

THE EDITION OF LATIN GRAMMAR OF FATHER MANUEL ÁLVARES, JAPAN, 1594: BRIEF NOTE AND BIBLIOGRAPHICAL REFERENCES

Carlos Assunção*

Abstract: The scope of this study is to develop some fundamental brief notes regarding the edition of Latin grammar of Father Manuel Álvares, printed in Amacusa, Japan, in 1594. Apart from a brief note about editing, an analysis of the *bibliographical references* of the work will be carried out, which includes the implementation of the programme Nooj.

Keywords: Grammar. Latin. Japanese.

1.¹Without permission, the Japanese edition *Emmanuelis Alvares e Societate Iesu de institutione grammatica libri tres: Coniugationibus accessit interpretatio Iapponica* was published eleven years after the author's death, in 1594, by the Japanese Jesuit College of Amakusa². Only two copies of this edition are known to exist: one in the Biblioteca Pública de Évora, Portugal, and the other in the Biblioteca Angelica, Rome, Italy.

The Amakusa (Japanese) edition is in a 4^o format (22.7 x 15.7 cm in the Évora copy³), and occupies a total of 170, mostly paginated, folios printed on a Japanese *torinoko*⁴ paper. The Évora copy has a cover of calfskin-pasted board, with four raised cords. A paste-down and two fly-leaves of western paper are added on both sides of the cover respectively. A fragment of a unidentified incunabula of *Missa de nomine Iesu* (possibly *Missale ad usum insignis ecclesiae eboracensis*) is used as a lining of the front cover⁵.

* Doutor em Linguística pela Universidade de Trás-os-Montes e Alto Douro (Utad) – Vila Real – Portugal. Professor catedrático na mesma instituição. E-mail: cassunca@utad.pt

1 Followed to describe the point 1 (ASSUNÇÃO; TOYOSHIMA, 2012, p. 261-262).

2 Amongst other scholars, Matos (1987, p. 164) maintains that there may have been a 1593 edition of the Amakusa grammar. However, as the author states that “um exemplar encontra-se na Biblioteca Angélica, de Roma, e outro na Biblioteca Pública de Évora”, it seems fair to conclude that the indication on the title page M. D. XCIII. (instead of a traditional M. D. XCIV) might have led to the misinterpretation.

3 The Angelica copy is slightly smaller, 21.0 x 15.0 cm, because it is (sometimes too) deeply trimmed during binding.

4 Torinoko is a specific type of Japanese paper with smooth surface, lustrous, thin, and quite resistant. The Jesuits in Japan used torinoko (so called in the Jesuit inventories) papers for their publications by Latin script (for publications in Japanese script, they used minogami, i.e. mulberry papers), in official records, and in correspondences to Rome.

5 On the front cover is a printed ticket “IV Centenário da fundação da Universidade de Évora”, attached on the occasion of the exposition (1959) commemorating the 4th centenary of the University of Évora, in which the Amakusa edition was the 86th

The Angelica copy also had a cover of calfskin-pasted board, but has recently been re-furbished, with fragments of the old cover pasted on the new modern cover, with three raised cords. A modern pastedown, and two contemporary fly-leaves of a very rough (probably Japanese) paper are respectively added on both sides of the cover. The paste-down used to contain a fragment of a printed leaf of *Guia do Pecador* (1599, Nagasaki) as a lining, but is currently detached, and has been replaced by a modern paste-down. On the beginning fly-leaf is a dedication “Anno Domi [...] 1605 / Petrus Antonius A[...] Iapp[...] / hũc librũ dono dedit / Bibliothecae / A[...]ae / [.....]” (*in the year of the Lord 1605, a Japanese Petrus Antonius A[raki]?⁶ gave this book to the Augustine Library as a gift*). The last line of the dedication (which may have been a signature) is crossed out, and is illegible.

The two copies differ slightly, in the correction of typographical errors: 3v (the Angelica copy corrects the error in the Évora copy), 20r (E corrects A), 23v (E and A are both in error), 30r (E corrects A), 100v (A corrects E). Occasional interpolations exist only in the Évora copy.

