), Guast (2023).

8 Sophistopolis is the well-known nickname attributed to it by Russell (1983), 21–39.

9 Cf. Cribiore (2001), 225–230; Agosti (2005).

10 Cf. Cribiore (1996), 259–262 (n° 344–357); Cribiore (2001), 221–230; Agosti (2005), 55; Amato / Ventrella (2005), 223–225.

- tradition: most extant *progymnasmata* and declamations survived in medieval manuscripts, but a relevant number of *progymnasmata* have been discovered in papyri (and tablets);[11]
- completeness: apart from *progymnasmata* and declamations which are entirely or fragmentarily extant, hundreds of themes of *progymnasmata* and declamations are mentioned in other works (rhetorical manuals, commentaries to rhetorical manuals, biographies of rhetoricians etc.).[12]

Other criteria of classification include genre, author, epoch, character(s), and theme(s).

The Database *Hypotheseis*

Reason and Aims

The idea of starting a database of *progymnasmata* and declamations originated from two main needs:

- firstly, a comprehensive overview of Greek rhetorical exercises could allow to understand more deeply these literary genres, their role in rhetorical teaching, their relationship with other genres and their chronological evolution; however, the existing lists are partial, each covering only a part of *progymnasmata* and declamations (defined by genre and/or epoch), and a systematic extraction of themes mentioned in other works is still missing;
- secondly, existing lists of Greek rhetorical exercises are textual publications, so they sort the entries according to one criterion; consequently, sorting the entries according to a different criterion requires a relevant amount of work, and analysing them according to multiple criteria combined together is basically impossible without a tabular reshaping of the data.

Thus, the main aims of the database are, in the long-term, providing a comprehensive database of *progymnasmata* and declamations, both extant and mentioned in other works, and making it possible to analyse them according to the multiple classification criteria previously outlined.

Hypotheseis also aims to make data about *progymnasmata* and declamations FAIR,[13] specifically in the following ways:

- providing both *progymnasmata* and declamations and the Named Entities to which they are connected (authors, characters, places, etc.) with PIDs;
- making the data retrievable in multiple ways (mainly through a web interface and a SPARQL endpoint);
- interlinking all the entities, wherever possible, with *Wikidata*, in order to make it easier to compare the two databases and to make federated SPARQL queries on the two databases;
- releasing the data in CC0 license,[14] in order to allow their widest possible reuse.

11 Cf. Morgan (1998), 198–226; Cribiore (2001), 225–230; Hock / O'Neil (2002), 5–48, 56–66, 94–97.

12 Cf. Jacobs (1899) and Kohl (1915), which extract themes not only from extant texts but also from these sources.

13 Cf. https://www.go-fair.org/fair-principles/ (last access 11.07.2025).

14 https://creativecommons.org/public-domain/cc0/ (last access 11.07.2025).

Software Choice: *Wikibase*

The software chosen to create the database *Hypotheseis* is *Wikibase*,[15] a set of *MediaWiki* extensions allowing collecting structured data. *Hypotheseis* can thus be defined as a *Wikibase* instance, like *Wikidata*. The main reasons of this choice are the following:

- *MediaWiki* and its extensions, including *Wikibase*, are open source software, and they are regularly maintained[16] by a wide community of volunteers, by WMDE, which developed *Wikibase* (and *Wikidata*) from their launch[17] and has currently a dedicated development plan for *Wikibase* (and *Wikidata*),[18] and by the WMF;

- a *Wikibase* instance can have a SPARQL endpoint[19] allowing to query data, besides offering also other ways to retrieve them (the web interface, the special page EntityData,[20] the *MediaWiki* Action API,[21] and the REST API[22]); it is easy to make federated queries with *Wikidata*,[23] provided that the entities described in *Hypotheseis* declare their equivalents in *Wikidata*,[24] and also with other *Wikibase* (especially *Wikibase Cloud*) instances;

- the user interface is intuitive, which makes potentially easy to involve in the process of data curation scholars with little previous experience in digital humanities and databases; additionally, two user-friendly and open source tools are available to perform massive imports of data, i.e. *QuickStatements*[25] and the software *OpenRefine*[26], and it is also possible for more advanced users to program bots in Python using the *Pywikibot* library;[27]

- all edits are tracked in the page histories, so it is easy to reconstruct both the evolution of pages and the contributions of each editor, potentially also creating statistics about them; moreover, mistakes can easily be undone and no damage is permanent (at least considering damages made through the user interface), since also page deletions can be quickly reverted;

- data can be entered in hundreds of languages,[28] thus constructing a multilingual database which could potentially be a good hub for collaboration between scholars from different linguistic backgrounds;

---

15 Cf. https://wikiba.se/ (last access 11.07.2025) and https://www.mediawiki.org/wiki/Wikibase (last access 11.07.2025).

16 Cf. https://www.mediawiki.org/wiki/Version_lifecycle (last access 11.07.2025).

17 Cf. https://meta.wikimedia.org/wiki/Wikidata/Technical_proposal (last access 11.07.2025).

18 https://www.wikidata.org/wiki/Wikidata:Development_plan (last access 11.07.2025).

19 https://hypotheseis.wikibase.cloud/query/ (last access 11.07.2025).

20 https://hypotheseis.wikibase.cloud/wiki/Special:EntityData (last access 11.07.2025).

21 https://hypotheseis.wikibase.cloud/w/api.php (last access 11.07.2025).

22 Cf. https://doc.wikimedia.org/Wikibase/master/php/repo_rest-api_README.html (last access 11.07.2025).

23 https://www.wikidata.org/wiki/Wikidata:SPARQL_query_service/Federated_queries (last access 11.07.2025).

24 The equivalence with *Wikidata* is stated through the properties "Wikidata property" (P1) for properties and "Wikidata item" (P2) for items.

25 https://hypotheseis.wikibase.cloud/tools/quickstatements/ (last access 11.07.2025).

26 https://openrefine.org/ (last access 11.07.2025).

27 Cf. https://www.mediawiki.org/wiki/Manual:Pywikibot (last access 11.07.2025).

28 Specifically, in all the languages used in Wikimedia wikis (https://meta.wikimedia.org/wiki/Special:SiteMatrix [last access 11.07.2025]) and, more generally, in all languages which have a Wikimedia language code (https://www.wikidata.org/wiki/Help:Wikimedia_language_codes/lists/all [last access 11.07.2025]).

- the properties used to construct statements are not predefined, but created by users according to their needs, so the data model is highly flexible (cf. below *Data model*);

- all properties can have more than one value on the same item, also when it would be logically or empirically wrong (e.g. two birth dates), because of the high flexibility of the software. Consequently, it is possible to store redundant and obsolete values, besides the best ones, whenever needed (e.g. when important sources have made influential mistakes or when, despite their obsolescence, it is anyway useful to store them for historical purposes). In order to mark the quality of concurring values, it is possible to use ranks,[29] specifically marking the best value(s) as preferred and wrong value(s) as deprecated, whilst the standard rank is normal; SPARQL queries allow selecting only 'truthy' statements (i.e., for each property, all the preferred-ranked values, or, if absent, all the normal-ranked values), or all statements, or only the statements with a specific rank;

- statements can be provided with qualifiers[30] and references:[31] qualifiers are used to provide context on statements, enabling the expression of nuances and details; references provide sources for the data, and the possibility of associating them with each statement, rather than with the entire entity, makes them much easier to verify.

