

# West meets East

Papers in historical lexicography and  
lexicology from across the globe

Edited by

Geoffrey Williams

Mathilde Le Meur

Andrés Echavarría Peláez

World Histories of Lexicography  
and Lexicology 1



## World Histories of Lexicography and Lexicology

Editors: Fredric T. Dolezal (University of Georgia, USA), Roderick W. McConchie (Helsinki University, Finland)

In this series:

1. Williams, Geoffrey, Mathilde Le Meur & Andrés Echavarría-Peláez (eds.).  
West meets East: Papers in historical lexicography and lexicology from  
across the globe.

# West meets East

Papers in historical lexicography and  
lexicology from across the globe

Edited by

Geoffrey Williams

Mathilde Le Meur

Andrés Echavarría Peláez



Geoffrey Williams, Mathilde Le Meur & Andrés Echavarría Peláez (eds.). 2025. *West meets East: Papers in historical lexicography and lexicology from across the globe* (World Histories of Lexicography and Lexicology 1). Berlin: Language Science Press.

This title can be downloaded at:

<http://langsci-press.org/catalog/book/458>

© 2025, the authors

Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):

<http://creativecommons.org/licenses/by/4.0/> 

ISBN: 978-3-96110-503-8 (Digital)

978-3-98554-132-4 (Hardcover)

DOI: 10.5281/zenodo.15020225

Source code available from [www.github.com/langsci/458](http://www.github.com/langsci/458)

Errata: [paperhive.org/documents/remote?type=langsci&id=458](http://paperhive.org/documents/remote?type=langsci&id=458)

Cover and concept of design: Ulrike Harbort

Typesetting: Sebastian Nordhoff

Proofreading: Amir Ghorbanpour, Brett Reynolds, Christopher Straughn, Conor Pyle, David Carrasco Coquillat, Diana Lewis, Elliott Pearl, Jeroen van de Weijer, Katja Politt, Mary Ann Walter, Nicoletta Romeo, Rebecca Madlener, Sebastian Nordhoff, Vladan Sutanovac

Fonts: Libertinus, Arimo, DejaVu Sans Mono, Arabtype, Source Han Serif ZH

Typesetting software: Xe<sub>La</sub>TeX

Language Science Press

Scharnweberstraße 10

10247 Berlin, Germany

<http://langsci-press.org>

[support@langsci-press.org](mailto:support@langsci-press.org)

Storage and cataloguing done by FU Berlin

# Contents

1	<b>Introduction</b> Geoffrey Williams	1
2	<b>On closure and its challenges: Examining the editors' proofs of OED1</b> Lynda Mugglestone	11
3	<b>Dictionaries in the web of Alexandria: On the dangerous fragility of digital publication</b> Daphne Preston-Kendal	31
4	<b>A dictionary of the languages of medieval England: Issues and implications</b> Gloria Mambelli	43
5	<b>The treatment of English high-frequency verbs in the <i>Promptorium Parvulorum</i> (1440)</b> Kusujiro Miyoshi	59
6	<b>Disattributing the <i>Encyclopédie</i> article on <i>définition en logique</i> from Jean-Henri-Samuel Formey</b> Alexander K. Bocast	77
7	<b>Project Cleveland: Documenting the lexicographic output of 20th-century Slovenian immigrants in the US</b> Alenka Vrbinc, Donna Farina & Marjeta Vrbinc	99
8	<b>The incorporation of proper nouns of non-Slavic origin into the 16th-century Slovenian literary language</b> Alenka Jelovšek	141
9	<b>Dictionnaires manuscrits dans l'histoire de la lexicographie croate : Des recueils de mots aux trésors linguistiques et culturels</b> Ivana Franić	163

*Contents*

- 10 Évaluer la dette : L'étendue de la présence de Richelet dans le  
*Dictionnaire universel* de Basnage (1701)  
Clarissa Stincone 193
- 11 De Félibien à Boutard : L'évolution du dictionnaire artistique entre le  
XVIIème et le début du XIXème siècle  
Rosa Cetro 211
- 12 La valeur pragmatique des langues dites « orientales » dans le  
*Dictionnaire universel* de Trévoux (1721)  
Georgios (Yorgos) Kassiteridis 229
- 13 Musical terms of Greek and Italian origin in Ottoman Turkish  
lexicography  
Agata Pawlina 245
- 14 Exploring the unique method for encoding sinograms in the first  
known Chinese–Polish dictionary  
Andrzej Swoboda 271
- 15 Les travaux lexicographiques de Carlo da Castorano et ses tentatives  
pour faire imprimer un dictionnaire européen de chinois  
Mariasaria Gianninoto & Michela Bussotti 297
- 16 The bilingual dictionary as a mediator between West and East: The  
beginnings of English-Polish lexicography  
Mirosława Podhajecka 321
- 17 Lexicon Lapponicum Bipartitum.....ungarice scriptum: Hungarian  
aspects of North Saami dictionary writing  
Ivett Kelemen 349
- 18 Les exemples dans les dictionnaires français–hongrois à travers les  
siècles  
Gábor Tillinger 363
- 19 « Sul finir d'imparare la Grammatica Francese, fa d'uopo studiar il  
Dizionario delle Frasi » : Deux recueils phraséologiques bilingues  
franco-italiens de la première moitié du 19e siècle  
Michela Murano 399

<b>20</b>	<b><i>The Lexicon of Oriental Words in Ancient Greek</i></b>	
	Rafał Rosół	417
<b>21</b>	<b>The discovery of a Russian-Tajik dictionary</b>	
	Abdusalom Mamadnazarov & Bahriddin Navruzshoev	427
	<b>Index</b>	443



# Chapter 1

## Introduction

Geoffrey Williams

Université de Bretagne Sud; Université Grenoble Alpes

This work consists of a series of papers written following the International Conference for Historical Lexicography and Lexicology (ISHLL) held in Lorient, France in May 2022. The selection is in two languages – English and French, the language of the host country – but is organised thematically so as to underline content. The papers illustrate the vitality of this field by analysing historical dictionaries that reflect languages from across the world.

As we all know, lexicography in its many forms is a very old practical discipline solving practical problems concerning word usage. “Word” seems more appropriate than “language” as lexicography addresses more questions relating to what we now call lexicology than other issues. As with all areas of human endeavour, what grew gradually through trial and lots of error is eventually subjected to a theoretical framework. The role of historical lexicography is to look back on the development of these highly varied word lists to see how we reached the tremendous variety that characterises practice throughout the world. One issue is the word “dictionary” as this in fact severely reduces our scope of study, and this is why we insist on the sister discipline of historical lexicography as the analysis of word-usage and the description of words in formalised reference works are totally intertwined.

This volume is both a selection of papers from one conference on historical lexicography and lexicology held under the aegis of the International Society for Historical Lexicography and Lexicology (ISHLL) in Lorient, France in May 2022 and also the first in a new series dedicated to the area and which represents a collaboration between two sister associations ISHLL and HEL-Lex.

Our current vision of what we now term “dictionaries” is a relatively modern phenomenon developing largely from the late 16th century with major works



such as Robert Estienne's (1539) *Dictionnaire françois-latin* being called a "dictionary", whereas he had named his previous work a "thesaurus". As late as 1604, Robert Cawdrey called his work *A table alphabeticall* (1604), emphasizing simply the order in which the words were given, whilst at the same period the Accademia della Crusca were calling their far more consequential work a *Vocabolario* (Accademia della Crusca 1612). These are the origins of the European tradition of lexicography with the different terms gradually becoming stabilized as pointing to different objects. However, the use of such essential reference works is far older and can be traced through all written civilisations, notably China and India (Yong & Peng 2008). Whilst the chapters in this volume essentially deal with European style dictionaries, both monolingual and bi- or multilingual, this international perspective must be borne in mind.

The papers published here seek to demonstrate a world perspective in the array of languages covered, moving quite literally from west to east. This is a selection of papers presented at the *12th International Conference of Historical Lexicography and Lexicology* held in Lorient, France, in 2022 under the patronage of the *Académie française* and with the assistance of the French Ministry of Culture. This is a collective work with the chapters being rewritten and expanded, reflecting the interaction and dynamism of the event. Being held in France, the languages of the conference were English and French, two internationally used *linguae francae*, with French being very much the dominating language in the 18th century, a fact reflected in the fact that three of the chapters concern 18th century French lexicographical works. The event was also sponsored by the BasNum project<sup>1</sup> based at the *Université Grenoble Alpes* that is digitising the major encyclopaedic work of Antoine Furetière, the *Dictionnaire universel* in its 1701 edition.

The theme of this event was "West meets East" with the aim of looking at historical dictionaries right across Europe from Portugal and Great Britain in the West to Georgia in the East and in global terms from across the world.

The theme is also a play on place names as the named city in which it was held is literally "The Orient", Lorient. Lorient was founded *ex nihilo* in 1666 as a port to build ships for the French East Indies Company, la *Compagnie des Indes*. At this time, the bay had no coastal town, just dispersed fishing villages. The first ship built was the *Soleil de l'Orient* and the workers gradually became referred to as those from the l'Orient with the name sticking so that the future parish and burgeoning town adopted this name. Lorient is thus a city open to the world with strong international connection making it an ideal location to celebrate the East.

---

<sup>1</sup><https://anr.fr/Projet-ANR-18-CE38-0003>

The chapters represent a linguistic diversity through a variety of language combinations. Ordering them thematically is clearly a necessary impossibility in that the parameters are numerous. We have decided to order first by monolingual and bilingual and then to regroup by relatedness in the themes. However, it must be noted that whilst a thematic grouping has been attempted, we have deliberately decided to mix papers published in the two languages of the event: French and English.

If dictionaries are tools for the understanding and diffusion of language, then multilingualism as expressed through the Helsinki Initiative<sup>2</sup> is an important issue. The language of lexicography is not English, but English is just one of many vehicles for discussing research related to dictionaries. This is why this work has been compiled according to themes rather than language and why we have not created artificial sections by language. As in any international event, the majority of papers tend to be presented in English, whatever the native language of the writer. Thus, we consider that, as English and French are both international languages, it is not necessary to group productions. It is also important to stress that texts in French are not simply those of French native speakers, but scholars who work in this language in the same way as other non-native speakers adopt English. True multilingualism is not about privileging one language above another, but promoting variety and choice. However, the abstracts are systematically in English so that a wider audience can access introductory information.

The first paper is that of Lynda Mugglestone *On Closure and its Challenges: Examining the Editors' Proofs of OED1*. This chapter represents an entirely new look at the genesis of the *Oxford English Dictionary* using recently discovered the final Editors' Proofs that show the process of closure in detail. Not only do these show often amusing details as to fixing the text, they shed a great deal of light on the finalising of this important text in the pre-digital hot-press era. This clearly shows the trials and tribulations of putting a printed work to press with the question of access to historical data in the digital age.

In *Dictionaries in the Web of Alexandria: On the Dangerous Fragility of Digital Publication*, Daphne Preston-Kendal discusses long-term access to data in a situation where a work may be constantly updated, but where no paper reference edition exists. This "web of Alexandria" highlights the complexities of digital management and citing entries which are not necessarily as stable over time as a print edition. For example, if we look up any keyword like, say, 'dictionary', in any previous edition of OED, we can give an exact reference, but a dynamic evolving work can see its entries change from one day to the next. It is also a

---

<sup>2</sup><https://www.helsinki-initiative.org>

question of long term access as the number of copies of a file is limited to a small number of servers. This may seem a technical issue, and it is, but in a digital world such issues have very major consequences.

Two papers address issues relating to early English dictionaries: *A dictionary of the languages of medieval England: issues and implications* (Gloria Mambelli) and *The Treatment of English High-frequency Verbs in the Promptorium Parvulorum 1440* (Kusujiro Miyoshi). Mambelli tackles the complex issue of triglossia through the analysis of three historical dictionaries of English: the *Middle English Dictionary* (MED), the *Anglo-Norman Dictionary* (AND), and the *Dictionary of Medieval Latin from British Sources* (DMLBS). The discussion looks at how they handle a situation where a literate medieval Englishman had to master his variety of English, Anglo-French, which was the language of the court, and Latin, the language of the church, law and administration. The outcome is that existing historical dictionaries have a highly inconsistent approach to the three language groups and that a more recent work is much needed. Miyoshi tackles a different controversy, the treatment of high-frequency verbs in English. The preeminent place is traditionally awarded to Samuel Johnson, but here it is posited that a wider definition of dictionaries than simply monolingual works should be adopted and that in this case, the first lexicographer to tackle such an issue would be the medieval monk and grammarian Galfredus Grammaticus, the probable compiler of the otherwise anonymous *Promptorium Parvulorum*. Whatever position is adopted, this paper demonstrates the debt we have to early lexicographical endeavours.