In accordance with the subtitle’s announcement (“*Conjugationibus accessit interpretatio japonica*” [Japanese interpretation added to the conjugations]) the Japanese edition distinguishes itself from the “simple” editions of the Álvares’ grammar by the addition of the Japanese conjugations to the Latin and Portuguese conjugations that already were a part of the Portuguese editions of the “arte grande” and the “arte pequena”.

2. The electronic version of Padre Manuel Álvares’ *De Institutione Grammatica* edited in Amacusa in 1594, was explored using version 2.5 of NooJ, a linguistic development environment developed by Max Silberstein of the University of Franche-Comté and available at www.nooj4nlp.net.

In the approach taken, no features of automatic linguistic analysis were implemented, given the weak adequacy of existing resources to the specificity of the text under study: the plethora of abbreviations used, coupled with the fact that this is a text of the late sixteenth century, written in Portuguese, Latin and Japanese would require the adaptation of the available linguistic resources. Notwithstanding the multiplicity of languages for which modules of linguistic resources are available, Japanese is not yet among them. In addition, multi-language NooJ features are best adapted to studies that fit the scheme language-of-origin *vs* language-of-destination, whether they are of a comparative-contrastive nature, or they are regarding automatic translation, thereby making it necessary to reconfigure the software for the analysis of a multilingual corpus.

Thus, the NooJ features used in this study were of a statistical approach and of research and construction of concordances from the unannotated corpus (raw corpus), disregarding the potential of the metalinguistic information-based approach for the reasons provided.

A comprehensive statistical portrait of the text will be the first step, focusing attention on the bibliographical references, given the profusion of authors and works cited throughout *De Institutione Grammatica*, both from a statistical perspective and from the compliance of spelling variations relating to works and authors.

exhibit. On the flyleaf of the back cover a card of the library is pasted saying “Esteve esposto na Europália 89 / Japão” (“This was exhibited in the Europalia [1989 [Belgium/] Japan”).

6 The paper is broken, and readable as “Am” or “Ar”. In the former case “amicus” (a friend), and if the latter, the donator may be Petrus Antonius Arachius Iaponensis, i.e. Thomas Araki (? - 1646?), who is known to have studied in the Pontifical Roman Seminary to be ordained there.

According to Nooj, the .rtf file of the work under review consists of a total of 415,695 characters, including 77,588 blank spaces, 23,370 other delimiters (namely punctuation marks, dashes and parentheses) 1,443 digits and 313,294 letters and, of this total, there are 86 characters, 4 types of blank spaces, 9 delimiters, 10 digits and 63 different letters.

If the blank spaces and punctuation marks are excluded, the ten most frequent characters are, in descending order: *e* (34,228 occurrences), *i* (31,551), *u* (29,681), *a* (26,109), *t* (23,338), *s* (20,909), *r* (19,651), *o* (18,624), *n* (16,451) and *m* (15,953). Impressive is the fact that the sum of the total of occurrences of the 10 most frequent letters (236,495) corresponds to 75% of the total sum of letters of the text. The first capitalized letter, in terms of frequency, occupies 18th place, with 2,928 occurrences (*A*).

At the opposite extreme, with only one occurrence throughout the text, are the following characters: *â*, *ê*, *í* and *ô*.

Regarding the total of 57,496 words in the text, it appears that, in 16,362 different forms, the ten most common are the following:

Table 1 – Frequency of different forms

Order	Form	Frequency
1 st	et	836
2 nd	ut	794
3 rd	uel	782
4 th	in	723
5 th	lib	550
6 th	a	513
7 th	Idem	482
8 th	est	476
9 th	Vare	402
10 th	A	400

Source: Elaborate by the author.

Seven of these ten forms belong to the group of so-called grammatical or functional words – conjunctions, prepositions, pronouns – which are usually the most frequent in any corpus; given their nature. Nor is it surprising to find the forms *est* and *Vare* among the most frequent, as they correspond to the most frequent verb form in all languages. Finding the abbreviation *lib.* in fifth place is unusual, and is an indicative symptom of the importance that citations occupy in the work. If we continue to look at the list, the signs of such importance do not take long before multi-plying: among the hundred most frequent forms are specific references to works and authors such as:

Table 2 – Frequent authors

Order	Form	Frequency
17 th	Cic. (Cicero)	265
20 th	Aen. (<i>Aeneid</i>)	224
27 th	Virg. (Virgil)	184
60 th	Atti. (<i>Letters to Atticus</i> , by Cicero)	91
63 th	Ouid. (Ovid)	87
67 th	Terent (Terentius)	82

Source: Elaborate by the author.