Presently, the two ways available to create a *Wikibase* instance are *Wikibase Suite* and *Wikibase Cloud*:

- *Wikibase Suite*[32] consists of the possibility of downloading *Wikibase* and managing it independently to create a self-hosted *Wikibase* instance; it allows high margins of personalisation (e.g. through the installation of extensions[33]), but it also implies that the creator of the instance is fully in charge of its maintenance, both technically (i.e. software updates) and financially (i.e. server cost);

- *Wikibase Cloud*[34] is a platform hosted by WMDE since 2022,[35] where each user can freely create up to six cloud-hosted *Wikibase* instances; thus, WMDE manages the technical and financial costs of maintaining these instances; the main drawback, in comparison with *Wikibase Suite*, is the limited degree of possible personalisation for each instance.

According to data extracted from *Wikibase World*,[36] a *Wikibase* instance designed to be a census on a voluntary basis of *Wikibase* instances, as of February 16th 2025 the number of existing *Wikibase* instances with properties was 777, most of them (711) hosted in *Wikibase Cloud*.[37]

The growing interest for *Wikibase* in the field of Digital Humanities is proven also by the establishment in September 2024 of the DHwiki Working Group inside DARIAH-EU, which "is a space for

29 Cf. https://www.wikidata.org/wiki/Help:Ranking (last access 11.07.2025).

30 Cf. https://www.wikidata.org/wiki/Help:Qualifiers (last access 11.07.2025).

31 Cf. https://www.wikidata.org/wiki/Help:Sources (last access 11.07.2025).

32 https://www.mediawiki.org/wiki/Wikibase/Docker (last access 11.07.2025).

33 For the list of extensions used in *Hypotheseis* (and in all *Wikibase Cloud* instances), cf. https://hypotheseis.wikibase.cloud/wiki/Special:Version (last access 11.07.2025).

34 https://www.wikibase.cloud/ (last access 11.07.2025).

35 For historical remarks, cf. https://addshore.com/2024/09/2-years-of-wikibase-cloud-by-wmde/ (last access 11.07.2025).

36 https://wikibase.world/ (last access 11.07.2025).

37 Cf. https://addshore.com/2025/02/visualizing-wikibase-ecosystem-using-wikibase-world/ (last access 11.07.2025).

discussion and dissemination of use case experiences and best practices around *Wikibase* and *Wikidata* in a DH context, and for contributions to further developing the *Wikibase* software and related tools".[38]

Hypotheseis was created by the author of this paper on March 27th 2024 using *Wikibase Cloud*[39], since it did not rely on any financing, and the limited personalisation options were considered a minor issue. Its data are released in CC0 license.[40]

Data Entry and Statistics

As of March 13th 2025, the database *Hypotheseis* has been edited only by the author of this paper, with the username Epìdosis;[41] the total number of registered users is five.[42] According to the automatic statistics,[43] 11440 edits have been made and the number of pages is 1753, including 48 properties[44] and 1636 items.[45] The total number of triples is 63832 and they can all be downloaded from the SPARQL endpoint.[46] The most significant steps in data entry, and in general in the evolution of the database, are recorded in Italian in the page "Cronistoria".[47]

Most data have been added using the aforementioned tool *QuickStatements*, which allows making massive additions of data previously collected in a CSV file. Specifically, *QuickStatements* has been used to import massively most of the *hypotheseis*, providing them with basic data; the "tagging" with "character(s)" (P25), "object(s)/concept(s)" (P35), "action(s)/state(s)" (P26), "place" (P27), and "time" (P28) has been done manually in the great majority of cases. Labels have been added systematically in Italian first; since April 2024, they have also been added in English. Most entities (excluding *hypotheseis*, works, and expressions) also have labels in Ancient Greek.

The order in which statements appear in the items is not casual, nor it is based on the order of addition, but it relies on the order of properties established through the page "MediaWiki:Wikibase-SortedProperties", through which it can also be quickly modified (the properties not listed in that page appear after the ones listed in the page, in order of addition).[48] Conversely, qualifiers and references presently appear, for each statement, in the order in which they have been added.[49]

For 523 items the equivalent *Wikidata* item is specified through "Wikidata item" (P2);[50] they represent the majority of items, excluding *hypotheseis* (because they would not fall into *Wikidata*'s notability

38 Cf. https://www.dariah.eu/activities/working-groups/dhwiki/ (last access 11.07.2025) and https://dhwiki.wikibase.cloud/ (last access 11.07.2025).

39 https://hypotheseis.wikibase.cloud/ (last access 11.07.2025).

40 https://hypotheseis.wikibase.cloud/wiki/Project:About (last access 11.07.2025).

41 Cf. userpage: https://hypotheseis.wikibase.cloud/wiki/User:Epìdosis (last access 11.07.2025).

42 Cf. https://hypotheseis.wikibase.cloud/wiki/Special:ListUsers (last access 11.07.2025). New accounts can be requested using https://hypotheseis.wikibase.cloud/wiki/Special:RequestAccount (last access 11.07.2025).

43 https://hypotheseis.wikibase.cloud/wiki/Special:Statistics (last access 11.07.2025).

44 https://hypotheseis.wikibase.cloud/wiki/Special:ListProperties (last access 11.07.2025).

45 https://hypotheseis.wikibase.cloud/wiki/Special:AllPages?namespace=120 (last access 11.07.2025).

46 Cf. SPARQL query: https://tinyurl.com/2xq49u2l (last access 11.07.2025).

47 https://hypotheseis.wikibase.cloud/wiki/Cronistoria (last access 11.07.2025).

48 https://hypotheseis.wikibase.cloud/wiki/MediaWiki:Wikibase-SortedProperties (last access 11.07.2025).

49 Cf. request to order also qualifiers and references in the same way as main statements: https://phabricator.wikimedia.org/T169960 (last access 11.07.2025).

50 Cf. SPARQL query: https://tinyurl.com/24m56wup (last access 11.07.2025).

policy[51]) and expressions (because they are not currently recognised in *Wikidata*'s data model for books, which collapses the two concepts of "expression" and "manifestation" into "edition"[52]).

The works already consulted to extract *hypotheseis* are listed in the page "Opere".[53] As of March 2025, all the extant *progymnasmata* and declamations (including four still unpublished[54]) known to the author of this paper have been added to *Hypotheseis*, with two exceptions: *progymnasmata* in verses and *progymnasmata* discovered in papyri.

Data Model

The definition of the data model was the initial focus of the editing activity in *Hypotheseis*[55] and is presently defined in Italian in the page "Modello dei dati".[56]

The basic unit of the data model is the ὑπόθεσις (*hypothesis*) of a *progymnasma* or a declamation, i.e. the topic of the rhetorical exercise; it is used as a title in most extant *progymnasmata* and declamations,[57] although it is sometimes absent.[58] The choice of using the *hypothesis* as basic unit, instead of the entire *progymnasma* or declamation, is motivated by the attempt to model similarly extant *progymnasmata* and declamations and the ones whose themes are only mentioned in other works, for which therefore only the *hypothesis* is known.[59]

An item about a *hypothesis* has mainly the following statements:

- "instance of" (P4) with one of the four recursive subclasses of "hypothesis" (Q2) as value;[60]
- "text of the hypothesis" (P6) with the ancient Greek text of the *hypothesis* as value; the edition used to extract the text is specified through the qualifier "transcribed from the expression" (P7); it is possible to add multiple values, extracting the *hypothesis* from different editions and marking the value extracted from the reference edition with preferred rank;[61]
- "taken from" (P10) with the work from which the *hypothesis* is taken as value; the qualifiers "citation (hypothesis only)" (P11) and "citation (entire *progymnasma*/declamation)" (P36),

51 https://www.wikidata.org/wiki/Wikidata:Notability (last access 11.07.2025).