Alexander Bocast's chapter *Disattributing the Encyclopédie article on définition en logique from Jean-Henri Samuel Formey* takes us across to 18th century France and another controversy. In this case, the issue is whether the attribution of the *définition en logique* (definition in logic) in the great Encyclopaedia of d'Alembert and Diderot to Jean-Henri Samuel Formey, a German pastor of French Huguenot origin can be justified. After looking at whether Formey could have written, translated or edited the article, the author concludes that even if Formey did indeed provide material to Diderot it was not his text that was finally used.

This initial set of chapters discusses works aimed essentially at scholarly audiences, the contribution entitled *Project Cleveland* by Alenka Vrbinc, Donna Farina & Marjeta Vrbinc looks at the lexicographical production of Slovenian immigrants in the USA in the 20th century. The paper looks at several bilingual Slovenian-English and English-Slovenian dictionaries in order to ascertain their pedagogical role in the rapid language assimilation of this immigrant community. The paper provides a valuable insight into dictionary use, and also into the arrival and integration of Slovenians into the New World. It provides important

input as to how dictionaries can be developed to handle the language needs of very precise communities in particular historical contexts. Whilst it is fascinating from a historical perspective, there are also important lessons to be learnt about how immigrant communities achieve integration through language whilst maintaining their own cultural specificity.

Remaining with the Slovenian language, Alenka Jelovšek takes us far further back in time to 16th century Slovenian through *The Incorporation of Proper Nouns of Non-Slavic Origin into the 16th-Century Slovenian Literary Language*. The previous chapter sees Slovenians adapting to a new and emerging cultural environment, but here we see how dictionaries adapt proper nouns from other cultural environments, including different scripts, into the Slovenian language. Nowadays ISNI<sup>3</sup> assists modern scholars in handling variations across languages, but begs the question as to how they enter a language, which is what is handled in this text through the theory and practice of proper noun integration through the development of a Slovenian literary language. The texts, often religious in character, allow an analysis as to how Semitic, Romance language and German were brought into Slovenian allowing the move from a largely vernacular language to a written literary one.

The chapter entitled *Dictionnaires manuscrits dans l'histoire de la lexicographie croate : Des recueils de mots aux trésors linguistiques et culturels* by Ivana Franić deals with early Croatian manuscript dictionaries. As with the previous chapter, we see the essential role of dictionaries in the evolution of a language through the work of the Dubrovnik circle. The dictionaries discussed are a series of bilingual works which are in a manuscript format which the authors now bring to the attention of a wider public.

By the end of the 17th century, France had already a well established tradition of lexicography, but one hampered throughout the century by the monopoly of the *Académie française*, which forbade the publication of any monolingual dictionary of the French language until, and for twenty years after, the publication of their own dictionary, which took over sixty years to see the light of day. There were two serious breaches of this monopoly, that of Pierre Richelet published in Switzerland in 1680 and that of Antoine Furetière's *Dictionnaire universel* which, after a notorious scandal in which the Academy endeavoured to prevent its publication, was published posthumously in Rotterdam in 1690. The paper *Évaluer la dette : l'étendue de la présence de Richelet dans le Dictionnaire universel de Basnage 1701* by Clarissa Stincone, working within the context of the BasNum project, deals with the second edition of the *Dictionnaire universel* which had been totally

---

<sup>3</sup><https://isni.org>

revised and updated by Henri Basnage de Beauval in 1701. Working within very tight deadlines, Basnage was a brilliant compiler making use of many specialised works. Stincone posits that Basnage owed a great debt to the second edition of the Richelet for language and grammar related entries to which he failed to give full justice.

If Basnage had recourse to many specialised sources, one of these was Félibien, whose role in the development of terminology relating to art is discussed by Rosa Cetro in her paper *De Félibien à Boutard : l'évolution du dictionnaire artistique entre le XVIIIème et le début du XIXème siècle*. This paper charts the development of artistic terminology from the *Des Principes de l'Architecture, de la Sculpture et de la Peinture et des autres Arts qui en dépendent* of Félibien to that of Boutard, the *Dictionnaire des arts du dessin, la peinture, la sculpture, la gravure et l'architecture*. As the author had already digitised the work of Félibien (Cetro 2022), she was able to trace words in the non-digital works based on a sample drawn from the letter C. The evolution of this type of dictionary also reflects changing audiences from a restricted élite to a wider public for whom the history of art is a factor. The challenge facing the author is also the width of the fields covered from architecture through painting and sculpture, which requires looking at the terms from a terminological as well as a lexicographical point of view.

A link to Henri Basnage also appears in the text of Georgios Kassiteridis who looks at *La valeur pragmatique des langues dites « orientales » dans le Dictionnaire universel de Trévoux*. The so-called *Dictionnaire de Trévoux*, a title the Jesuits only accepted much later, was a plagiarised version of Basnage's 1701 *Dictionnaire universel*, which first came out in 1704. Published in the city of Trévoux by an anonymous group of Jesuits, the work only differed from the original by the addition of Latin equivalents and the rewriting or removal of usage that intolerant Catholics might see as heretical because linked to Protestantism. Their initial aim was simply to undermine the market of the Protestant publishers based in the Netherlands. The edition under discussion here is the second one, which was already moving away from being a pale copy to an encyclopaedic work which would later try to attack the great *Encyclopédie* of Diderot and d'Alembert, which the Jesuits also saw as heretical. The work of George Kissiteridis looks particularly at the introduction of Hebrew, Arabic, and Syriac, which at the time were termed "oriental languages". As the author points out, this 1721 edition demonstrates an increase in religious terminology, but also reflects input from other fields such as geography and botany. However, what is studied here is the large increase in the use of non-European languages. The notable inclusion of Hebrew demonstrates the clear influence of the theologian and bible critic, Richard Simon, who had always tried to hide his involvement in this work. Simon's knowledge

of the bible and his mastery of Hebrew and other “oriental” languages, explains an improved etymology of bible-related language and the increased presence of Hebrew.

Agata Pawlina’s chapter brings an interesting change of perspective as we move from “oriental” language in a French dictionary to the use of Italian and Greek in a Turkish lexicography, in this case Ottoman Turkish. Turkey has a unique place in Europe in being an interface between cultures of East and West, but also importantly in its covering much of the Roman Byzantine provinces and hence with a strong cultural contact with Italian and Greek. This work is based on a corpus of ten Ottoman Turk dictionaries. Her work is an exciting multidisciplinary study based on a strong knowledge of both music and its terminology and the history of the Turkish language, which allows a full analysis of the socio-cultural aspects. The result is a fascinating insight into Turkish lexicography.

The following chapter, *Exploring the unique method for encoding sinograms in the first known Chinese-Polish dictionary* from Andrzej Swoboda is again an interface between very different languages and cultures, but in this case through a single bilingual dictionary of Chinese-Polish. The work under consideration is from 1936, *Słownik chińsko-polski do czytania tekstów chińskich bez przygotowania*, and represents the first attempt to create an interface between these two languages. With reference to earlier studies, it analyses the challenges faced by the author of the dictionary, Doman Wieluch, and the means he employed to overcome these difficulties and structure his dictionary.

The chapter on *Les travaux lexicographiques de Carlo da Castorano et ses tentatives pour faire imprimer un dictionnaire européen de chinois* by Mariarosaria Gianninotoa & Michela Bussotti continues with the European language-Chinese interface, but takes us back to the 17th–18th century with work by an Italian Franciscan, Carlo Orazi da Castorano, and his difficulties in publishing a trilingual Latin-Italian-Chinese dictionary. The work was not published, as European printers were not yet handling Chinese characters well, but the attempts of da Castorano mark a change in the handling of non-European languages. The chapter traces the origin of dictionaries of Chinese by missionaries from the 16th century up to the work of da Castorana. Choices inevitably meant which Chinese, notably Mandarin, and the difficulty in finding suitable equivalents. It then goes on to the trials and tribulations of getting the work published in print, a task in which da Castorano ultimately failed.

The chapter by Mirosława Podhajecka on *The bilingual dictionary as a mediator between West and East: The beginnings of English-Polish lexicography*, yet again a manuscript dictionary, this time from the 18th century. Amongst the interesting factors is the date, which underlines the influence of French in Eastern Europe

at a time when English was already becoming the dominant language of trade. A further twist in the tail is that analysis shows that the author based much of his work on a series of English-French dictionaries, including Boyer's dictionary of 1699. The article begins with the historical context of a large central European country that finds itself divided up amongst other powers, leaving Polish as a national language without a country. It also explains the dominance of French and German in the educational system, leading to a lack of interest in the English language. It then describes the manuscript of the anonymous *Do Słownika Angielsko-Polskiego Dodatek ktorego układanie przedsięwzięte zostało Roku 1799 Dnia 17 czerwca* [An Addendum to an English-Polish Dictionary Whose Compilation was Undertaken on 17 June 1799] before considering issues relating to the compilation of this work.

The chapter *Lexicon Lapponicum Bipartitum.....ungarice scriptum: Hungarian aspects of North Saami dictionary writing* by Ivette Keleman deals with the very different language combination of Northern Saami, spoken in Norway, and Hungarian, both of the Finno-Ugric group. The historical context is of great importance as missionaries sought to convert the Lapps and as scientific exploration also took place. This paper thus describes the work of two Jesuit astronomers, Maximilian Hell and János Sajnovics, their discovery of the affinities between the two languages and their creation of a dictionary using Hungarian spelling conventions. It provides a history of Saami lexicography and the creation of a new writing system.

Remaining with Hungarian, Gábor Tillinger looks at French-Hungarian dictionaries in a paper entitled *Les exemples dans les dictionnaires français-hongrois à travers les siècles*. Starting from 1844, Tillinger takes us through the microstructure of dictionaries with an emphasis on examples. In a series of highly used verbs. The collection under study represents a series of large general language bilingual dictionaries to the exclusion of smaller specialised works. It then traces the choice of examples in these works and points to the particularities engendered by comparing two very different language structures. Covering a period from 1844 to 2007, examples show clearly the socio-cultural contexts.

Michela Murano takes us to the comparison of two French-Italian dictionaries of the 19th century in her chapter entitled "*Sul finir d'imparare la Grammatica Francese, fa d'uopo studiar il Dizionario delle Frasi*": *deux recueils phraséologiques bilingues franco-italiens de la première moitié du 19e siècle*. The chapter, in French, looks at two phraseological works, the *Fraseologia italiana-francese* (Trucchi 1825) and the *Dictionnaire des phrases françaises-italiennes* (Lanza 1837). In 19th century Italy, the teaching of French affected all levels of society and works like those under study sought to approach grammar through phraseology. This is

very much French as a Foreign Language, currently an area that attracts many students in France, but more than a century before the current perspectives. This chapter takes us through the macro and micro structures of dictionaries that were revolutionary in their time, and which contain lessons for current practice.

As can be seen in this collection of papers, historical lexicography and lexicology is about legacy dictionaries and also about works that describe past usage of language. Thus Rafał Rosół takes us to the foundations of Western European culture with a dictionary concerning the world of Ancient Greek, the *Lexicon of Oriental Words in Ancient Greek*. The lexis discussed here covers a period of some two thousand years from the Mycenaean period to the 6th century AD, also termed Common Era or CE. The language covered concerns the incredible melting pot of civilisations in the Middle East up to the end of the Roman Empire. In many ways, this chapter shows the total relativity East meets West, because the question is as to whose East and whose West. The Greeks placed themselves in the centre and regarded the remainder of the world as barbarians, that is, non-Greek speakers. We are in a constant continuum of centre as we cross the world, which is why a historical approach to language is so important.

The final chapter, *The discovery of a Russian-Tajik dictionary* by Abdusalom Mamadnazarova & Bahridin Navruzshoev, raises several issues which include the East to West continuum where Tajikistan is at the cross-roads between what we term Europe and what we term Asia, and the key issue of forgotten lexicographers that has pervaded this entire work. In this case, the author, Sh. Khatymtayev, is known, as is the date of publication; it is simply that the dictionary remains only in a fragmentary form. After a brief discussion of the history of Tajik dictionaries, the paper looks into the structure of this forgotten work and its equally forgotten compiler.

Dictionaries are tools of mediation across time and cultures. They are windows on the past. Historical lexicography and lexicology are vibrant, dynamic disciplines that open these windows to allow us to better understand the development of lexicography and the world in which we live. This collection of papers gives insights into a variety of worlds with shifting centres as languages evolve over time. East to West across Europe, East to West across the world.

## References

Accademia della Crusca. 1612. *Vocabolario degli accademici della Crusca*. Venezia: Giovanni Alberti. <http://www.accademiadellacrusca.it/en/digital-shelves/crusca-online>.