The list of the sequences of the two most frequent words reveal the importance of the bibliographical references: in the first five places, three sequences explaining works or authors are found, and in the first 100 most frequent sequences there are 26 that explicitly refer to works or authors:

Table 3 – Frequency of the references

Order	Form	Frequency
1 st	Virg Aen (<i>Aeneid</i>)	137
4 th	Atti lib (<i>Letters to Atticus</i> , by Cicero)	89
5 th	Idem Aen (<i>Aeneid</i>)	84
13 th	Idem in (author mentioned above)	56
14 th	Idem ad (author mentioned above)	51
15 th	Mart lib (Martial)	49
17 th	Cic Atti (<i>Letters to Atticus</i> , by Cicero)	47
18 th	Cic in (Cicero)	47
23 rd	Cic de (Cicero)	45
24 th	Ouid Met (<i>Metamorphoses</i> , by Ovid)	42
29 th	lib c (abbreviation for book and chapter)	38
31 st	Verr lib (Cicero, <i>Against Verres</i>)	38
39 th	in Verr (Cicero, <i>Against Verres</i>)	33
44 th	Idem pro (author mentioned above)	31
46 th	Anton Philip (<i>Philippics</i> , Cicero)	30
57 th	Cic pro (Cicero)	27

(continua)

Table 3 – Frequency of the references (*continuação*)

Order	Form	Frequency
60 th	Idem lib (author mentioned above)	27
61 st	Liu Bel (Titus Livius)	27
62 nd	Lucan lib (Lucanus)	27
76 th	Ab Vrb (Titus Livius)	21
85 th	Cic ad (Cicero)	20
87 th	Idem Ad (author mentioned above)	20
96 th	de Orat (Cicero, <i>Do orador</i>)	19
98 th	Liu Ab (Titus Livius)	19
99 th	Plin lib (Pliny the Elder)	19
100 th	Virg Eclog (<i>Eclogues</i> , by Virgil)	19

Source: Elaborate by the author.

Using the information provided by the listing of the most frequent forms, including the sets of two words, and undertaking routine and filter research based on the forms starting with a capital letter, it was possible to determine that there are 56 authors/authorities explicitly referred to or quoted in the text, namely:

Table 4 – Frequency of authorities references

Anonymous (<i>Reth. Ad Her.</i>)	Isocrates	Quintus Curtius Rufus
Aristarchus	Justinian	Rongo
Aristocrites	Juvenal	Sallustius
Aulus Gellius	Lucanus	Seneca
Aulus Hirtius	Lucretius	Silius Italicus
Cato	Manuel Álvares	Suetonius
Catullus	Marcellus	Sulpicius
Caesar	Martial	Tacitus
Cicero	Marcus Valerius Probus	Thales
Columela	Ovid	Theocritus
Demosthenes	Persius	Terentius
Diomedes	Pythagoras	Tibulus
Donatus	Plato	Titus Livius
Enio	Plautus	Valerius Flaccus
Statius	Pliny the Elder	Valerius Maximus
Feiqe	Pomponius Mela	Varro
Focas	Priscianus	Verrius Flaccus
Herodotus	Propertius	Virgil
Horatius	Quintilianus	

Source: Elaborate by the author.

As to the frequency with which each author is explicitly mentioned, the list of most frequent forms provides an indication, as shown, that the form *Cic.* (Cicero) has 265 occurrences, occupying first place. This does not mean, however, that Cicero, as the author, is mentioned 265 times in the text because, besides the times that the author is referred to through the expressions *idem* and *ibidem*, Cicero is identified in the text under review, through seven graphic variations for each of which number of occurrences is given: *Ciceronis* (3), *Cic.* (265), *Cicero* (13), *Ciceronem* (4), *Marcus Tullius Cicero* (1), *Ciceroni* (4) and *M. Tullium* (1). We, therefore, conclude that Cicero is explicitly named 291 times, considering the graphic variants presented.