52 https://www.wikidata.org/wiki/Wikidata:WikiProject_Books (last access 11.07.2025).

53 https://hypotheseis.wikibase.cloud/wiki/Opere (last access 11.07.2025).

54 Two *progymnasmata* of Konstantinos Akropolites (https://hypotheseis.wikibase.cloud/entity/Q792 [last access 11.07.2025] = https://pinakes.irht.cnrs.fr/notices/oeuvre/16480/ [last access 11.07.2025] and https://hypotheseis.wikibase.cloud/entity/Q793 [last access 11.07.2025] = https://pinakes.irht.cnrs.fr/notices/oeuvre/20274/ [last access 11.07.2025]) and two declamations of Thomas Magistros (https://hypotheseis.wikibase.cloud/entity/Q1461 [last access 11.07.2025] = https://pinakes.irht.cnrs.fr/notices/oeuvre/75/ [last access 11.07.2025] and https://hypotheseis.wikibase.cloud/entity/Q1462 [last access 11.07.2025] = https://pinakes.irht.cnrs.fr/notices/oeuvre/7664/ [last access 11.07.2025]). I am currently working on the critical edition and translation of the two declamations of Thomas Magistros.

55 Cf. first edit: https://hypotheseis.wikibase.cloud/w/index.php?title=Pagina_principale&oldid=1 (last access 11.07.2025).

56 https://hypotheseis.wikibase.cloud/wiki/Modello_dei_dati (last access 11.07.2025).

57 E.g. Lib. Eth. 1 (https://hypotheseis.wikibase.cloud/entity/Q453 [last access 11.07.2025]).

58 E.g. G.Kyp. Prog. 1 (https://hypotheseis.wikibase.cloud/entity/Q729 [last access 11.07.2025]).

59 E.g. Aphth. Prog. p. 35 Rabe = 11.2 Patillon (https://hypotheseis.wikibase.cloud/entity/Q1213 [last access 11.07.2025]).

60 Cf. SPARQL query presenting a scheme of the types of *hypothesis*: https://tinyurl.com/29onqdo5 (last access 11.07.2025).

61 E.g. Aphth. Fab. 8 Hausrath – Hunger = Ps.Nicol. Fab. 1 Walz (https://hypotheseis.wikibase.cloud/entity/Q133 [last access 11.07.2025]).

the second one used only for extant *progymnasmata* and declamations, specify as a string of text the passage of the work where the *hypothesis* is found;

- "literary genre of the hypothesis" (P13) with the literary genre of the *hypothesis* as value; "literary subgenre of the hypothesis" (P47) can also be added with one or more values;[62]

- "character(s)" (P25), "object(s)/concept(s)" (P35), "action(s)/state(s)" (P26), "place" (P27), and "time" (P28) with one or more items as values, to tag the *hypothesis* by topic; as specified in references through the property "determination method" (P34), most of these tags have been assigned on the basis of the *hypothesis* (Q83, "inferred from the hypothesis"); the qualifier "role" (P24) allows giving further details about the role of the value of the statement, e.g. "fictional speaker" (Q101) for characters and "setting place" (Q1344) for places;[63]

- "same hypothesis as" (P19), "comparable with" (P20) and "opposite of" (P30) with one or more *hypotheseis* as values, to connect *hypotheseis* that are thematically related.

The Italian and English label of the *hypotheseis* contain the abbreviated bibliographical reference to the *hypothesis*; the Italian aliases also include other possible bibliographical references to the *hypothesis* (especially in cases of reattributed texts) and a brief summary of the *hypothesis* in Italian.[64]

Works have "instance of" (P4) "work" (Q3), are connected to their author through "author" (P21) and always have one or more abbreviations associated through "abbreviation" (P31). The abbreviations are usually derived from a reference work, specified through the qualifier "reference work" (P17); however, the abbreviation is created *ex novo* if absent in the reference works considered (LSJ and LBG, consulted in their TLG digitisations),[65] and it is readapted if stylistically incoherent with the other abbreviations in the use of dots.[66]

Editions have "instance of" (P4) "expression" (Q4), are connected to the corresponding work through "expression of" (P23) and can have a "transcription URL" (P9: used to link to a transcription of the edition in the TLG) and/or a "digitization URL" (P32: used to link to a downloadable PDF file of the edition in Internet Archive or Google Books).[67] The choice of modelling editions using the term "expression" is based on its definition as given in the conceptual model IFLA LRM (s.v. Expression, LRM-E3):[68]

> "An *expression* is the specific intellectual or artistic form that a work takes each time it is 'realized'. *Expression* encompasses, for example, the specific words, sentences, paragraphs, etc. that result from the realization of a *work* in the form of a text, or the particular sounds, phrasing, etc. resulting from the realization of a musical work. The boundaries of the entity *expression* are defined, however, so as to exclude incidental aspects of physical form, such as typeface and page layout for a text, unless, due to the nature of the work, these are integral to the intellectual or artistic realization of the work as such."

62 E.g. Lib. Decl. 26 (https://hypotheseis.wikibase.cloud/entity/Q1393 [last access 11.07.2025]).

63 E.g. Lib. Decl. 18 (https://hypotheseis.wikibase.cloud/entity/Q1385 [last access 11.07.2025]).

64 E.g. Sev. Eth. 10 Amato = Lib. Eth. 26 (https://hypotheseis.wikibase.cloud/entity/Q250 [last access 11.07.2025]).

65 E.g. G.Kyp. Decl.av. (https://hypotheseis.wikibase.cloud/entity/Q1253 [last access 11.07.2025]).

66 E.g. Pach. Decl. (https://hypotheseis.wikibase.cloud/entity/Q1250 [last access 11.07.2025]).

67 E.g. the text of Pach. Decl. established by Boissonade (https://hypotheseis.wikibase.cloud/entity/Q1280 [last access 11.07.2025]).

68 IFLA LRM 2017, 23.

Editions of rhetorical works, in *Hypotheseis*, are considered as realizations of the respective works in a concatenation of "specific words, sentences, paragraphs" as defined by the critical editor(s), but excluding the "incidental aspects of physical form", since it does not differentiate between expressions as materialized in their printed manifestation and in their different digital manifestations.