- Cawdrey, Robert. 1604. *A table alphabeticall*. London: Edmund Weauer.
- Cetro, Rosa. 2022. *La démarche terminologique d'andré félibien. La systématisation du lexique artistique en français*. Torino: L'Harmattan Italia.
- Estienne, Robert. 1539. *Dictionnaire françois latin contenant les motz et manières de parler françois, tournez en latin ([Reprod.]*) Paris: R. Estienne. <http://gallica.bnf.fr/ark:/12148/bpt6k505878> (22 November, 2017).
- Lanza, Nicolas. 1837. *Dictionnaire des phrases françaises-italiennes*. Napoli: Seguin.
- Trucchi, Giuseppe. 1825. *Fraseologia italiana-francese*. 2nd edn. Napoli: Seguin.
- Yong, Heming & Jing Peng. 2008. *Chinese lexicography: A history from 1046 BC to AD 1911*. Oxford: Oxford University Press.

## Chapter 2

# On closure and its challenges: Examining the editors' proofs of OED1

Lynda Mugglestone


Pembroke College, University of Oxford

This chapter examines the newly discovered Editors' Proofs of the first edition of the *Oxford English Dictionary*. Long considered lost (with the exception of a single volume in the OED archives at Oxford University Press), these were a serendipitous find in the Bodleian Library in 2021 where they were filed under the heading "Amusements". Signed and dated, as well as richly annotated, they offer a complex narrative of the final stages of the making of the OED, documenting the working practices of the different editors, and presenting a heteroglossic space where the interactions of editors, assistants, and press workers reveal a unique community of practice, united – and at times divided – by the different agendas at stake. On one hand, these offer compelling insights into the material historicity of the original text, where pen, ink, and the demands of hot metal can reveal an uneasy symbiosis. On the other, the proofs emerge as a site of new historical inquiry, marked by the visible shaping of the final stage of the text in which e.g. the demands of in-text justification can drive a range of conflicted decisions while the salience of substitutive revision – by which any addition must be compensated by an equivalent process of loss – impacts in critical ways on the presentation of evidence, labelling, semiotic analysis, as well as entries as a whole.

### 1 Introduction

*Closure* ('a bringing to a conclusion; end, close'), edited by James Murray, appeared in the fascicle *Clo–Consigner* in 1891 as part of the serial publication of the *New English Dictionary on Historical Principles* (henceforth OED1) (Murray et al. 1884–1928). The opening fascicles *A–Ant*, and *Ant–Batten* had been published, respectively, in 1884 and 1885. If Murray had been appointed as editor on



Lynda Mugglestone. 2025. On closure and its challenges: Examining the editors' proofs of OED1. In Geoffrey Williams, Mathilde Le Meur & Andrés Echavarría Peláez (eds.), *West meets East: Papers in historical lexicography and lexicology from across the globe*, 11–29. Berlin: Language Science Press. DOI: 10.5281/zenodo.15394471 

the assumption that the dictionary would be completed in ten years, closure remained elusive. Henry Bradley (appointed as a second editor in 1886), produced his second independent fascicle, *Everybody–Ezod*, in 1894 (three years after *E–Every*). Ten years later, Murray had advanced to *O*, Bradley was at work on *M*, and another editor, William Craigie, promoted in 1901 from his earlier role as dictionary assistant, had completed work on *Q–Ree*. Only after another decade did closure start to seem a realistic proposition. “We may... reckon that the end of 1916 will see the *Dictionary* finished”, Murray wrote to Walter Skeat in 1912: “If I live to then, I shall be 80, and it will also be my Golden Wedding: Let us hope that the Grand Conjunction of all these cycles will really take place”<sup>1</sup>. A flurry of activity ensued. *Su–Subterranously*, the first independent fascicle by a fourth editor, Charles Onions, appeared on New Year’s Day 1915. This was followed, if discontinuously, by William Craigie’s *Spring–Squoyle* in April, and Bradley’s *Standard–Stead* in October. Murray’s *Trink–Turn–down*, a fascicle published in July 1915, was, however, to be his last. He died on July 26<sup>th</sup>. *T* was completed by Craigie, with the assistance of Murray’s Scriptorium staff, as well as Murray’s extant drafts. Closure, for the *Dictionary*, eventually took place in 1928. Craigie and Onions were, by that point, the only surviving editors.

A richly documented history of OED1 (see e.g. Brewer 2007, Gilliver 2016, Mugglestone 2005, Murray 1977) traces this lengthy gestation, aided by substantial archival collections in both Oxford University Press and in the Bodleian Library, Oxford. This article, however, will focus on the contributions of a newly-discovered resource, the final “Editors’ Proofs” of OED1, found by chance in July 2021 in the Bodleian Library where it had been tagged under the heading “OED. Amusements”. “Every so often we get lucky, and a long-lost work turns up” writes Lynch (2016: 158) in a chapter on lost linguistic projects and the fragility of the historical records we possess. Lynch’s words were prompted by the discovery, also in the Bodleian Library, of James Boswell’s manuscript dictionary of Scots. The Editors’ Proofs reveal similar fortuity, as well as undeniable salience in bibliographic as well as linguistic terms. Long considered lost (apart from a single volume preserved at Oxford University Press), they present a vast and hitherto unexamined resource, spanning twenty-three volumes from the mid-1880s to 1928. Documenting the penultimate versions before each fascicle went to press, they shed new light on a text on the very point of closure, alongside the sustained challenge of shaping definitive histories on the page. Pasted “slips” containing new material as well as in-text epistolarities that, in various ways, negotiate matters of both form and inclusion are common. Marginal asides, in a

---

<sup>1</sup>MP/22/7/1912. James A. H. Murray to Walter Skeat. Murray Papers, Bodleian Library, Oxford.

variety of hands, create an intriguingly heteroglossic text. The Proofs are replete with annotations, deletions, and revisions at a range of levels.

As this suggests, the Editors' Proofs exist as a text of complex materiality, making plain not only the practical difficulties of fixing the text but the role (and significance) of different communities of practice as part of this process. As such, they present a useful corrective to the familiar historical alignment of OED1 with the history of Murray himself – a consequence both of archival materials (see e.g. the *Murray Papers* in the Bodleian Library) as well as popular print history in which *Murray's Dictionary* was early established as an alternative title. “The college to subscribe to James Augustus Henry Murray’s *Dictionary of the English Language*”, as the Chapter Book of Corpus Christi College (CBCCC)<sup>2</sup> recorded in 1890 (“Chapter Book”, 22). That Walter Worrall, employed as an assistant on the *Dictionary* since 1885, was the *de facto* editor of *W* after Bradley’s death in 1923 is, for example, unambiguous in the Proofs. So, too, are the decisive contributions of, say, Charles Balk who worked on the *Dictionary* for some forty years, initially alongside Murray (and later with Onions). Across the Proofs, the surviving records of in-text correspondence between different editors and dictionary assistants, as well as the contributions of press-room and compositors, vividly document a range of alternative histories behind the finished text.

## 2 Examining the text in time

“The great dictionaries” of the past “were created using basic technologies; pen, paper, and index cards for the lexicography, hot metal for the typesetting and printing” write Michael Rundell and Adam Kilgariff (Rundell & Kilgariff 2011: 258). This sense of print history is palpable in the Editors’ Proofs. References to the “foundry” (“To be sent to the foundry at once”, as the corrected sheet of *contemplation* declares), or to the need to “impose for Electro” (‘used *colloq.* as an abbreviation for [...] ELECTRO-PLATE’, as Bradley explained in the relevant fascicle) are common. Press-room warnings that “type wanted” (such that the text must be finalized, printed off, and type released for further sections) reveal other imperatives of historical production. Apparent, too, is the occasional striking instance of “printers pie”. “Please have this page reset, they have had an accident and dropped the matter out before moulding”, as an in-text missive written alongside *devil* records in March 1895. The final column was immaculate. The “matter” of the previous two was missing, presumably scattered on the composing-room floor.

---

<sup>2</sup>[https://www.corpus.cam.ac.uk/sites/default/files/downloads/chapter\\_book\\_8.pdf](https://www.corpus.cam.ac.uk/sites/default/files/downloads/chapter_book_8.pdf)

Rundell & Kilgarriff (2011) nevertheless describe a far more binary process of historical composition than is presented by the Editors' Proofs. On one hand, they suggest, is the "lexicography", vested in pen and ink in the primary collection of data. On the other, as they indicate, was the press, and the technology of hot metal on which it relied. The Editors' Proofs, however, routinely place lexicography in pen and ink as an ongoing process – as revision in proof, and in symbiosis with "hot metal" to the last possible moment. As in-text annotation confirms, for example, the extent to which lexicographical desiderata could be accommodated before closure of this kind was a recurrent concern. Hand-written injunctions such as Bradley's "If it is not too late, I should like the enclosed correction to be made" (here with reference to a last-minute revision of the etymology of *lanchara*) or "Please make this correction and addition if in time" (a comment written by Murray above *to-queme* in a change that affected both derivation and cognate forms) routinely reveal the provisionalities (and hazards) of late-stage change.

The iterated use of "if" claims particular significance in this respect. Murray's "Please alter this if it can be done" (under *dragonate*) or, say, "I shall be glad if the enclosed alteration can be made on the latest final", written by Craigie alongside the entry for *quandary* in January 1902 (by which an earlier attributive sense was to be removed) stand as other marginal annotations, directed to the press-room, and in which the pragmatic contingency of emendation is overt. Balk's "If it is not too late (i.e. the sheet printed off) we should like to have this correction made, as the original statement is incorrect", inscribed alongside *discipline* in the Editors' Proofs (and counter-signed by Murray), is, even so, particularly illuminating in the detail it provides. On one hand, the proposed emendations were self-evidently important for the authority (and authoritativeness) at which OED aimed. On the other, the information provided might, as Balk recognised, remain a casualty of production, and the processes it required. "If not too late, please make corrections marked", as another annotation by Balk, here against *obmit*, records: "if too late i.e. the pages already electro'd, please return".

The text was, in this light, routinely shaped by its own unfolding history even as it endeavoured to tell the full history of words. In-text annotations such as "needed for the date" (placed alongside *disc* (n.) where Murray wanted to include lengthy if last-minute citational evidence from the *Engineer* of 1856) can, for instance, set out an intentionally persuasive rationale for the changes to be made – even if, of course, the individual compositor might think otherwise given the resetting and laborious readjustment that was, in turn, required. Bradley's use of "small", in an in-text missive pinned to *fight* in 1895 ("Will you kindly make the small corrections indicated on the enclosed cuttings from the pages") was

equally tactical. His further comment (“I fear some of the pages in question will have been printed off, but the alteration might be made in the plate so that subsequent copies printed may be correct”) arrestingly suggests the process of in-text change even as the text was printing, and by which variant versions might be incorporated in a single print run. Seen in hindsight, the decision, made in Proof, to delete the evidence and entry for *final proof* as a further aspect of late-stage change acquires a sharply ironic charge, “*Printing*. The last proof sheet before sending to press”, the relevant sheet still records. “Final”, as other comments confirm, could be a highly relative term.

Visibilities of this kind vividly inscribe the complex realities of collaborative and multi-author composition while reminding us of the gaps that, as part of late-stage emendation, could intervene between lexicographical intent and the text as set. A surprising number of revisions remain, for example, within the confines of the Editors' Proofs where editorial “last words” have, for whatever reason, failed to be reproduced in print. Craigie decided, for instance, to undertake a stringent process of revision in the fascicle *Unright-Uzzle*, deleting the entry (and evidence) he had previously included for *unsonned* and adding new evidence (and a new reference) to *unstoicized* instead. A new quotation from Robert Wallace's biography of the Scottish historian George Buchanan was carefully written in the bottom margin: “Had he been all his detractors call him, that would not have \*unstoicized him”. At the same time, the previous evidence, from George MacDonald's *Weighed and Wanting* of 1882 was crossed through (“~~Determined that his son, having \*unsonned himself, should no longer be treated as a son~~”). Yet neither change made its way into print or, indeed, into later iterations of the OED. Even in OED Online, *unsonned* remains in place while *unstoicized* remains unrecorded and unknown.

The final revisions of *clitoris* (n.), retained in full on the Editors' Proofs (see Figure 1), provide an even more striking example, problematizing the standard narrative on this contested word in OED history. Edited by Murray in a fascicle published in 1889, its definition as ‘a homologue of the male penis, present, as a rudimentary organ, in the females of many of the higher vertebrata’ is, as by Rees & Stedall (2019), prototypically seen as a product of Victorian reticence as well as androcentrism. “The iconic English dictionary's misogynist descriptions were ripe for revision”, they assert. In the modern OED, a new definition (‘The female genital organ located in the anterior part of the vulva, which contains numerous nerve endings and plays a major role in sexual arousal and pleasure’) articulates meanings which were hitherto obscured or denied, and in which new interpretative trajectories in relation to female sexuality displace the hegemonies of the past.