For Virgil, the second author to appear in the table of the most frequent (*Virg.*, in 27th place, with 184 occurrences), the following seven variations are used, in a total of 196 occurrences: *Virgilii* (2), *Virg.* (184), *Virgilius* (4), *Virgil.* (2), *Virgilis* (1), *Virgilium* (2) and *Aut.* (4).

Considering that the authors who are among the hundred most frequent forms of text, the cases of Ovid and Terentius should be referred to. As for Ovid, who appears in 63rd place in the form *Ouid*, with 87 occurrences, he is also referred to as *Ovidius* (5) and *Ouidi* (1), a total of 93 explicit references. Terentius, whose form *Terent* comes in 67th place with 82 occurrences, is also referred to as *Terentii* (one occurrence).

Regarding the graphical diversity, besides those already mentioned, the following authors can also be named since they are contemplated with more than a graphical form:

- Caesar (*C. Caesarem*, *Caesaris*, *Caesar*, *Caes.* and *Caesare*), in a total of 28 occurrences;
- Diomedes (*Diomedis* and *Diomedes*), in a total of 3 occurrences;
- Donatus (*Donati* and *Donatus*), in a total of 3 occurrences;
- Horatius (*Horatii*, *Horatius* and *Horat.*), in a total of 36 occurrences;
- Lucanus (*Lucanus* and *Lucan.*) in a total of 28 occurrences;
- Martial (*Mart.*, *Martialis*, *Martial.* and *Martiale*), in a total of 58 occurrences;
- Persius (*Persius* and *Pers.*), in a total of 5 occurrences;
- Plato (*Platoni*, *Platonis* and *Plato*), in a total of 5 occurrences;
- Plautus (*Plaut.* and *Plautus*), in a total of 16 occurrences;
- Pliny the Elder (*Plinium*, *Plin.* and *Plinius*), in a total of 16 occurrences;
- Pomponius Mela (*Pomponius. Mela*, *Pompon. Mela* and *Pomp. Mela*), in a total of 3 occurrences;
- Priscianus (*Prisciani* and *Prisc.*), in a total of 4 occurrences;
- Quintilianus (*Fabii Quintiliani*, *Quint.*, *Quintil.*, *Quintilianus* and *Quintiliano*), in a total of 25 occurrences;
- Sallustius (*Sallust.*, *Sallustius*, *Sallusti* and *Sall.*), in a total of 15 occurrences;
- Seneca (*Seneca* and *Senec.*), in a total of 12 occurrences;
- Silius Italicus (*Silius*, *Silium* and *Sil.*), in a total of 12 occurrences;
- Suetonius (*Sueto*, *Suet.* and *Suetonius*), in a total of 5 occurrences;

- Sulpicius (*Sulpitius e Sulpitio*), in a total of 6 occurrences;
- Titus Livius (*Liuii, Liuius and Liu.*), in a total of 56 occurrences;
- Varro (*M. Varronis and M. Varronem*), in a total of 2 occurrences.

This graphic diversity extends to the reference titles of cited works: for Cicero only, the most cited author in the grammar, 40 different titles through 96 different forms are named. Cicero is, indeed by far, the author with the greatest variety of titles cited. Of Virgil, the second author in the list of the most frequently mentioned, there are only three works cited: *Aeneid*, *Georgics* and *Bucolics*, in a total of five variants. Of Ovid there are four titles mentioned (*Ex Ponto*, *Metamorphoses*, *Tristia* and *Amores*) in 8 different forms. To highlight just the authors who appear among the hundred most frequent forms, it should be noted that there are six comedies by Terentius mentioned, in a total of 13 different forms.

As already noted, the authors and works are often referred to through the forms *idem* and *ibidem*, so that any counting made from the forms that unambiguously refer to each author may not be reliable. To contextually resolve the ambiguity of the form *idem*, considering that this is usually followed by a reference to the title, four grammars were put together, each one dedicated to each of the authors that are mentioned in the text. These grammars are meant to solve the semantic ambiguity of the form *idem* when applied in the context of bibliographic citation, also returning references, when applied to text, already identified with each of the individual authors, which allows for the retrieval of a total of occurrences that results from the sum of the forms already identified and the contexts in which the use of *idem* refers to the author concerned.