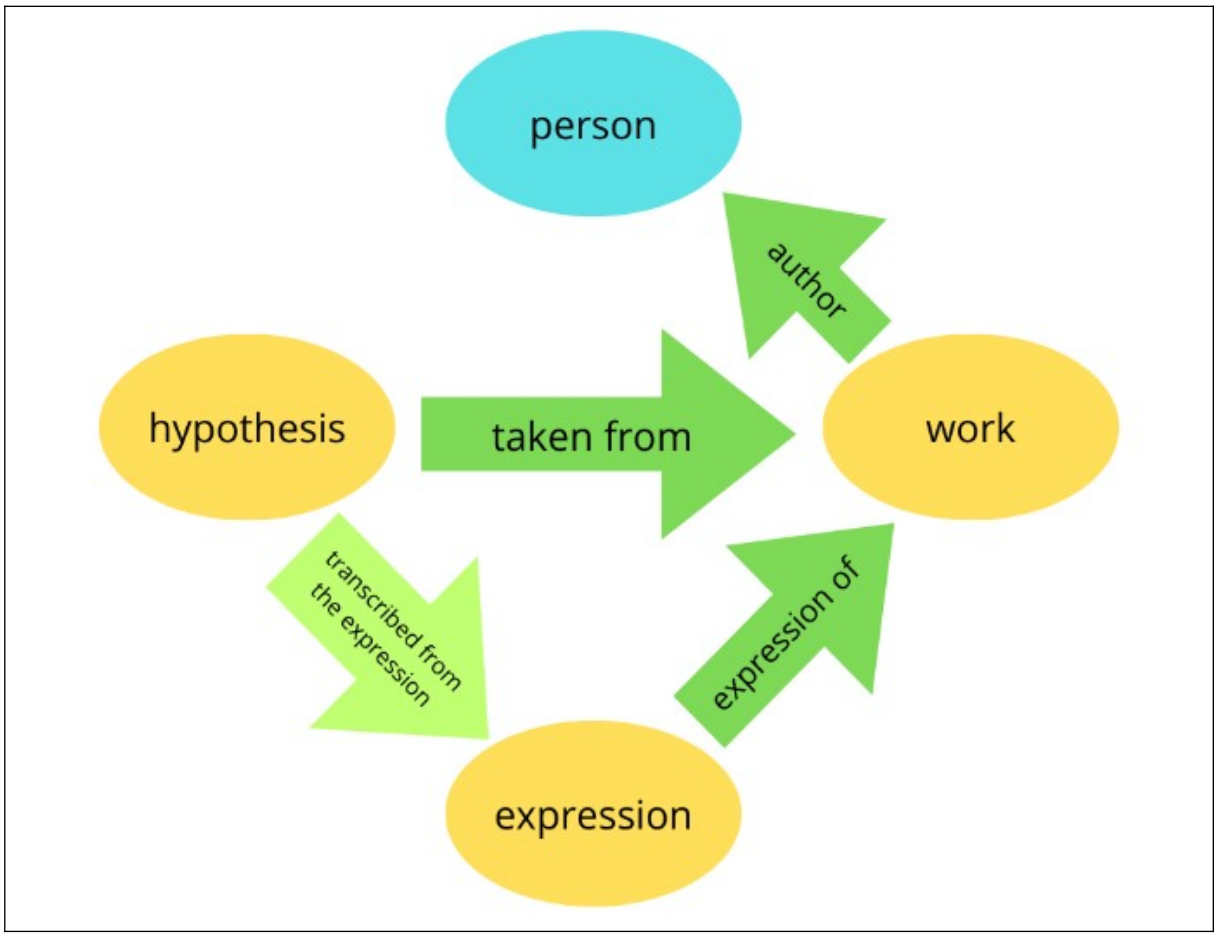

Fig. 1: Schema of the relations between *hypothesis*, work, expression and author.

For all other types of entities, statements are mostly limited to "instance of" (P4) and/or "subclass of" (P5), which are used as in *Wikidata*:[69]

- "instance of" means that the subject of the item is an individual appertaining to the class defined by the object; e.g. "war hero" (Q998) is an instance of "simple human type" (Q59) and "son-war hero" (Q1506) is an instance of "composite human type" (Q591);

- "subclass of" means that the subject of the item is a class appertaining to the class defined by the object; e.g. "son-war hero" (Q1506) is a subclass of "son" (Q1330) and of "war hero" (Q998).

The use of "subclass of" allows the inference of which items appertain, directly or recursively, to a certain class, e.g. all types of birds;[70] it is possible to use these inferences to group *hypotheseis* according to certain criteria, e.g. all *hypotheseis* mentioning birds.[71]

On authors, the use of "epoch of the author" (P48) allows a classification by epoch, and consequently also to filter works and *hypotheseis* by epoch; for the 34 authors added to the database so far, the four

69 Cf. https://www.wikidata.org/wiki/Help:Basic_membership_properties (last access 11.07.2025).

70 Cf. SPARQL query: https://tinyurl.com/243kfqj5 (last access 11.07.2025).

71 Cf. SPARQL query: https://tinyurl.com/2xmdgtue (last access 11.07.2025).

epochs currently in use are “imperial age” (Q1676), “late antiquity” (Q1677), “middle-Byzantine age” (Q1678), and “late-Byzantine age” (Q1679).[72]

Data Analysis and Visualisation

As previously said, the main – although not sole – way to query the structured data in a *Wikibase* instance is its SPARQL endpoint. In order to use the SPARQL endpoint of a *Wikibase Cloud* instance, it is necessary to explicitly define the prefixes used at the start of each query, or to use the full URIs, which is very inconvenient for readability.[73] A collection of precompiled SPARQL queries on *Hypotheseis*, and a list of all usable prefixes, is available in the page “Query” with Italian titles.[74] Results mentioned in this paragraph have been obtained running queries on March 14th 2025.

The main aim of queries on *Hypotheseis* is making analyses of the structured data. A first possible type of analysis is creating lists of entities:

- using one criterion, e.g. all extant *progymnasmata* (451 results)[75] and all extant declamations (120 results);[76]
- using two criteria, e.g. all extant *progymnasmata* and declamations mentioning Achilles (24 results)[77] or war heroes (10 results);[78]
- using more criteria, e.g. all extant late antique *progymnasmata* and declamations regarding the Trojan war (51 results)[79] or all the extant Byzantine *progymnasmata* and declamations with Biblical characters (33 results).[80]

A second possible type of analysis is making statistics on entities; these statistics can be visualised both as tables and as graphs inside the SPARQL endpoint, or data can be exported and used to create different visualisations with external tools. Some meaningful statistics can already be drawn from the entered data about extant *progymnasmata* and declamations (for the completeness of these data, cf. above *Data entry and statistics*), dividing them by author, by epoch and by genre:

- by author: considering *progymnasmata*,[81] the biggest corpus is Libanius (144), followed by Pseudo-Nicholas (111), Nikephoros Basilakes (56), and Aphthonius (54), with a total of 18 authors plus anonymous texts;[82] considering declamations,[83] the biggest corpus is again

72 Cf. SPARQL query: https://tinyurl.com/26ncftsb (last access 11.07.2025).

73 Cf. request to allow defining a complete list of prefixes for SPARQL queries from the dashboard of each *Wikibase Cloud* instance: https://phabricator.wikimedia.org/T335448 (last access 11.07.2025).

74 https://hypotheseis.wikibase.cloud/wiki/Query (last access 11.07.2025).

75 SPARQL query: https://tinyurl.com/2xs5h4gm (last access 11.07.2025).

76 SPARQL query: https://tinyurl.com/238bo7fo (last access 11.07.2025).

77 SPARQL query: https://tinyurl.com/2dafhs45 (last access 11.07.2025).

78 SPARQL query: https://tinyurl.com/24p4t9sp (last access 11.07.2025).

79 SPARQL query: https://tinyurl.com/2bs6bvhk (last access 11.07.2025).

80 SPARQL query: https://tinyurl.com/2dq5qe8w (last access 11.07.2025).

81 Cf. SPARQL query: https://tinyurl.com/237fl5ho (last access 11.07.2025).

82 It should be considered that 18 *progymnasmata* attributed to both Libanius and Pseudo-Nicolas are counted for both authors, as well as 12 *progymnasmata* attributed to both Aphthonius and Pseudo-Nicolas, and 2 *progymnasmata* attributed to both Severus of Alexandria and Libanius; cf. SPARQL query for a list of the 32 double attributions (https://tinyurl.com/24k5hfm9, [last access 11.07.2025]).