The Editors' Proofs, however, provide their own contested readings of this entry in ways which affect both evidence and definition. In one version of the surviving text, dated October 1889, Murray's definition had already been set. It drew closely on Thomas Huxley's *Manual of the Anatomy of Vertebrated Animals* of 1871 (from which Murray's final illustrative citation also derived). "In the female sex, the homologue of a penis frequently makes its appearance as a *clitoris*", as Huxley (1871: 111) had explained. Yet, as Figure 1 illustrates, pasted palimpsestically to the same page in the Proofs is an alternative entry in which evidence and definition assume a different shape. Murray's final citation ("In some few mammals (e.g. the *Lemuridæ*) the clitoris is traversed by a urethral canal"), included from Huxley in the original version, was now displaced. Evidence from Todd's *Cyclopaedia* ("The clitoris lies in the upper part of the vulvar fissure") takes its place. Redefinition meanwhile removed the male frame of reference, and the further link to Huxley, even if a certain euphemistic coding remains in place: "Part of the female generative organs in the mammalia", the revised entry explained. "Please make correction to final", states a neat annotation by Balk above these revisions. In effect, all subsequent editions, from the fascicle of 1899 to the reprinted text of 1933, as well as OED2 (Simpson & Weiner 1989), have been based on a version that had already been rejected.

Precisely why the published text assumed the form it did nevertheless remains unknown. Balk was, by 1889, an experienced co-worker whose notes on the Proofs routinely reflect the consensus of Scriptorium (and Murray) on the changes to be implemented. Evidence from other sources, as in the compositorial walk-out that attended the printing of Farmer and Henley's *Dictionary of Slang*, or indeed in contemporary Philological Society discussions (whereby it was decided that entries for *cunt* and *condom* should be suppressed in OED1 (see e.g. Mugglestone 2007: 1–22), clearly suggest, however, the presence of extra-textual pressures by which the immediacies of print history could be shaped. Other evidence in the Editors' Proofs readily confirms the effect of related sensibilities on the matter of cultural propriety and linguistic representation, evident in e.g. the tactical elision within *squitter* ('To ~~have or suffer from diarrhea~~; void thin excrement') just before it went to press or the deletion from *lap* of "Lap it hot about the privy parts" (a citation from Gervase Markham's *The English Housewife* (1615)) while equivalent if less indelicate material from Moxon, Goldsmith, and Swift in the same entry remained intact.

Murray's careful rewriting (and decorous code-switching) in sense 20 of *cock* (= 'penis') some fifty pages after the disputed entry for *clitoris* presents a further instance where contemporary socio-cultural taboo arguably reshaped the form it later assumed in print. Murray's earlier version – "Only in low vulgar use,

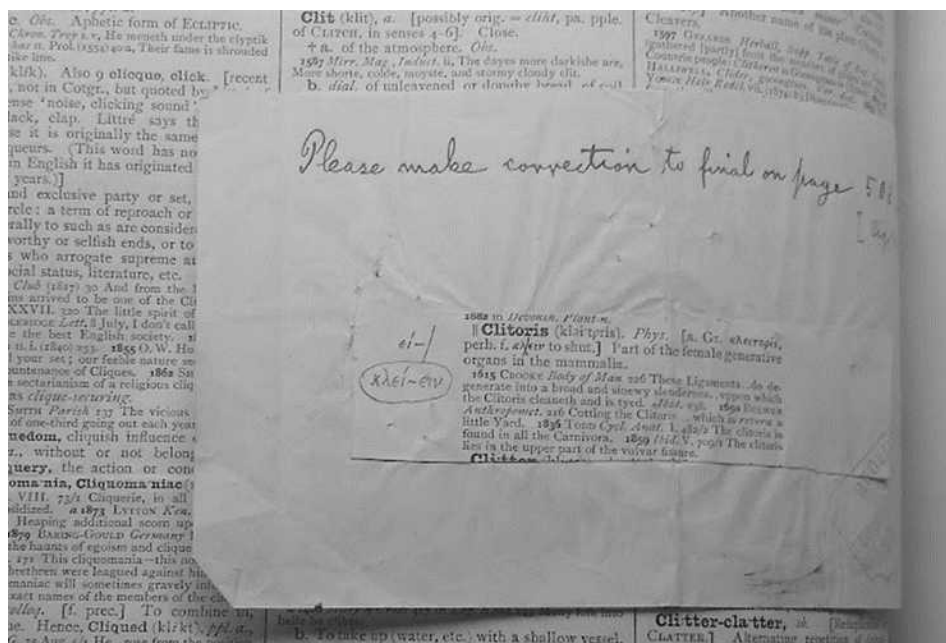


Figure 1: Editors' Proofs: revised entry for *clitoris*, November 1889. Reproduced with permission of the Bodleian Library Oxford from *A New English Dictionary on Historical Principles* (annotated proof text) 30255 c. 17 (2), p. 508.

and considered indecent”, with its embedded acknowledgement of contemporary language attitudes – was edited out. “*Pudoris causa*, inadmissible in polite speech or literature” (complete with Latinate indirection) took its place, courtesy of a series of in-text annotations in Murray’s hand. This, too, documents a history of change and revision that has hitherto remained hidden in the ancillary histories that the Proofs provide. We don’t, of course, have a variorum edition of OED1. At a range of points, however, the Editors’ Proofs suggest that, in whole or in part, this might constitute an interesting and, indeed, productive resource.

### 3 Rereading the text in Proof

How to use data of this kind nevertheless brings its own challenges. Auden’s skepticism about the value (and insight) of academic scrutiny with reference to late revision was palpable. Revisions he made while preparing his *Collected Shorter Poems* for the press in the early 1960s were, as he wrote to Neville Coghill, “a gift to any anal-minded American Ph.D. student” (Crook 1990). Nevertheless,

as with pre-print versions of Sally Rooney's *Beautiful World, Where Are You?* (published in 2021), such anterior texts have, in recent years, often claimed considerable attention in relation both to their monetary cost (Strzyżyńska 2021) as well as the critical exegesis they might afford. Pre-print text with annotations, serving to highlight e.g. revisions between different versions or which are seen as indicative of a writer's later thoughts, are typically regarded as even more collectible. Proof revision or the existence of an emended text in galley proofs can, in similar ways, claim a valued materiality, especially in an era of born-digital texts where the processes of writerly revision and rethinking are all too easily elided by the affordances of new technology. The unexpected recovery of the Editors' Proofs of Murray et al. (1884–1928), with all their complex materiality, will not in future years be paralleled by those of OED Online.

Revision in proof, and the survival of texts of this kind, can, for example, provide us with useful testimony of what Lillis (2013: 103) describes as the text trajectory of a particular work – in other words, with historical evidence of its contextualization, its coming into being as text – and, by extension, of what is, or not, suddenly deemed expendable. Proofs and page proofs, as Hannah Sullivan argues in her book *The Work of Revision* (Sullivan 2013), can often be seen as constituting a site of continued writerly experimentation, complete with visible evidence of re-composition, whether this involves the rewriting of a particular section or phrase or wider patterns of restructuring. On one level, as Sullivan comments (Sullivan 2013: 4), the process of late-stage revision clearly signals “the possibility of the text being fixed in some material form”. On another, as she adds, evidence of this kind also remains profoundly evocative of the dialectic between “what changes and what stays the same” – and, crucially, why this might be so.

Sullivan's critical focus is, of course, directed to literary rather than linguistic exegesis in which surviving proofs and their embedded revision by writers such as Henry James and Virginia Woolf are explored as “classic documents” of high modernism. If the text in proof occupies, as she argues (Sullivan 2013: 6), a “materially intermediate” status (“neither manuscript nor print”), it is also made reflective of the on-going “friction between possibility and constraint”, and a locus of new insights that encompass both the status of the published text, and the interpretative trajectories that anterior in-text annotation reveals. “High modernism”, not least given the OED's beginnings some forty years before Victoria's death, might seem remote from the critical debates that inform historical lexicography. There are nevertheless some illuminating correspondences in the visible tensions of fixity and flux, and “friction” and “constraint”, that the Editors' Proofs

variously attest – alongside the decisions that the final and published text can, in turn, reveal.

The status of Editors' Proofs per se has, in this respect, particular salience. They exist, by definition, at a point of "almost closure" or "almost finality" – a liminal state where, as contemporary print manuals affirmed (see e.g. Ford 1854, Graham 1848), only minor emendations were expected. Any whole-scale revision was disfavoured, especially if it threatened to disturb the pagination that had been carefully set up (and, importantly, in terms of the OED1, calibrated across different fascicles for which printing was, as we have seen, often alphabetically discontinuous). The fact that the Editors' Proofs drew on a pre-existing sequence of first and second proofs, third proofs when necessary, as well as revises and second revises, likewise meant any major processes of revision should already have taken place. Some remaining changes were, even so, obligatory. Each page was attentively read by the press reader whose oversight ensured the systematic correction of typography and font in a text where being "eloquent to the eye" (Murray 1888: vi) was not only aesthetic but functional, enabling the ready discrimination of different kinds of information. The OED, with its twenty-one contrasts of form and font, was, as Luna (2004) notes, particularly "structure-rich" in this respect.

Other forms of obligatory change are, however, also apparent. Transposition (often indicated by a marginal "tr.") was a surprisingly common injunction, triggered by the need to maintain dictionary order. The entries for *distraintment* and *distrainor* were, for example, awry until just before publication, as were those for *cradler*, *cradle-hood*, and *cradling*. Historically-motivated transposition (and the diachronic sequencing of evidence within a given sense or entry) claimed similar importance, especially given the historical principles on which the OED was based. "Should have been put in according to date", Murray firmly instructs alongside the entry for *beholding* where, in the Proofs, evidence from Shakespeare's *Merry Wives of Windsor*, dated 1598, unaccountably came after that from Henry More's *An Antidote Against Atheisme* of 1662. "[I] shall be glad if the reader will point out accidental derangements of date, such as this, when he sees them", another marginal note from Murray confirms. Even in emendations of this kind, the "friction between possibility and constraint" can be perceptible. Reading the Proofs of *D*, Murray had urgently written to alert the press-room that the entry for *denourishment* had been misplaced such that it followed *denunciatory* rather than appearing four pages earlier (and following *denouncing*). Yet, as an in-text and hand-written response from the press-room confirms, change was no longer possible. The relevant pages had already been cast. Even in the reprinted text of 1933, the disrupted order remains in place.

In general, however, emendations that aimed to secure an authoritative text had self-evident priority. Craigie, for example, rewrote *quick time* on a new slip pinned to the side of the relevant page on realizing that his earlier facts had been awry (it was not, he confirmed a “rate of marching [that] consists of 120 steps of 30 inches each (= 300 feet)” but “128 paces of 33 inches (=118 yards) or four miles an hour”). Bradley eliminated other late-stage errors under *lychnic*. As his deletions confirm, this was not “an office which accompanies the lighting of lamps, forming part of the vespers on certain occasions” but referred, with marked specificity, to “the introductory part of the vespers” alone. The text was, in both cases, reset accordingly. Early evidence for *flump* (“To fall or move heavily with a dull noise”), which had previously dated the first English usage to 1677, was likewise removed. “Found in the Bodleian today”, an annotation in the adjacent margin by George Sykes (one of the editorial staff) announced. The relevant word was, in fact, *slump* while the submitted evidence (and the entry up to this point) had been based on a misreading of long tailed <s> as <f>. *Flump* in English was suddenly postdated to 1816. “The perfection of the dictionary in its *data*” was vital, Murray (1880: 129) had early stressed. Here and elsewhere, rereading continued until the text went to press.

#### 4 Willed and willing change

Changes of this kind constitute what we might see as willing editorial revisions, impelled, if in different ways, by a sense of the language, and the intent to secure the linguistic (and historical) rigour on which the OED project was founded. Etymology, semantics, definition, and sense-history (as well as bibliographical information) can all be reworked to this end. Murray’s anxious revision (in Proof) of *toast-stand* shortly before he died is part of this. Was ‘a ~~stand~~ utensil for standing toast, etc. by the fire’ best, as a hand-written emendation in the margin records? Or ‘a stand for toast, etc. by the fire’? The former was crossed though. Bradley’s in-text revision of *glaciarium* (not a ‘skating rink with artificial ice’ but ‘a skating rink with ice artificially produced’, he decided) was similar; artifice was, rightly, relocated to the means of production rather than the ice itself. Across the Editors’ Proofs, Craigie proved a particularly inveterate rephraser changing e.g. ‘noodle, simpleton’ to ‘thickhead, dolt’ in revising *numskull*, and amplifying the already expansive definition of *narrow-minded* (‘lacking in breadth of mind; illiberal, bigoted, prejudiced’) by adding ‘incapable of broad views’. If the tensions of possibility and constraint thereby remain in view, so do related questions of where revision in proof could or should legitimately stop.