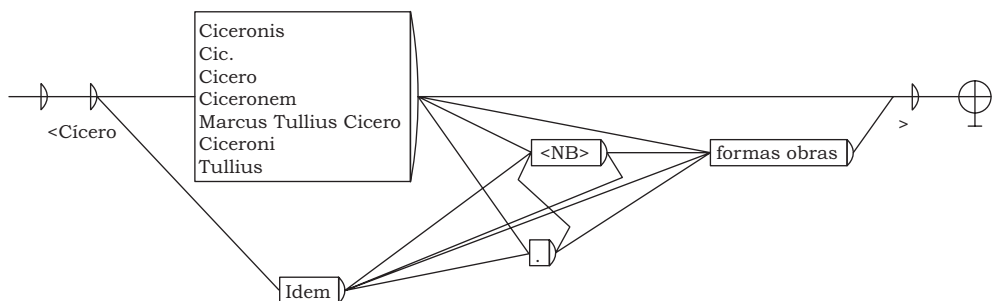


Figure 1 – Image of disambiguation grammar

Source: Elaborate by the author.

As example, the disambiguation grammar allusive to Cicero:

1. Reference to the author via the identified graphical variants of his name.
2. Contexts in which a graphical variant of the author's name is followed by one of the 96 forms, identified as referring to titles of his works (listed in the sub-graph *formas_obras*, which appears shaded in the image, and not illustrated due to its excessive size), with comma(s) or numeral(s) placed or not between the author's name and title of work.
3. Contexts in which the form *idem* (or *Idem*, or *IDEM*) appears before one of the 96 identified forms as referring to titles of works of the author, with or without comma(s) or numeral(s) between the form *idem* and the title of the work.

Note that the priorities for implementation of each of the possible paths in the graph avoid duplicate results being returned at the outset, since, in every situation, only the longest path is considered by the program. For example: the sequence *Cic., 5, Tuscul* is identified and annotated by path 2 mentioned above, and will not be considered in path 1 that covers only those situations where the author's name does not precede the title of the work.

As each of these disambiguation grammars is applied to the text, the results are as follows:

- for Cicero, 386 occurrences, 95 of which with the form *idem*;
- for Virgil, 296 occurrences, 100 of which with the form *idem*;
- for Ovid, 106 occurrences, 13 of which with the form *idem*;
- for Terentius, 93 occurrences, 11 of which with the form *idem*.

The validity and effectiveness of the disambiguation grammars can be verified by comparison with the figures resulting from the accounting made to the total of occurrences of the various ways each author is referred to, because the results presented here are the exact sum of occurrences mentioned above with the number of disambiguated cases of the form *idem*.

In percentage terms, the author most referred to through the form *idem* is Virgil, with 34%, although the data seem to suggest a relationship of progressive increase between the number of times that an author is mentioned in the corpus and the number of times that the form *idem* refers to the same author. That is: the more times an author is referred to in a text, the greater the percentage of times that this author is identified through *idem*.

A EDIÇÃO DA GRAMÁTICA LATINA DO PADRE MANUEL ÁLVARES, JAPÃO, 1594: NOTA BREVE E REFERÊNCIAS BIBLIOGRÁFICAS

Resumo: Este estudo tem como escopo desenvolver algumas breves notas fundamentais sobre a edição de Gramática Latina do Padre Manuel Álvares, impressa em Amacusa, Japão, em 1594. Além de uma breve nota sobre a edição, uma análise das referências bibliográficas do trabalho será realizada através da implementação do programa Nooj.

Palavras-chave: *Gramática. Latim. Japonês.*

REFERENCES

ÁLVARES, M. (P.). *De Institutione Grammatica libri tres*. Amacusa: Collegio Amacusensi Societatis IESV, 1594.

ASSUNÇÃO, C.; TOYOSHIMA, M. Edicion ant Introduction. In: ASSUNÇÃO, C.; TOYOSHIMA, M. *Emmanuelis Alvari e Societate Iesu*. De Institutione Grammatica Libri Tres. Tokyo: Yagi Bookstore, 2012. p. 261-262.

MATOS, M. C. de. Humanismo e evangelização no Oriente no século XVI. *Revista ICALP*, v. 7 e 8, p. 41-72, mar./jun. 1987. Disponível em: <<http://www.nooj4nlp.net>>. Acesso em: 31 mar. 2014.

Recebido em fevereiro de 2014.

Aprovado em março de 2014.