83 Cf. SPARQL query: https://tinyurl.com/25gqshu2 (last access 11.07.2025).

Libanius (51), followed by Georgios Pachymeres (13), Choricius and Aelius Aristides (12), with a total of 16 authors; these data can also be visualised as bubble charts;

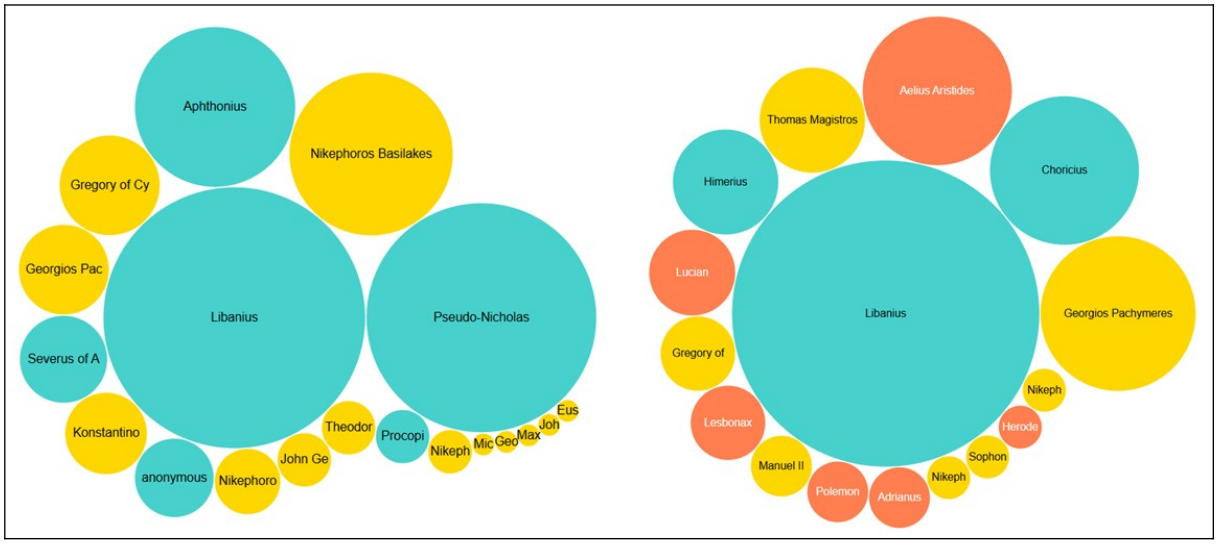

Fig. 2: Bubble charts representing the authors of extant *progymnasmata* (on the left) and of extant declamations (on the right) by number of texts and coloured by epoch.

- by epoch: for both extant *progymnasmata*[84] and extant declamations[85] there is a significant prevalence of late antique texts (313 out of 451, i.e. 69.4%, for *progymnasmata*, and 69 out of 120, i.e. 57.5% for declamations), followed by late-Byzantine texts (73, i.e. 16.2% for *progymnasmata*, and 26, i.e. 21.7%, for declamations); the share of middle-Byzantine texts is relevant for *progymnasmata* (65, i.e. 14.4%) but nearly non-existent for declamations (1, i.e. 0.8%); imperial age declamations complete the overview (24, i.e. 20%);

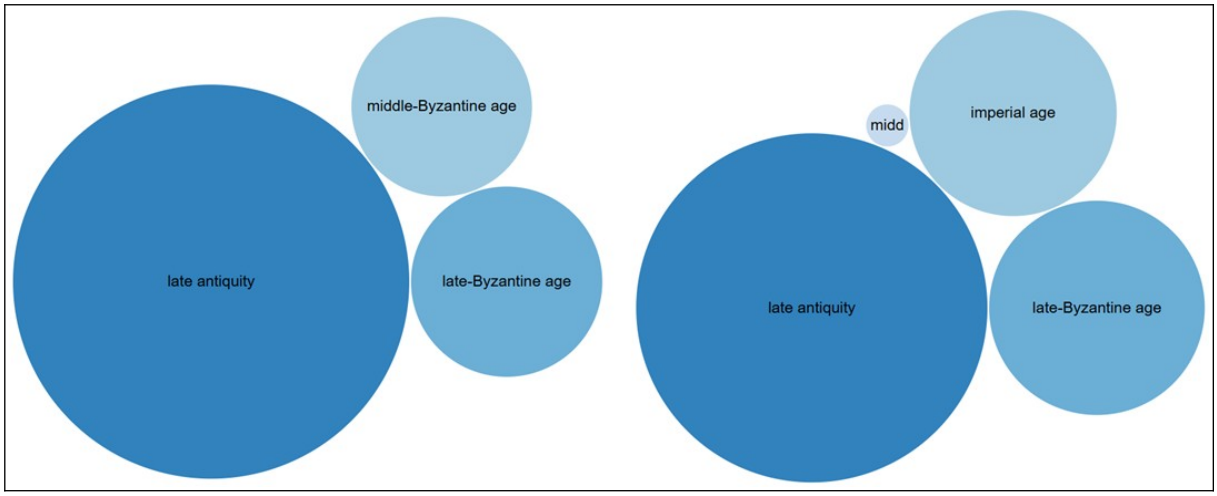

Fig. 3: Bubble charts representing the epochs of extant *progymnasmata* (on the left) and of extant declamations (on the right) by number of texts; the colours are the default ones.

- by genre: for *progymnasmata*,[86] the most common genres for extant texts are by far *ethopoeia* (94 out of 451, i.e. 20.8%), narration (86, i.e. 19.1%), and fable (81, i.e. 18.0%), whilst the least common are thesis (7, i.e. 1.6%) and introduction of a law (4, i.e. 0.9%); for de-

84 Cf. SPARQL query: https://tinyurl.com/28j3tus7 (last access 11.07.2025).

85 Cf. SPARQL query: https://tinyurl.com/26lho4aa (last access 11.07.2025).

86 Cf. SPARQL query: https://tinyurl.com/299svjxs (last access 11.07.2025).

clamations,[87] historical (57, i.e. 47.5%) and stock characters (52, i.e. 43.3%) declamations are far more common than mythological ones (11, i.e. 9.2%).

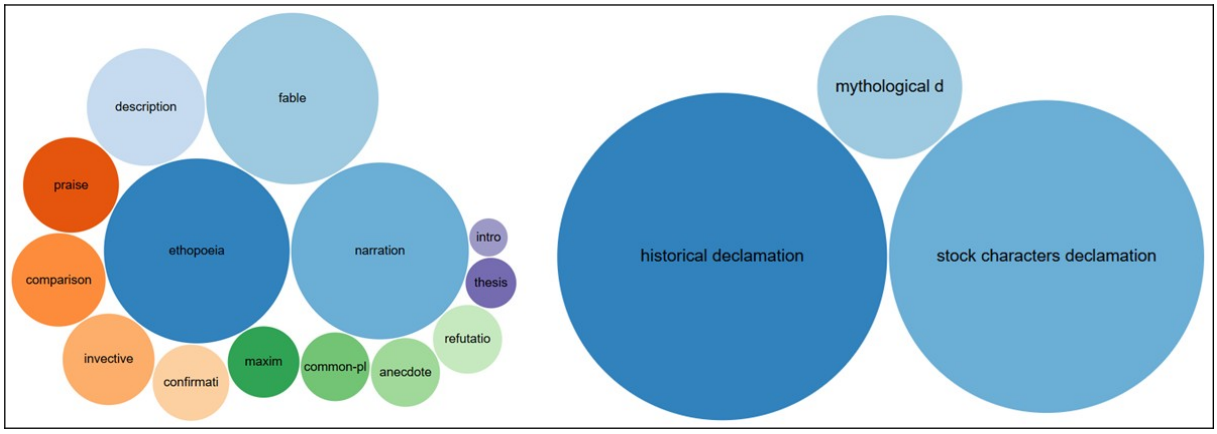

Fig. 4: Bubble charts representing the genres of extant *progymnasmata* (on the left) and of extant declamations (on the right) by number of texts; the colours are the default ones.