Nevertheless, as the annotated text confirms, each dictionary-maker and assistant also worked with the conscious knowledge that any new addition had a cost – not only monetary (though this had its own significance) but, more critically, a spatial cost, enacted within a textual economy that had to be rigidly maintained. A new page was not forthcoming. Space itself was a limited commodity in ways that yield not only other manifestations of constraint but processes of textual reshaping by which evidence and exposition can be forcibly recalibrated on the page. Antedatings and the emergence, in Proof, of new evidence on the first use of a given word or sense provide a useful text case in this respect. Linguistically, these had obvious salience; “quotations illustrating the first appearance” of a word or sense, and data on “a word’s first appearance in the language” had been identified by both Murray (1900: 46) and Richard Chenevix Trench (Trench (1857: 23) as primary requisites for a robust historical lexicography. Late-stage additions of this kind were pinned or pasted to a wide range of entries in the extant Proofs. A new citation from Mulcaster’s *Positions Concerning the Training Up of Children*, published in 1581, appeared, for instance, in the margin alongside *gay* ((vb.) ‘to make gay...to give a bright and pleasant look to’), here in ways that served to revise its early use by some sixty years. *Gather* and *coaxingly* provide similar examples. An antedating from George Gascoigne’s *The Steel Glass* (1576) secured a new date of birth for the former (versus a previous first usage in 1704). Likewise, under *coaxingly*, an antedating from Susannah Centlivre documented a life-history that now began over a century earlier than had previously been supposed. “Prithee, my dear, moderate the passion (*coaxingly*)”, as she had written in 1714 in *The Wonder: A Woman Keeps a Secret*.

Information of this kind served, at least in principle, to supplement text and historical record alike. Nevertheless, as the Proofs attest, the wider textual consequences can be arresting. Equally visible, for example, is not only the presence of newly-inserted citational evidence but a corresponding process of elision by which any material added must, given the immediacies of historical production, be accommodated by excising the equivalent space from elsewhere. As under *coaxingly* therefore, if Centlivre was added, a citation from Charles Lamb’s *Letters* in 1835 (“There was a rough earnest to the request though it was put *coaxingly*”) was removed while, in parallel ways, the decision to add Gascoigne under *gather* necessitated the loss of Colley Cibber’s *The Careless Husband* in a citation recorded from 1704. Mulcaster’s new-found evidence of *gay* yields a corresponding narrative of cause and effect. The relevant life-history of *gay*, sense 1, was now beneficially extended to the late sixteenth century. Nevertheless, a minor entry (*gay*, defined as an obsolete form of *go*) as well as earlier embedded references to *gayed* and *gaying* as subordinate forms were all deleted in response.

As the Proofs confirm, the process known as substitutive revision was widely adopted as a means of crafting late-stage lexicography on the page. As the antedatings discussed confirm, the text might therefore be supplemented, and the historical record improved. But gain could not take place without loss, nor insertion without excision.

Surviving in-text annotation can make this process of exchange strikingly overt. As under *convent* (n.), sense 6, a new slip was pinned to the left-hand margin, providing an antedating from Thomas Trusler's *The Distinction between Words Esteemed Synonymous*, dated 1795, and revising evidence of its earliest use (hitherto recorded from Stanley Lane-Poole's biography of Lord Stratford de Redcliffe in 1814). An explanatory note in Balk's hand, however, sets out the cuts and consequences which must, in turn, take place. As the press-room was instructed, if Trusler is added, a later citation from Guilt's *Architectural Glossary* (1842) was to be deleted to secure the requisite space. As this suggests, excisions of this kind were inherently opportunistic, driven by patterns of revision that play out at the local level of column or page. "If necessary to save line", Murray likewise wrote alongside *legitimacy*, for example, excising a quotation from the *Contemporary Review* in 1885 in order to include new evidence under *legister* instead. Similar was e.g. a new antedating from Nathan Bailey introduced under *demonographer* (which took usage back to 1736) while previously inserted evidence from the British traveller Rachel Harriette Busk was deftly truncated instead in order to accommodate the desired change.

An identical trajectory of gain and loss played out between Thomas Erskine and Jane Austen under *blowzy* (adj.), sense 2, defined as 'Of hair, dress: Dishevelled, frowzy, slatternly'. In the Editors' Proofs as originally printed, Austen's *Pride and Prejudice* (courtesy of Miss Bingley's description of Elizabeth Bennett) had supplied a prized first citation: "her hair was so untidy, so blowzy". Newly-arrived data, in an antedating from Erskine's *The Barber* of 1778, meant, however, that substitutive revision again took its toll, here in ways that displaced Austen and her evidence from the published text. Austen's *The Watsons* (1804) revealed a similar vulnerability under *full* (sense 1a). Bradley deleted Austen's "quite as full as ever...there seemed no vacancy whatever" under the pressures of space that were imposed by an antedating he introduced under *full*, sense 1e (referring to 'Of an animal: Pregnant. Of a fish: Charged with roe'). Changes of this kind might, as we have seen, be opportunistic but Bradley's choice can, even so, seem surprising.

The extent to which certain types of text seem more vulnerable to the forms of attrition that occur is, however, clearly of interest. Shakespeare in a new citation from *Timon of Athens* could, as under *tithed* ('Subject to, charged with, or

liable for the payment of tithes') extend his numerical dominance in the OED, courtesy of a citation written by hand in the upper margin of the relevant page. Yet while the entry for *tithed* was adjusted accordingly (its "birth" moved backwards by almost 250 years), at the other end of the entry a different process was taking place such that evidence from the *Manchester Evening News* in 1884, referring to 'The heavily rented and tithed British farmer', was crossed though, having been judged expendable against Shakespeare's greater claim. Craigie's insertion of a new citation from Wordsworth's *Prelude* in 1805 mid-way through *quartered* ('Through three divisions of the quartered year') can likewise be set against the corresponding loss of the *Daily News*, in final position, in what had originally provided a far more modern citation from 1900. In a related process of exchange, a decision to include a new citation from Blackstone mid-way though *tour* (referring to the tour or circuit made by the sheriff of a county) was clearly facilitated by the selective loss of the *Times*, in a 1904 citation that had previously concluded the previous entry for *tourmaline*. Likewise, evidence from the *Leeds Mercury* in 1887 disappears under *fleeting* (here, as prompted by the decision to include an antedating from Sir Thomas Smith), a citation from *Harper's Magazine* in 1887 disappears under *burgraviate* (sb.), displaced by one from 1762, taken from Patrick Murdoch's translation of *A New System of Geography*, and a citation from the *Nation* in 1892 disappears under *give* while adjacent evidence from Scott, Thackeray, and Wilkie Collins remains intact. "Please make this correction if not too late", an accompanying note from Bradley confirmed. While the canonical and gender prosodies of OED1 have long attracted comment (see e.g. Mugglestone 2005, Brewer 2012), the process of late-stage revision in the Proofs suggest other productive avenues for investigation in this respect, not least in the cuts and consequences that are revealed.<sup>3</sup>

In principle, of course, evidence across OED1 was treated with impartiality. "The literary merit or demerit of any particular writer, like the comparative elegance or inelegance of any given word, is a subject upon which the Lexicographer is bound to be almost indifferent", as Herbert Coleridge, the *Dictionary's* first editor, had affirmed (Coleridge 1858: 2). Canonical and non-canonical writers were targets of substitutive revision in the Proofs; if a new quotation from Longfellow was inserted under *birdlet*, it was, for example, Keats who was deleted under *bird-lime* in order to make room. Even so, the relative patterns of loss and gain in relation to e.g. inserted evidence from Scott are striking. As under *name*,

---

<sup>3</sup>The transfer of the data from the Editors' Proofs to a database (funded by the John Fell Fund at Oxford University) will, as part of the next stage of the project on which this paper is based, enable far closer and systematic scrutiny of patterns of this kind.

Scott is added, both in a personal letter and in a citation from Lockhart's *Life*, while evidence from the minor Anglo-Caribbean writer Annie Holdsworth (in a citation from 1890) is taken away. In related ways, in a textual battle for space between Surtees and Scott inscribed in the entry for *bartizan*, Scott (and a citation from *The Eve of St. John*) clearly wins. So does, under *neighbour*, a new mid-entry citation inserted from Scott's *The Fortunes of Nigel* (1822) which preferentially displaced the pithy "Neighbours especially are often far from neighbourly", as previously recorded from Charles Spurgeon in 1874. Under *qualifying*, too, a similar process can be seen at work such that a new citation from *Rob Roy* in 1817 drove the loss of a mid-entry citation from Worcester under the adjacent *qualify*; under *dine*, a new citation from *Guy Mannering* displaced one from the *American* in 1884. Under *spinster*, new evidence from Scott's *Rob Roy*, inserted in final position, likewise triggered the compensatory clipping of evidence that, under the adjacent *spinous*, had hitherto been recorded from the *Proceedings of the Berwickshire Naturalists' Club* in 1854. After Shakespeare, Scott was, Brewer (2020) records, the most cited individual writer in OED1. This was, self-evidently, reinforced rather than reduced in late-stage editing of this kind. On one level therefore, substitutive revision provides valuable evidence of the recomposition of the text under constraints of space. But, on another, it also makes particular aspects of editorial choice open to our view. Each editor (and editorial assistant) had, in reality, some latitude as to where the revising hand might fall.

The iterated vulnerability to attrition of e.g. citations from periodical literature, from non-UK writers, and minor writers more generally when aggregated across the Proofs can thereby both document the fact of wider patterns of representation than later appeared in the published text, as well as the praxis by which such representativeness was reduced. As in Craigie's re-editing of *nameless*, for example, evidence from Pope, Cowper, Tennyson, and Ruskin was, expansively, retained in full while, in order to satisfy the need for space, a citation from Adelaide Procter's *Legends and Lyrics* (1859) was instead reduced to the sharply truncated "Over [...] a nameless grave". In parallel ways, Edward Bulwer Lytton was clearly deemed of greater salience under *bulky* than the evidence (from the *Scotsman*) that Murray now deleted under the adjacent *bulking* while, under *object*, it was Frances Anne Kemble's *Journal of a Residence on a Georgia Plantation* – now recognized as a key text within the literature of American slavery – that was deleted (also by Murray) while evidence from J. H. Newman and Mark Pattison remained intact. Such instances can be multiplied across the extant text. Editing was, as the Proofs widely confirm, a process in which push routinely came to shove under the conditions of substitutive revision. Whether, given adequate space, such cuts would have been made is doubtful.

## 5 On being too long

No single change had, in reality, a monopoly over substitutive revision and the impacts it had. Etymology, sense-history, and definition (as well as citational range, as we have seen) can all be affected. As in the reshaping of *dweomercraeft* (an obsolete word signifying 'Jugglery, magic art') and *dwelth* (an obsolete word for 'Error, delusion'), both evidence and entries are visibly renegotiated on the page. "1 line more", as a comment in the margin attests, was necessitated by the addition of new etymological detail for *dweomercraeft*. "1 line less" was, however, secured by cutting equivalent detail under the adjacent *dwelth* – even if the information so cut had, in earlier iterations of the text, been deemed essential for the history and derivation that the OED should provide. Given the textual limits on space, the editor (Murray in this instance) had to choose and cut accordingly. Lexicographical fluency in the praxis and metalanguage of omission was conspicuous.

Particularly interesting, however, is the related evidence that the Proofs provide of a set of changes that were driven, in entirety, by the processes of the press and the calibration of allowable space that language history might occupy. What we might regard as the trivial matter of textual justification – and the desire to secure the requisite neatness of the footer or bottom margin – can, in this light, emerge as a further directive for obligatory change. Signalled by the press-room annotation "long" (or, conversely, "short") inscribed at the bottom of a printed column of text, press-room desiderata overtly took precedence while the dictionary-makers were compelled to adjust evidence and entries accordingly.

"Short", for example, routinely appeared in marking a column that, under the conditions of printing, fell short of its companions on the same page. It was, inevitably, far easier to remedy, given that it provided a rare opportunity for expansion. As on the Proof sheet of *narratively*, for example, Craigie's handwritten annotations ("In a narrative manner; ~~after the style of~~ also considered as a narrative") still illustrate his careful crafting of a definition that, under earlier pressures of space, had previously been omitted. Other examples can be located in e.g. *cockney* where Murray likewise seized the opportunity for amplification "to make up the line or 2 lines short", here by means of a new cross-reference as well as new embedded supporting evidence for *cock's egg*. He detailed a range of variables that the compositor might make use of, depending on the precise space required. Similar forms of delegated decision-making in relation to the published text are recorded elsewhere. "These are merely expansions & fill up the space" and so "may be dealt with as is most convenient to you", Murray likewise instructed under *commission*. Not all, as he realized, could be included. But the act

of selection was formally transferred from dictionary-maker to press-room in the final setting of the text.

Compensatory expansion of this kind was, however, the exception rather than the rule. “Long” was far more prevalent as a press-room annotation while definitions, evidence, etymology, usage notes and labelling could all be affected by the excisions that result. As under the Proof text of *oliva*, a word can be made to lose the senses it was previously thought to require or, as under *qualm* (v.) or *shiva*, to shed the chains of derivation they had previously possessed. In relation to definitions, too, a range of words were suddenly identified as targets for compression, gaining, by virtue of the injunction “long”, interpretative nuances that the relevant editor had not originally intended. Conditions of this kind attend, for example, the redefinition of *outwittal* as ‘the point of outwitting’ (from the earlier ‘outdoing or frustrating by trickery’) or dictated the appearance in print of ‘to preach above and beyond’ (under *overpreach* (v.) sense 1) instead of the earlier and more specific ‘To preach above or beyond (a person’s capacity)’. Craigie’s decision to compress ‘The study of the clouds as a branch of meteorology’ to ‘The study of the clouds’ (under *nephology*) was part of the same process, as was the truncated definition of *quietistic* so that if it applied to “quietists”, the earlier elaboration in terms of “quietism” was now removed. Other repercussions can be seen in the changed representation of supporting evidence, as dictated by an inscribed “long”, such that e.g. a vivid citation from Hawthorne’s *Scarlet Letter* (used in illustration of *frizzle*) was now reduced to its final five words: “~~Nothing was left of my respected predecessor, save an imperfect skeleton ...~~and a wig of majestic frizzle”. Related casualties included e.g. “~~His spontaneous emotion is the result of two over-night glasses of brandy~~” (under *overnight*); ‘~~A somewhat uncommon variety of the verruca simplex is the subungual wart~~’ (under *subungual*), ‘~~Their heads were full of the languid dreams of commentators~~’ (under *languid*), or ‘~~When He healed the sick, that was an earnest... of a sickless state~~’ (under *sickless*). The relevant headword was, in each case, still illustrated. But, at a number of points, loss arguably extends beyond the process of textual remediation that was pragmatically required.

Labelling, and evidence of in-text decision-making on whether field or status labels could be retained, reveal related forms of textual adaptation. Prefatory matter for the individual fascicles of OED1 often provided a neat tabulation and quasi-quantification of the labels that were deployed. Nevertheless, scrutiny of the Proofs confirms that deleting a label was early recognized as a relatively easy way to secure an additional line in order to meet compositorial requirements of this kind – even if this also problematizes our understanding of whether a word was (or was not) considered as obsolete or nonce, or attributive, or regional in

OED1. The entry for *siest* ('to take a siesta') was, for example, "nonce" until the injunction "long" appeared on the relevant page proof; Murray's accompanying marginal confirms that the resulting excision was "to enable the line to be saved" and makes the direction of willed change overt, alongside the priorities that textual revision had enforced. Similar is e.g. the annotation "omit if needed" which appears alongside a range of labels in the Proofs (and where "need" was delegated to the press-room, and decisions on what the final print text might require). Rare words can, by the same token, lose their rarity, obsolete words their obsolescence, and, as under *squatmore*, dialectal words their regional status.

Examples of this kind can be multiplied over the course of the Editors' Proofs. The textual imperatives of closure, on various levels, drove a range of changes as part of late-stage editing, and in the intent to secure the neatness of the individual print page. Nevertheless, in examining the intricate play of form and revision across the annotated text, it is interesting to contemplate whether a better historical record, or more representative data, can in fact be found in these antecedent Proofs and in, say, the definitions that existed before enforced compression took place, or in the citational range which can likewise suggest other more inclusive stories before space and time took their toll. In the published fascicles of OED1, the print text might thereby be finalized and closure achieved. But across the Proofs, the underlying dialectics of choice and change are made visible, presenting the possibility of new forms of exegesis as well as alternative histories that undercut the apparent certainties of the published text, and the depersonalized "the dictionary" that results.

## Archival material

Bodleian Library, Oxford. 30255 c. 17. Murray, James A. H., Henry Bradley, William Craigie & Charles Onions, *A New English Dictionary on Historical Principles* (annotated Editors' Proofs). 1882–1928.

OED/B/1/1/2. Herbert Coleridge. [c. 1858]. Draft report on the New English Dictionary. OED Archives at Oxford University Press.

MP/22/7/1912. James A. H. Murray to Walter Skeat. Murray Papers, Bodleian Library, Oxford.

## References

- Brewer, Charlotte. 2007. *Treasure-house of the language: The living OED*. London & New York: Yale University Press.
- Brewer, Charlotte. 2012. Happy copiousness? OED's recording of female authors of the eighteenth century. *Review of English Studies* 63. 86–117.
- Brewer, Charlotte. 2020. *Top Sources in OED1: Sir Walter Scott*. <https://oed.hertford.ox.ac.uk/quotations/outline/top-sources/#walter-scott>.
- Coleridge, Herbert. 1858. *Draft report on the New English Dictionary*. Oxford: OED Archives at Oxford University Press. (OED/B/1/1/2).
- Crook, Stephen. 1990. Auden on the record. *The W. H. Auden Society Newsletter* 6. <https://audensociety.org/06newsletter.html>.
- Ford, Thomas. 1854. *The compositor's handbook: Designed as a guide in the composing room*. London: Simpkin, Marshall, & Co.
- Gilliver, Peter. 2016. *The making of the Oxford English Dictionary*. Oxford: Oxford University Press.
- Graham, John. 1848. *The compositor's text book: Or, instructions in the elements of the art of printing*. Glasgow: Richard Griffin.
- Huxley, Thomas. 1871. *A manual of the anatomy of vertebrated animals*. London & New York: J. & A. Churchill.
- Lillis, Theresa. 2013. *The sociolinguistics of writing*. Edinburgh: Edinburgh University Press.
- Luna, Paul. 2004. Not just a pretty face: The contribution of typography to lexicography. In Geoffrey Williams & Sandra Vessier (eds.), *Proceedings of the 11th EURALEX international congress, 6-10 July 2004*, 847–858. Lorient, France: Université de Bretagne Sud. [https://euralex.org/wp-content/themes/euralex/proceedings/Euralex%202004/092\\_2004\\_V3\\_Paul%20LUNA\\_Not%20just%20a%20pretty%20face\\_the%20contribution%20of%20typography%20to%20lexicography.pdf](https://euralex.org/wp-content/themes/euralex/proceedings/Euralex%202004/092_2004_V3_Paul%20LUNA_Not%20just%20a%20pretty%20face_the%20contribution%20of%20typography%20to%20lexicography.pdf).
- Lynch, Jack. 2016. *You could look it up: The reference shelf from ancient Babylon to Wikipedia*. London: Bloomsbury.
- Mugglestone, Lynda. 2005. *Lost for words: The hidden history of the Oxford English Dictionary*. London & New York: Yale University Press.
- Mugglestone, Lynda. 2007. 'Decent Reticence': Coarseness, contraception, and the first edition of the OED. *Dictionaries: Journal of the Dictionary Society of North America* 28. 1–22.
- Murray, James A. H. 1880. The President's Annual Address for 1880. *Transactions of the Philological Society* 18. 117–176.

- Murray, James A. H. 1888. Preface to Volume I. In *A New English Dictionary on Historical Principles*, vol. I: A-B. Oxford: Clarendon Press.
- Murray, James A. H. 1900. *The Evolution of English Lexicography*. Oxford: Clarendon Press.
- Murray, James A. H., Henry Bradley, William Craigie & Charles Onions (eds.). 1884–1928. *A New English Dictionary on Historical Principles*, 10 vols. Oxford: Clarendon Press.
- Murray, K. M. Elizabeth. 1977. *Caught in the Web of Words. James A. H. Murray and the Oxford English Dictionary*. New Haven & London: Yale University Press.
- Rees, Emma & Ellie Stedall. 2019. Helping the OED to turn over a new fig leaf. *Times Higher Education*. <https://www.timeshighereducation.com/opinion/helping-oed-turn-over-new-fig-leaf> (21 November, 2022).
- Rundell, Michael & Adam Kilgarriff. 2011. Automating the creation of dictionaries: Where will it all end? In Fanny Meunier, Sylvie de Cock, Gaëtanelle Gilquin & Magali Paquot (eds.), *A taste for corpora: In honour of Sylviane Granger*, 257–282. Amsterdam: John Benjamins.
- Simpson, John & Edmund Weiner (eds.). 1989. *The Oxford English Dictionary*. 2nd edn., 20 vols. Oxford: Clarendon Press.
- Strzyżyńska, Weronika. 2021. Advanced copies of Sally Rooney's unpublished book sold for hundreds of dollars. *The Guardian* (20th August 2021). <https://www.theguardian.com/books/2021/aug/20/advanced-copies-sally-rooney-unpublished-book-sold-beautiful-world>.
- Sullivan, Hannah. 2013. *The work of revision*. Cambridge, MA: Harvard University Press.
- Trench, Richard Chenevix. 1857. *On some deficiencies in our English dictionaries*. London: John W. Parker & Sons.



## Chapter 3

# Dictionaries in the web of Alexandria: On the dangerous fragility of digital publication

Daphne Preston-Kendal

Humboldt-Universität zu Berlin

Digital publication is an attractive means of making dictionaries available to the general public, as has been widely documented by lexicographers and dictionary readers alike since the beginning of the “digital revolution” in dictionary publication over thirty years ago. However, commonly adopted means of publishing dictionaries pose dangers to scholarly integrity and to the long-term survival of lexicographical works. Adoption of practices used for the publication of some other online reference works and academic papers, indeed mirroring the survival characteristics of paper dictionaries, can help to mitigate these issues.

### 1 Models of digital reference publishing

The most common model adopted when publishing dictionaries online today is one which I have called the Oxford Model, after the *Oxford English Dictionary Online* which was probably the first to adopt it in full at its launch in 2000.<sup>1</sup>

---

<sup>1</sup>To be clear, I do not blame the OED editorial team nor Oxford University Press for adopting this model, though I am about to criticize it heavily. The digitization of the OED in the 1980s and development through the 1990s into the online dictionary which launched in 2000 was, in every way, a pioneering process. The OED website first went online during the comparatively early days of the World Wide Web, when it was not yet clear what the implications and dangers of this model would be.

This is also not to say I consider the Oxford University Press entirely responsible stewards of the OED in this regard. However, 23 years later, a shift to a more sustainable model should



In the Oxford Model, nobody can purchase their own complete copy of the dictionary to keep on their own computer. Instead, a “subscription” is offered for sale, allowing the subscriber to view individual dictionary entries over the Internet. These entries are loaded one at a time from the one and only complete copy of the dictionary, which exists only on the servers of the publisher.

Because there is only one complete copy of the dictionary held on the publisher’s servers, and because readers as subscribers can only load entries as they currently appear in that copy, readers have no guarantee that the entry they load today will have the same content as when they viewed it previously, or whether it might change (again) before they view it again in the future. The ability to publish the newest research immediately and perform ongoing revision of a dictionary is of course a great benefit of digital publication, but under the Oxford Model, readers have no way to go back and see exactly how an entry has been revised and when.

As subscribers to the dictionary website, rather than owners of a complete copy, readers are also beholden to the dictionary publisher on an ongoing basis. If they stop paying their subscription fee, they lose the ability to read the dictionary and cannot even see entries they have previously looked at, unless they had the foresight to save them to their local disk. On the other hand, if the publisher were to go out of business, decide that hosting the website was too unprofitable, or otherwise lose interest, they may simply take the website offline, at which point nobody can read the dictionary any more.

The Oxford Model in its complete form is the most irresponsible way to publish dictionaries online from the point of view of ensuring long-term, sustainable accessibility. The reasons for this, if they were not immediately clear from this description, will be explored in more detail below. Despite this, the Oxford Model is also very widespread: in the German national database of electronic academic resources, the *Datenbank-Infosystem*,<sup>2</sup> 1,373 resources described by “dictionary, encyclopedia, reference work” (*Wörterbuch, Enzyklopädie, Nachschlagewerk*) are accessed over the web in exchange for a licence fee as of writing.

There are a few variants of the Oxford Model in use by some dictionaries, each of which tempers the problems with it. Some online dictionaries are published under an Open Access Oxford Model, in which there is no subscription fee, but the publication in other respects follows the system outlined above. Others have

---

certainly have been made long ago. The Oxford University Press has declined requests from scholars for access to digital copies of older OED editions. Decision-making with regards to the OED in recent years has appeared to be increasingly profit-motivated, to the detriment of the long-term scholarly value of the work.

<sup>2</sup><https://dbis.ur.de/>

adopted an Oxford Model Without Revisions, where there is at least an informal policy that entries are not revised after they have been published (except perhaps to fix typographical errors and other, very minor details, of the same variety of change which may also be made between printings of the same edition of a paper dictionary). While both of these variants maintain some problematic aspects of the Oxford Model, they do at least mitigate some of the others. How and why will also be explored below.

## 2 The dangers of centralization

A key aspect of the paper model of publication is the distribution of complete copies for a one-time cost borne by the purchaser. While this may seem obvious, it is also notably different from the Oxford Model described above, to the extent that proposals to reinstate this crucial aspect within the context of digital publication can seem radical by comparison.