An important caveat to the data previously shown regards the attribution of these texts: each *hypothesis* is presently considered according to its belonging to a collection (e.g. all the 27 texts in the corpus of Libanius' *ethopoeiae* are counted as by Libanius), but many judgements of inauthenticity have been expressed by modern scholars for single texts in these collections, and taking these judgements into account could significantly modify these numbers. This is one of the possible future improvements to the database (see below *Possible future developments*).

Other interesting statistics can be drawn about the characters of the extant *progymnasmata* and declamations, e.g. exploring the most common characters in different contexts:

- in extant *progymnasmata* and declamations regarding the Trojan war,[88] Achilles is by far the most frequent character (23 appearances), followed distantly by Odysseus (9), Ajax son of Telamon (8) and Hector (5);

87 Cf. SPARQL query: https://tinyurl.com/279ptags (last access 11.07.2025).

88 Cf. SPARQL query: https://tinyurl.com/26tu2o7q (last access 11.07.2025).

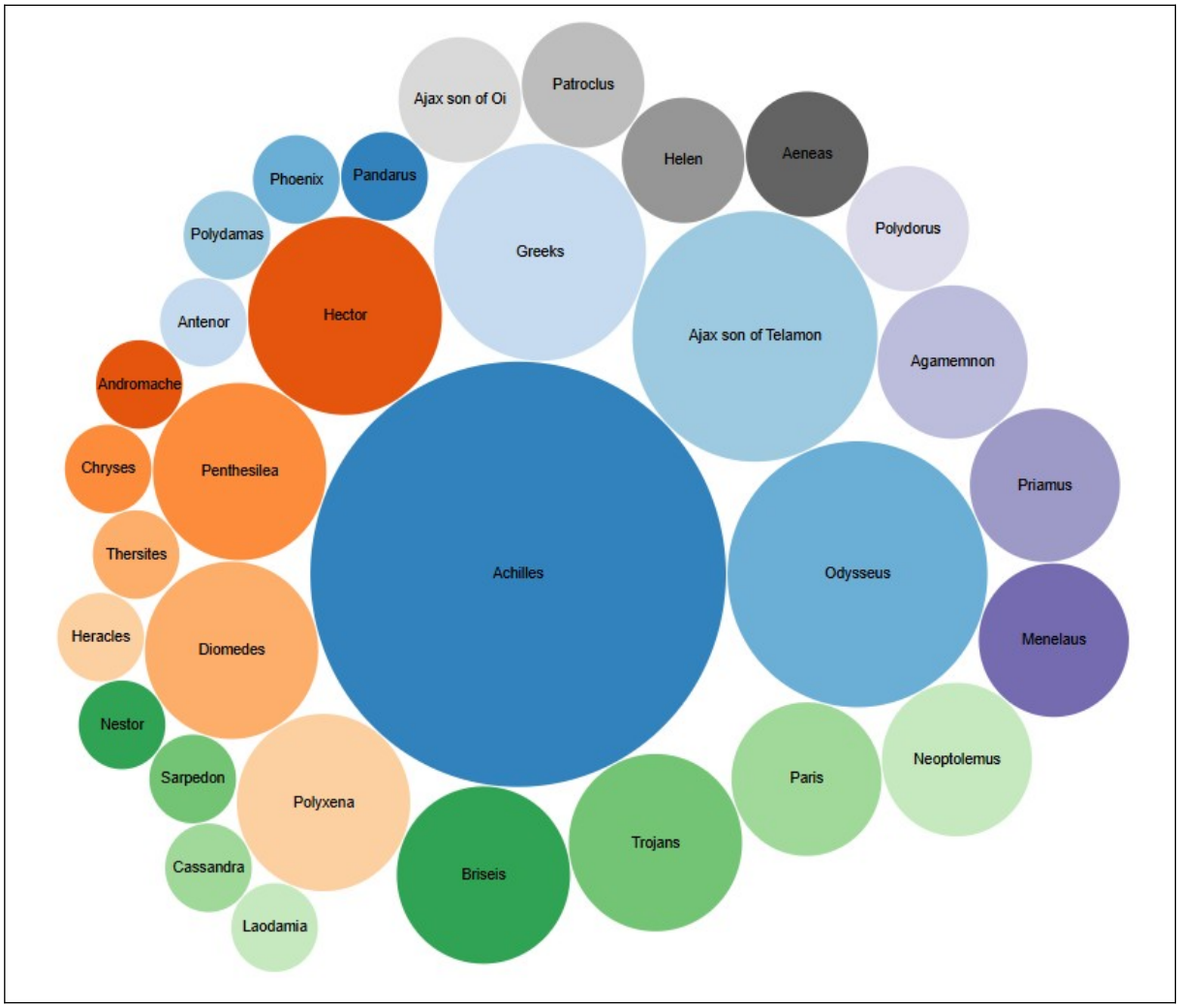

Fig. 5: Bubble chart representing the most frequent characters of extant *progymnasmata* and declamations regarding the Trojan war by number of appearances; the colours are the default ones.

- in extant *ethopoeiae* and declamations,[89] the most common named fictional speakers are Demosthenes (9 appearances), Achilles (8), Heracles (5), and Odysseus, Menelaus and Ajax son of Telamon (4).

89 Cf. SPARQL query: https://tinyurl.com/2cesrkj4 (last access 11.07.2025).

I conclude with an interesting visualisation of the most common values of P26 ("action(s)/state(s)") in all the extant *progymnasmata* and declamations:[90]