Why, then, is the distribution of dictionaries from a centralized authority only so problematic? Some problems of models in which the only complete copy of a text is held by the publisher are outlined above. The technological structure of the World Wide Web essentially encourages this model for the publication of all kinds of media online, but this situation is bad for the long-term accessibility of that media. New media researcher Bret Victor has informally called this situation the “Web of Alexandria”, in reference to the burning of the library of Alexandria, a disastrous loss of knowledge caused by the failure to make and distribute duplicate copies of works which were instead held in one central location (Victor 2015). While physical natural disasters or human accidents such as file deletion remain a danger to digital texts stored on the servers of a single institution, this model is insidious in other ways.

One has to do with how one of the advantages of digital publication – the ability for scholarly dictionaries to be continually revised to reflect the latest research – is applied in a centralized system. It has become quite normal for entries to be revised in their online editions. A description of how this occurs in practice under the Oxford Model is described in Simpson (2012): the OED1 and OED2 sub-entry for *market garden* was upgraded to a full entry in 2000, the first year of OED3 entry publication. At the time, the oldest quotation was from 1811, an improvement on the 1840 date given in the older edition. Two subsequent revisions pushed this back first to 1793, then to 1727. However, these revisions are “invisible” to the reader, and there is no way to trace when these antedatings were added. Indeed, there is no way to trace the specific dates of any changes

to OED3 articles after their initial publication, nor what they looked like at a particular previous point in time.<sup>3</sup> It is entirely conceivable that a study of, for example, word coinages of the early 19th century might have cited this OED3 entry at the point in time when it claimed the word was coined in 1811. Future readers of that study following this claim back to the source are left to wonder why its author missed or disregarded the earlier uses since added. This alone undermines basic scholarly practice.

Market forces and changing institutional interests can also pose a danger to the long-term accessibility of dictionaries under the centralized digital publication model. If a publisher decides a dictionary project is too unprofitable or loses interest, it may simply decide to remove the dictionary from the web entirely, so it becomes effectively entirely lost to history. We have already witnessed cases where publishers have made decisions like these: the tragic case of the *Historical Dictionary of American Slang* (henceforth HDAL)(Lighter 1994, 1997), cancelled by its publisher after only two of its planned four volumes, is well known. In a paper publication model, the completed dictionary parts are at least accessible by those who have already purchased copies; in an Oxford Model online environment, this kind of commercially-motivated decision would simply result in the total inaccessibility of the work already done. In fact, we have even seen a very similar dictionary become a victim of exactly this failure mode of the Oxford Model: *Partridge Slang Online*, an online edition of the *New Partridge Dictionary of Slang and Unconventional English* (Dalzell & Victor 2005), disappeared without trace from the web after about three years without any apparent explanation.<sup>4</sup> Fortunately in this case, the dictionary had also previously been published on paper, so the dictionary text remains accessible in libraries and on private bookshelves around the world, albeit without updates, advanced search tools beyond alphabetical headword lookup, and other benefits of digital publication.

Finally, while the ability to update dictionary entries on an ongoing basis is certainly a benefit of digital publication, the practice of doing this in a centralized

---

<sup>3</sup>Recently the OED3 website has added an indication of the month and year in which each entry was last revised, but there remains no indication of what was actually changed at each revision, nor any way to actually go back to an older version. This indication also often appears to be inaccurately forward-dated: as of writing, the entry for *market garden* claims the entry was last revised in March 2022: but there are no textual differences compared to the version of 2012 which appears in Simpson's paper.

<sup>4</sup>*Partridge Slang Online* was online at <http://www.partridgeslangonline.com/> from approximately January 2013 to September 2016 according to the Internet Archive WayBack Machine, which has archived the home page, front matter, and a small number of sample entries, but not the dictionary as a whole. After September 2016: only error pages are archived.

publication model makes the dictionary vulnerable to political attacks. Numerous dictionary projects, even those of a scholarly nature, have been the victims of political interference typically by totalitarian regimes,<sup>5</sup> but those regimes could not retroactively alter the contents of volumes of dictionaries already published (except by official edicts to libraries to censor their copies, which is reported to have happened in some cases). With centralized digital publication, this kind of change to entries for political reasons can be carried out immediately, affecting all users, and potentially without any trace of the older versions remaining.

None of this, however, need dissuade us from digital publication and its benefits entirely. Instead, an alternative model is required which retains these advantages over print while providing the long-term stability which was intrinsic to printed copies. As we will see, the low cost of digital publication can actually even lead to better and more widespread long-term accessibility than is possible with paper, provided the right technical and administrative choices are made about how the dictionary is actually published.

### 3 An alternative model

While digital dictionaries, especially scholarly dictionaries, are unique among digital texts in the extent and importance of their microstructure, lexicographers can still hope to learn from the experiences of other kinds of scholarly resources in exploring the possible models of digital publication.

A particularly instructive model, one which I hold to be the gold standard for digital publication of scholarly works, is that adopted by the *Stanford Encyclopedia of Philosophy* (SEP, Zalta 2006; see also Hammer & Zalta 1997, Allen et al. 2002). In the Stanford Model, an endowment has been established under a membership model which universities, libraries, and other institutions may join for a (significant, but one-time) fee. In exchange for this fee, the institutions obtain the right (and implicitly the responsibility) to archive the SEP on their own computers in perpetuity; if the SEP project ever closes down or becomes otherwise unavailable, they then gain the further right to re-host the entire encyclopedia on a public website to ensure general access to the text remains available. The encyclopedia is revised in versions which carry the dates of their publication explicitly, and all older versions of the encyclopedia and of the individual articles remain available in archived editions, both on the official website and in the

---

<sup>5</sup>As examples, see for descriptions of the ideological pressures exercised on lexicographers under the two dictatorial regimes in Germany in the 20<sup>th</sup> century, (Lea 2009) and (Zielinski 2010).

archived versions held by member institutions. Finally, because the membership scheme supports an endowment and only the returns on the investment are used to actually support the project, Stanford promises to return the membership fee with interest to members should the SEP project ever shut down. As of writing, over 500 institutions worldwide have become members of this scheme.<sup>6</sup>

This model in its entirety is incredibly ambitious. However, I will outline the most important aspects which can be applied to digital lexicography, dealing first with what might be termed the technical aspects, then an overview of the financial aspects.

The first technical aspect, as implied by the previous section, is that each member institution has under its control complete copies of the encyclopedia. They thus retain access to it, no matter what should happen to the producers and publishers. As applied to digital scholarly dictionaries, however, it is important not only that the text itself be made available, but that it be in machine-readable form; that is, with the microstructure encoded in a database file usually using descriptive XML or SGML markup. It is this machine-readable database structure which enables the advanced search capabilities which users of digital scholarly dictionaries have come to expect and rely upon. With machine-readable database files, this functionality can be made to work again long after the original software has become technologically obsolete.

The second is that this archival capability is, de facto, not strictly limited to member institutions; rather, because the SEP is available as open access in web page form without any authentication of users, the web pages themselves are picked up and archived by web crawlers such as the Internet Archive WayBack Machine and potentially by other private “webpage capture” services such as archive.is. These static web page archives can only be considered machine readable in a limited sense, and are not ideal for the use cases for machine-readable text previously outlined – but simply making an existing dictionary website open access represents an important and easy first step to allow archiving by third parties.

The third technical aspect is that not only do older versions happen to be accrued by member institutions over the period of their membership, continued distribution of older versions of the text is actually explicitly foreseen by the model. The ability to continue to distribute older editions and thus maintain the value of scholarly citations (as mentioned above) is an advantage of digital publication which has, in lexicography, remained thus far comparatively unexplored. It is another consequence of the fact that digital reproduction is close to free,

---

<sup>6</sup><https://plato.stanford.edu/support/commitments.html>

### *3 On the dangerous fragility of digital publication*

compared to the unattractive paper and ink costs associated with continually making older editions of a printed dictionary available.

These first two of these three technical aspects are not unique to the SEP but are in widespread use for the distribution of academic papers under the open access model.

I will now briefly discuss the financial aspects of the Stanford Model, though with an acknowledgement that the specific funding situations of individual lexicography projects differ significantly from one another.

Notably, the Stanford Model as applied for the SEP is concerned only with covering the costs of publication of the encyclopedia, not of paying the authors of its articles, which they write on a pro bono basis. Most lexicography projects involve a small integrated team, rather than a large distributed worldwide team of experts each writing articles in their own field of specialization – this difference necessarily implies that the SEP model, whereby authors consider their work on the dictionary to be part of their output of academic writing in their careers, is not workable for lexicography. While most European historical dictionary projects are funded as research projects under the auspices of a university or academy of sciences, and the costs of producing the text thus subsidized away for the eventual publisher who must only cover typesetting and printing costs, the OED and some other dictionaries of potential scholarly relevance (such as the “Unabridged” dictionary published by the Merriam-Webster company in the US) are commercial undertakings. This, indeed, was the case with HDAS, where it was likely not the printing but the editorial costs which motivated the early cancellation of the project.

It is difficult to know what to suggest in such cases, other than that the preparation of scholarly dictionaries in the Anglosphere, where currently carried out for commercial or quasi-commercial publishers, should move towards the European model. The association between the OED and the Oxford University Press goes back to the first edition, but perhaps it is time for this arrangement to be reconsidered, and for the editorial work of the OED to come under the aegis of and be funded by, for example, the University of Oxford’s Faculty of English, rather than its University Press. This is merely an idle suggestion, however, and made in the acknowledgement that European scholarly dictionary projects have historically moved much slower even than English ones, probably largely because of this model in which lexicographers are also employed as academics with their own teaching and research schedules, rather than as full-time dictionary writers.

## 4 Further work

It was suggested at the conference that some clear guidelines should be prepared for the sustainable publication of digital dictionaries. If lexicographers, lexicographical historians, and dictionary publishers feel such guidelines would be useful, an expert group could be assembled to use the experiences of lexicographers working on projects of various kinds, as well as those of librarians and archivists, to produce a thorough but realistic set of recommendations. Since scholarly dictionaries in particular are, in the digital age, increasingly being published “in house” by the department of the university or other institution which produces them, leaving decisions about publication mostly in the hands of lexicographers, such guidelines may be valuable especially for smaller teams which may lack specific expertise in this area. Several such smaller projects were presented at ICHLL12.

Nonetheless, I hope this paper’s recommendations – that wide distribution of machine-readable dictionary texts in their entirety and with their complete revision history should be made an integral part of the publication of lexicographical projects – offer a useful starting point until such guidelines are published.

The existing infrastructure for open access journal and data-set distribution provides a useful starting point for the technical means of achieving this. Another project of Stanford University, LOCKSS (Reich & Rosenthal 2000),<sup>7</sup> provides open-source software used by a number of “preservation networks” which automatically create copies of digital resources in linked member institutions around the globe. One network established by Stanford currently archives 52 million journal articles and nearly half a million ebooks; in addition, archives established by the governments of the United States and of Canada distribute and save those countries’ digital national archives, among other networks which preserve data for the future which cannot yet be made available to the public. LOCKSS is clearly a proven technology for this purpose, but given the unique requirements of digital dictionaries as structured texts, it would perhaps make most sense to establish a new network using this software which all historical lexicography projects could use to share their published work.

Until sustainable electronic dictionary publication as the norm becomes a reality, stopgap solutions must be found. One simple such stopgap is to continue to publish and buy dictionaries in paper form. This can be achieved through pressure on publishers and libraries to keep print alive even if digital publication is still perceived as the superior option. The paper publication model, as

---

<sup>7</sup><https://lockss.org/>

### 3 *On the dangerous fragility of digital publication*

outlined above, has intrinsically good survival characteristics: dictionaries such as *Partridge Slang Online* which have already died in digital form survive in paper form. “Retro-redigitalization” from a paper source of a dictionary which was originally made digitally, but whose digital form has been lost, may be a realistic last-resort possibility to recover a digital text.

In the era of online journal publication, it has become the norm for (sometimes extensive) supplementary material to be made available with the electronic version of an article. As a further stopgap measure, as long as Oxford Model reference resources exist, I propose that it should be made the norm that these supplementary materials should include PDF archival copies of the complete versions of any dictionary entries cited, when those entries were accessed from an Oxford Model publication. Given the low cost of electronic storage and distribution, this is technically plausible even for articles which cite very many dictionary entries, though in this case legal difficulties may arise due to copyright law. However, this practice alone should be sufficient to guard against the problem of being unable to trace back citations after an Oxford Model dictionary has been revised to obscure an original reference.

A further consideration related to the long-term retraceability of dictionary citations is the long-term reproduceability of research conducted using text search tools on online dictionaries. This implies that the exact search algorithms used need to somehow be reproducible. Even apparently simple algorithms can have very complex requirements in reality: a case-insensitive search is actually intrinsically language sensitive (as one example, ⟨i⟩ is not the lower case form of ⟨I⟩ in Turkish, unlike in most languages). This alone makes the task of searching an etymological dictionary, which typically mentions words from dozens of different languages in its text, much more complex. This complexity affects the reproducibility of research based on text searches, because the approaches used by a particular search software must be fairly exactly replicated in order to obtain the same results. The simplest approach would be to make all dictionary search software open source so that the same software can be used by everyone on their own computers. Code, though, intrinsically has worse survival characteristics than data: software platforms are constantly evolving, but data storage standards such as XML are essentially permanent. Making software which can still be used many years in the future is a broader problem in computing research, but the number of complicating factors in the design of linguistic applications makes the issue even trickier for our use cases.