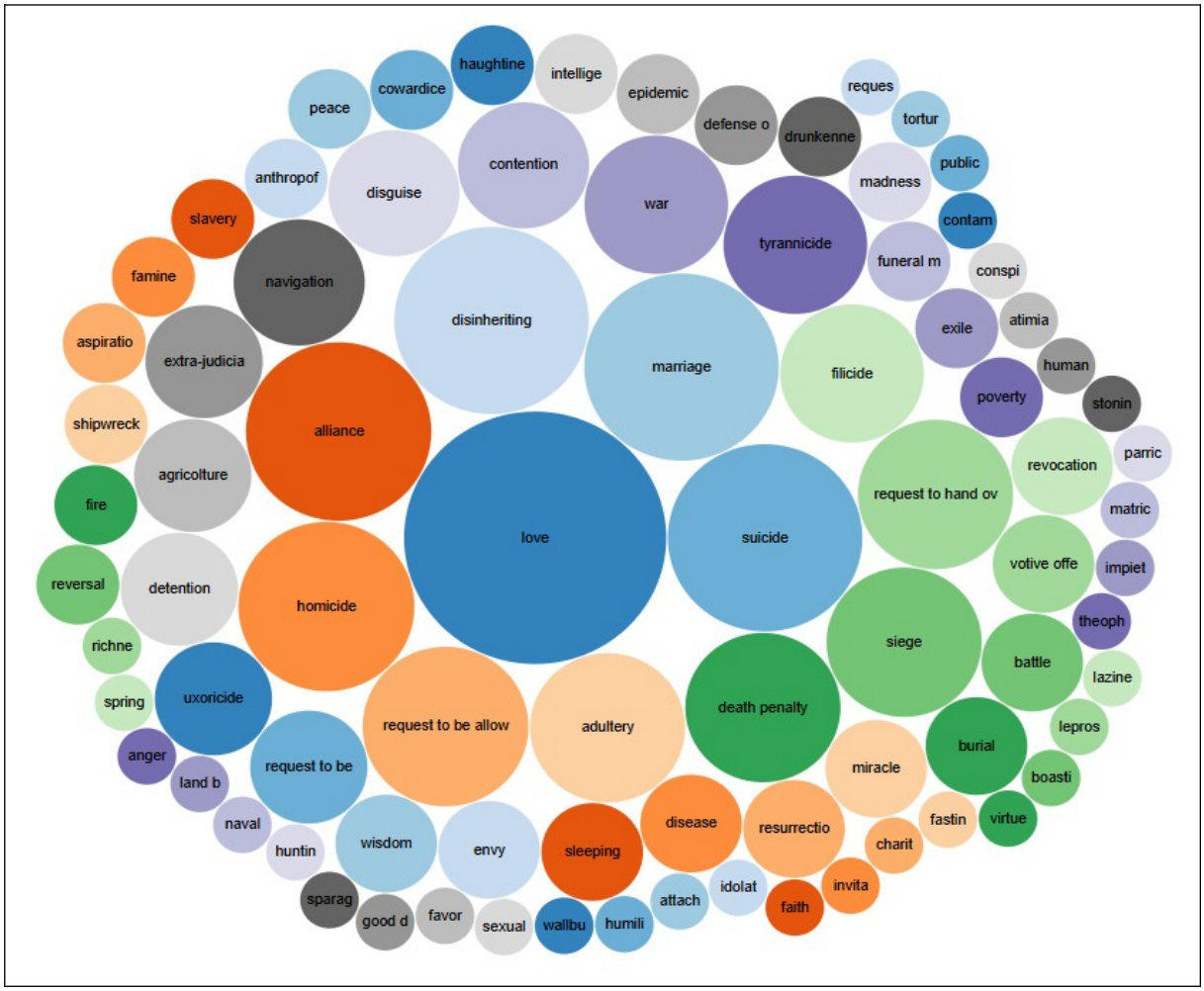

Fig. 6: Bubble chart representing the most frequent actions/conditions in the extant *progymnasmata* and declamations by number of appearances; the colours are the default ones.

The above data, despite the described limitations, already show that SPARQL queries on *Hypotheseis* could be an effective mean to statistically analyse this corpus of texts through the combination of multiple criteria and to visualise data about it.

Finally, SPARQL queries can also be used to find inconsistencies, incoherences, and mistakes in the data, so as to be an effective tool for the routine maintenance of the data quality in the database.

Possible Future Developments

The main path for expanding *Hypotheseis* is surely continuing the additions of *hypotheseis*, starting from the already mentioned groups of *progymnasmata* that are still missing in the database: *progymnasmata* in verses and *progymnasmata* discovered in papyri. Moreover, a systematic research of themes of *progymnasmata* and declamations mentioned in other works is yet to begin, but is one of the long-term aims of the project.

Other possible additions to the database include translating labels in other modern languages besides Italian and English and providing structured data about the editions – as of now, for expressions, only the name of the critical editor(s) is specified – and the translations of the works. The data about the translations could also be added directly to *Wikidata* and retrieved through federated SPARQL queries.

90 Cf. SPARQL query: https://tinyurl.com/2xjgb3hq (last access 11.07.2025).

It would also be possible to improve the data model of *Hypotheseis* to store data about the judgements of authenticity and inauthenticity that have been pronounced about specific *progymnasmata* and declamations, since the authorship of many of them is debated by modern scholars, and manuscripts themselves often show different attributions (see above *Data analysis and visualisation*).

The scope of *Hypotheseis* could also be expanded to include rhetorical exercises in languages other than Greek, especially Latin declamations, which would give interesting possibilities in terms of analyses of common themes. The addition of Armenian *progymnasmata* started in May 2025. The possibility of including also Greek literary works whose literary genre has some resemblances to *progymnasmata* and declamations, e.g. literary letters (Aelian, Alciphron, Philostratus, Aristaenetus, Theophylact Simocatta), could also be considered.

In order to ensure the accuracy and the coherence of the data of the database, it would also be useful to involve more editors, which could check already entered data and discuss more deeply the data model presently used. In fact, *Wikibase* is designed to create collaboratively edited databases, so its current use by one single editor unleashes only a fraction of its potential. When other editors will join the database, it will be a priority to translate documentation pages into English, since most of them are presently available only in Italian;[91] it is already possible to show a textual page in Italian or in English according to the interface language chosen by the user.[92]

As of now, *Wikibase* is not designed to host long texts as structured data, for various reasons. Firstly, the maximum length of labels, descriptions, and aliases, as well as of string and monolingual text values in statements, is 2500 characters. Secondly, inside these fields no markup is visually rendered (e.g. the string text is visualised in this exact way)[93] and it is impossible to store multiline texts.[94] However, it would be possible to store texts in *MediaWiki* textual pages and link them from items;[95] these options could be explored e.g. to store complete translations of rhetorical exercises made by the editors of the database. However, translations of *hypotheseis* only could well be stored directly as values of monolingual text statements, as they would remain within the 2500-character limit (the longest *hypothesis* presently stored in *Hypotheseis* is 1309 characters long,[96] and it is unlikely that longer ones will ever be added).

91 Cf. https://hypotheseis.wikibase.cloud/wiki/Category:Documentazione (last access 11.07.2025) and https://hypotheseis.wikibase.cloud/wiki/Category:Documentation (last access 11.07.2025).

92 Compare the first parts of https://hypotheseis.wikibase.cloud/wiki/Pagina_principale?uselang=it (last access 11.07.2025) and https://hypotheseis.wikibase.cloud/wiki/Pagina_principale?uselang=en (last access 11.07.2025).

93 Cf. request to create "a datatype capable of interpreting XML markup" in *Wikibase*: https://phabricator.wikimedia.org/T372326 (last access 11.07.2025).

94 Cf. request to create a datatype for multiline text in *Wikibase*: https://phabricator.wikimedia.org/T323705 (last access 11.07.2025).

95 To make this linking easier, it would be useful to have the possibility to create sitelinks from items to textual pages of the same *Wikibase* instance; cf. request to add this feature: https://phabricator.wikimedia.org/T330672 (last access 11.07.2025).

96 It is the *hypothesis* of Manuel Decl.: https://hypotheseis.wikibase.cloud/entity/Q1460 (last access 11.07.2025).

Conclusions

Greek *progymnasmata* and declamations are a wide but understudied corpus of texts, and structured data is required in order to apply a wide range of statistical analyses, which would have been significantly difficult or impossible with only the existing textual studies.

The software *Wikibase* has proven to be a good choice for creating databases in the field of Digital Humanities for a wide range of reasons: among the most relevant ones, being open-source and community-managed, giving the opportunity to create a customisable data model, enabling data entry through user-friendly interfaces (both manually and massively) and providing various ways to retrieve data (most notably a SPARQL endpoint). The platform *Wikibase Cloud* allows freely creating up to six cloud-hosted *Wikibase* instances, thus permitting scholars to focus exclusively on data curation, without worrying about costs and technical issues. This platform, in fact, has already been used to create many *Wikibase* instances in the field of classical studies,[97] and is likely to host others in the future.

The database *Hypotheseis*, started on March 27th 2024, aims to collect structured data about Greek *progymnasmata* and declamations, with the long-term target to describe also the themes mentioned in other works (especially rhetorical manuals) and not only extant texts. A data model has been designed to structure these data and has already been experimented by collecting data about most of the extant prose *progymnasmata* and declamations; SPARQL queries show that this data model is effective in allowing multiple kinds of statistical analyses about these texts.

The database could be developed in many possible ways, including cataloguing modern editions and translations of these texts and the scholarly debate about their authenticity, although the main purpose remains the extraction of themes mentioned in rhetorical manuals. However, in order to accomplish these aims, it will be fundamental to involve more editors in data curation, since as of now the database has been curated solely by the undersigned.

97 In 2024, apart from *Hypotheseis*, *DataLib* about Libanius' *Letters* (https://datalib.wikibase.cloud/ [last access 11.07.2025]) and *Greek Metrical Inscriptions* (https://greek-metrical-inscriptions.wikibase.cloud/ [last access 11.07.2025]) have also been created.

List of Abbreviations

API	Application Programming Interface
FAIR	Findable Accessible Interoperable Reusable
CSV	Comma-Separated Values
Fig.	Figure
LBG	Lexikon zur byzantinischen Gräzität
LSJ	Liddell-Scott-Jones
PID	Permanent Identifier
SPARQL	SPARQL Protocol and RDF Query Language
URI	Uniform Resource Identifier
WMDE	Wikimedia Deutschland
WMF	Wikimedia Foundation

Sources

Online Sources

https://hypotheseis.wikibase.cloud/ (last access 11.07.2025).

https://www.wikidata.org/ (last access 11.07.2025).

Digital Corpora

TLG database

Text Editions

Felten (1913): I. Felten (ed.), Nicolai progymnasmata, Leipzig 1913.

Kennedy (2003): G. Kennedy, Progymnasmata: Greek Textbooks of Prose Composition and Rhetoric, Leiden / Boston 2003.

Patillon / Bolognesi (1997): M. Patillon / G. Bolognesi (eds. trans.), Aelius Théon, progymnasmata, Paris 1997.

Patillon (2008): M. Patillon (ed. trans.), Corpus rhetoricum: Anonyme, Préambule à la rhétorique; Aphthonios, progymnasmata; Pseudo-Hermogène, progymnasmata, Paris 2008.

Reche Martínez (1991): M. D. Reche Martínez, Teón. Hermógenes. Aftonio. Ejercicios de retorica, Madrid 1991.

References

Agosti (2005): L'etopea nella poesia greca tardoantica, in: Amato / Schamp (2005), 34–60.

Amato et al. (2015): E. Amato / F. Citti / B. Huelsenbeck (eds.), Law and Ethics in Greek and Roman Declamation, Berlin 2015, https://doi.org/10.1515/9783110401882 (last access 27.12.2025).

Amato / Schamp (2005): E. Amato / J. Schamp (eds.), Ethopoiia: la représentation de caractères entre fiction scolaire et réalité vivante à l'époque impériale et tardive, Salerno 2005.

Amato / Ventrella (2005): E. Amato / G. Ventrella, L'éthopée dans la pratique scolaire et littéraire, in: Amato / Schamp 2005, 213–231.

Berardi (2017): F. Berardi, La retorica degli esercizi preparatori: glossario ragionato dei Progymnásmata, Hildesheim 2017.

Chiron (2017): P. Chiron, Les progymnasmata de l'Antiquité gréco-latine, Lustrum 59 (2017), 7–129.

Cribiore (1996): R. Cribiore, Writing, Teachers, and Students in Graeco-Roman Egypt, Atlanta 1996.

Cribiore (2001): R. Cribiore, Gymnastics of the Mind: Greek Education in Hellenistic and Roman Egypt, Princeton 2001.

Fernández Delgado (2025): J. A. Fernández Delgado, Homero en la enseñanza práctica de la retórica griega, Anuario de Estudios Filológicos 48 (2025), 109–130, https://doi.org/10.17398/2660-7301.48.109 (last access 27.12.2025).

Gibson (2004): C. A. Gibson, Learning Greek History in the Ancient Classroom: The Evidence of the Treatises on progymnasmata, Classical Philology 99/2 (2004), 103–129, https://doi.org/10.1086/423858 (last access 27.12.2025).

Guast (2023): W. Guast, Greek Declamation and the Roman Empire, Cambridge 2023, https://doi.org/10.1017/9781009297158 (last access 27.12.2025).

Hock / O'Neil (1986): R. F. Hock / E. N. O'Neil, The Chreia in Ancient Rhetoric 1: The progymnasmata, Atlanta, 1986.

Hock / O'Neil (2002): R. F. Hock / E. N. O'Neil, The Chreia and Ancient Rhetoric: Classroom Exercises, Atlanta 2002.

Hock (2012): R. F. Hock, The Chreia and Ancient Rhetoric: Commentaries on Aphthonius's progymnasmata, Atlanta 2012.

Hunger (1978): H. Hunger, Die hochsprachliche profane Literatur der Byzantiner, Munich 1978.

IFLA LRM (2017): P. Riva / P. Le Bœuf / M. Žumer, Consolidation Editorial Group of the IFLA FRBR Review Group, IFLA Library Reference Model. A Conceptual Model for Bibliographic Information, Den Haag 2017.

Jacobs (1899): J. Jacobs, De progymnasmaticorum studiis mythographicis, Marpurgi Cattorum 1899.

Kohl (1915): R. Kohl, De scholasticarum declamationum argumentis ex historia petitis, Paderborn 1915.

Morgan (1998): T. Morgan, Literate Education in the Hellenistic and Roman Worlds, Cambridge 1998.

Russell (1983): D. A. Russell, Greek Declamation, Cambridge 1983, https://doi.org/10.1017/CBO9780511897887 (last access 27.12.2025).

Ureña Bracero (1999): J. Ureña Bracero, Homero en la formación retórico-escolar griega: etopeyas con tema del ciclo troyano, Revista de Lingüística y Filología Clásica 67/2 (1999), 315–338, https://doi.org/10.3989/emerita.1999.v67.i2.178 (last access 27.12.2025).

Ureña Bracero (2005): J. Ureña Bracero, El uso de fuentes literarias, recursos retóricos y técnicas de composición en etopeyas sobre un mismo tema, in: Amato / Schamp (2005), 93–111.

Figure References

Fig. 1 Camillo Carlo Pellizzari di San Girolamo.

Fig. 2 On the left: from the SPARQL query https://tinyurl.com/262we83r (last access 11.07.2025); on the right: from the SPARQL query https://tinyurl.com/22p6y32l (last access 11.07.2025).

Fig. 3 On the left: from the SPARQL query https://tinyurl.com/259x5c6n (last access 11.07.2025); on the right: from the SPARQL query https://tinyurl.com/25neodcy (last access 11.07.2025).

Fig. 4 On the left: from the SPARQL query https://tinyurl.com/27f732ec (last access 11.07.2025); on the right: from the SPARQL query https://tinyurl.com/27886ong (last access 11.07.2025).

Fig. 5 From the SPARQL query https://tinyurl.com/2b3976m2 (last access 11.07.2025).

Fig. 6 From the SPARQL query https://tinyurl.com/2db9k66r (last access 11.07.2025).

Author Contact Information[98]

Camillo Carlo Pellizzari di San Girolamo
Scuola Normale Superiore
Piazza dei Cavalieri 7
56126 Pisa
E-mail: camillo.pellizzaridisangirolamo@sns.it

98