## 5 Summary

My purpose in this article is certainly not to dampen the great enthusiasm for digital publication which lexicography has shown over the last few decades. On the contrary, I welcome digital publication and recognize the great power it offers – much of it as-yet still totally unexplored by lexicographers. I instead wish to heed caution and encourage more careful thought about the precise means by which digital dictionaries as texts are made available to the public.

Fixing this problem will require changes in the technical means of dictionary publication, as well as potentially in how scholarly dictionary writing is financed. Some of the technical problems involved have already been solved by other digital publications facing similar issues; others will require active research to resolve. The financial problems, where they exist, will likely have to be resolved by individual dictionary projects, since sources of funding already differ greatly between projects. Above all, solving this problem requires a recognition by dictionary sponsors that a scholarly dictionary is not a source of profit, but a contribution to human knowledge, and must be treated as such.

Indeed, solving this problem is a necessity if dictionary making as a scholarly pursuit is to continue, because matters of basic scholarly integrity are at stake. We must save our dictionaries from the Web of Alexandria.

## References

- Allen, Colin, Uri Nodelman & Edward N. Zalta. 2002. The Stanford Encyclopedia of Philosophy: A developed dynamic reference work. *Metaphilosophy* 33(1). 210–228. DOI: 10.1111/1467-9973.00225.
- Dalzell, Tom & Terry Victor (eds.). 2005. *The new partridge dictionary of slang and unconventional English*. London: Routledge.
- Hammer, Eric M. & Edward N. Zalta. 1997. A solution to the problem of updating encyclopedias. *Computers and the Humanities* 31(1). 47–60. DOI: 10.1023/A:1000418920193.
- Lea, Henry A. 2009. Dictionary-making in the Third Reich: The case of *Trübners Deutsches Wörterbuch*. *Seminar: A Journal of Germanic Studies* 45(4). 369–386. DOI: 10.3138/seminar.45.4.369.
- Lighter, Jonathan E. 1994. *Random House historical dictionary of American slang*, vol. 1: A–G. New York: Random House.
- Lighter, Jonathan E. 1997. *Random House historical dictionary of American slang*, vol. 2: H–O. New York: Random House.

### 3 On the dangerous fragility of digital publication

- Reich, Vicky & David S. H. Rosenthal. 2000. LOCKSS: Lots of copies keep stuff safe. *New Review of Academic Librarianship* 6. 155–161. DOI: 10.1080/13614530009516806.
- Simpson, John. 2012. What has the OED become? *Nordiske studier i Leksikografi* 11. <https://tidsskrift.dk/nsil/article/view/19131>.
- Victor, Bret. 2015. *The Web of Alexandria*. <http://worrydream.com/TheWebOfAlexandria> (14 May, 2023).
- Zalta, Edward N. 2006. The Stanford Encyclopedia of Philosophy: A university/library partnership in support of scholarly communication and open access. *College & Research Libraries News* 67(8). 502–504. DOI: 10.5860/crln.67.8.7670.
- Zielinski, Lech. 2010. Ideologie und Lexikographie: Die Ideologisierung des *Wörterbuchs der deutschen Gegenwartssprache* von Ruth Klappenbach und Wolfgang Steinitz. In *Danziger Beiträge zur Germanistik*, vol. 31. Frankfurt am Main: Peter Lang.



## Chapter 4

# A dictionary of the languages of medieval England: Issues and implications

Gloria Mambelli

University of Verona, Italy

The coexistence of three languages in England during the post-Conquest period led to a complex multilingual situation. The fact that Middle English, Anglo-French, and Latin showed mutual borrowing and shared lexis as a result of multilingual practices makes them difficult to handle from a lexicographic perspective. Starting from the assumption that “monolingual lexicography in a multilingual environment does not work” (Trotter 2000a: 4), this paper illustrates the issues related to the treatment of these languages in the three major historical dictionaries and advocates the compilation of a multilingual dictionary, which would be able to provide a more accurate perspective of the linguistic situation of the time. The problems and implications connected to the designing of such a long-needed resource are tackled and some methodological proposals are presented by taking into consideration two existing databases which treat specific lexical domains related to everyday life in medieval England from a multilingual perspective. The aim of this contribution is to urge further research on the possible ways to reproduce the fuzziness of language boundaries in the post-Conquest period in order to make the compilation of a multilingual dictionary of the languages of medieval England an attainable goal.

## 1 Introduction

The situation of linguistic contact which gave rise to “the trilingual civilization of medieval England” (Rothwell 1998: 164) has attracted considerable attention in recent scholarship (Trotter 2000b, Wogan-Browne et al. 2009, Ingham 2010,



Jefferson & Putter 2013, Wright 2020b). Textual evidence shows that, during the centuries following the Norman Conquest, any form of literacy required “considerable, probably native-like, competence in Anglo-French and Latin, as well as in English” (Durkin 2014: 229). The coexistence of these three major languages<sup>1</sup> is generally referred to as triglossia in view of the fact that language choice usually depended on domains and text types; however, increasing interest in “the non-literary material dealing (broadly speaking) with everyday life” (Trotter 2006: 73) has revealed a more complex situation in which the use of these languages overlapped. Specifically, the multilingual condition of medieval England is not only reflected in monolingual texts drawn up in each of the three languages, often found in multilingual collections intended for a multilingual readership, but also in multilingual texts in which the languages are mixed by means of code-switching practices (intended as the mixing of different codes in the same text).

Mixed-language writing has been extensively discussed by Wright in her studies of business documents (Wright 2000, 2002, 2011, 2013b,a, etc.), a “text type where a mixing of two or more languages is the norm” (Wright 2000: 149). Accounts and inventories regularly include instances of code-switching and are characterised by considerable language overlap, mainly by means of the abbreviation and suspension system, which, besides allowing scribes to save parchment and ink, allowed readers to interpret words according to their linguistic competence (see Wright 2000). These findings show that, contrary to the traditional views of early language mixing as a “broken” or “macaronic” language,

mixed-language texts are typically not the product of incompetent or non-native authors or scribes, but rather reflect the – often very high – multilingual competence of the authors and scribes who wrote or copied them. (Schendl & Wright 2011: 20)

Mixed-language practices inevitably add to the complexities underlying the linguistic condition of medieval England, which was typified by mutual borrowing and shared lexis. Investigations of medieval multilingual texts have highlighted the difficulty of assigning word-forms to a specific language: according to Hunt, “from a lexical point of view attempts at a clear distinction of language identity is futile, for such distinctions were certainly blurred in medieval England” (Hunt 2011, 64). Even though the study of code-switching and borrowing relies on the separation of codes, according to Trotter, we cannot “assume [...] that

---

<sup>1</sup>These were not the only languages in use in post-Conquest England: the Celtic languages survived in peripheral areas (Cornwall, Wales, Scotland, etc.), and “Norse continued to be spoken [...] quite possibly into the twelfth [century] in some places” (Townend 2006: 66).

what is perceived as separate now (after centuries of standardization and monolingualism) was regarded as separate or indeed separable in the Middle Ages” (Trotter 2011: 157). If, as claimed by Rothwell, “a cut-and-dried divide between languages did not exist” (Rothwell 1998: 545), then medieval speakers might not have been aware of language boundaries.

From a lexicographic point of view, the fuzziness of language boundaries poses problems in terms of language labels and leads to the following question: how can the medieval languages be effectively represented in lexicographic terms?

The following section addresses the issues and implications of a lexicographic treatment of the languages in use in post-Conquest England by presenting an analysis of existing historical dictionaries.

## 2 Medieval multilingualism and lexicography

At present, the lexicon of the languages of medieval England is recorded in three historical dictionaries: the *Middle English Dictionary* (MED, Lewis 1952–2001), the *Anglo-Norman Dictionary* (AND), and the *Dictionary of Medieval Latin from British Sources* (DMLBS, Ashdowne et al. 2018), all of which have been transferred into their online versions. The work of linguists heavily relies on these lexicographic resources, which, in the absence of actual speakers of the languages in question, represent essential tools for historical linguistic studies and investigations of medieval texts.

Unfortunately, historical dictionaries do not comprehensively record the extant attestations of the medieval languages, most of which are scattered in archives and not easily accessible. A considerable number of historical documents still remains unpublished and therefore its content has yet to find a place among dictionary citations. As a result, what we are offered by the aforementioned dictionaries is skewed towards attestations found in published texts.

The entries provided by the MED, AND, and DMLBS include Present-Day English definitions and related quotations. In addition, some of them provide information concerning spelling variants (MED, AND), etymology (MED, DMLBS), and links to other dictionaries (MED, AND). Users interested in one of the three languages are likely to be used to consulting the relevant dictionary and they certainly benefit from its content; however, it is only by consulting these three dictionaries concurrently that we are offered a glimpse of the overlapping use of the medieval languages and the complexities arising from it. For the reasons illustrated below, these dictionaries should necessarily be considered complementary resources.

As a result of the multilingual environment of post-Conquest England, each of these dictionaries inevitably includes words derived from the languages treated in the other two. The DMLBS, for instance, contains a considerable number of vernacular words (i.e. of French or English etymology) with Latin inflections, and, similarly, it is not unusual to find terms that are derived from English in the AND and French- or Latin-origin ones in the MED. This is due to the situation of language contact which characterised the Middle English period and, more specifically, to the existence of mixed-language attestations. As mentioned before, administrative documents include instances of code-switching, i.e. vocabulary from a language that is integrated into a different matrix language (generally vernacular words embedded in a Latin text). As a result, Middle English and Anglo-French words found in Latin texts are included in the DMLBS, and Latin and English words found in Anglo-French texts are included in the AND.

Table 1 illustrates a sample of lemmata of English native origin drawn from the *Oxford English Dictionary* (OED) which are found in two of or all the three dictionaries. These terms were selected as examples of vocabulary relating to concepts and activities connected to everyday life in medieval rural areas. They are all of native origin, either totally (e.g. *frith*, *hide*) or partially (e.g. *gavel*, from Old English and Medieval Latin, and *selfode*, probably combining an Old English element with an Old Norse or Middle English one), and are recorded in attestations displaying different matrix languages. It is intended as a sample of shared lexis, since the same word is recorded in different historical dictionaries, although in different forms. With respect to form, it is necessary to bear in mind that Middle English and Anglo-French spelling was not standardised, hence spelling variation is a prominent feature in surviving texts; the word-forms reported in Table 1 are the variants found as headwords.

The word *hayward*, an English term denoting a manorial or town officer in charge of keeping fences and enclosures under surveillance, is found in the three dictionaries. The AND entry is linked to the OED, MED, and DMLBS: this indicates that we are in the presence of a word of non-French etymology, which does not appear in any dictionary of Old or Middle French. The following citations are recorded in the AND:

- (1) Le **hayward** [...] deit tard e tempre espier, e environer, e garder lez boys, e lez blez, e lez prez (c.1270)
- (2) Un malvois ribaut baillif ou **hayward** par cohercion fra un povre om fere une suite (1310)
- (3) messatorc: **haiuard** (MS: s.xii/xiii)

Table 1: Sample of English-origin words recorded in more than one historical dictionary

OED	MED	AND	DMLBS
acreman	aker-man	-	acremannus
craft	craft	craft	-
croft	croft	croft	crofta
frith	frith	frith	frithum
gavel	gavel	gavel	gabulum
hayward	hei-ward	hayward	haiwardus
hide	hide	hide	hida
husbandry	hus-bondri(e)	husbandrie	husbonderia
mill	milne	mille	-
ridding	ridding(e)	-	riddinga
selfode	self-ode	-	selfodis
steward	steward	estiward	steuardus

The DMLBS records the Latinised form *haiwardus* and provides the Middle English form *heiward* at the beginning of the entry. The word is attested with different spelling forms, a typical feature of Middle English, in Latin documents dating between 1209 and 1410. Two examples are given below.

- (4) **haywardus** de B. adquietat annuatim de gabulo per servicium suum xx d., et debet habere in defenso domini, quamdiu animalia fuerint ad herbam, j bovem vel j vaccam (1234)
- (5) **heiwardus** habet penes se particulas et nomina plegiorum (1247)

The Middle English term *steuard* (Present-Day English *steward*), denoting an official in charge of the administration of a manorial estate, is recorded as *estiward* in Anglo-French and *steuardus* in British Medieval Latin. Both the AND and DMLBS signal its etymological origin, the former by referencing the corresponding entries in the OED and MED and the latter by reporting the Old and Middle English corresponding forms within square brackets before providing the definition. A different picture emerges when we look up the term *husbandry*, used to refer to the management of a household or estate and agricultural operations, in the three dictionaries. The AND explicitly invites the user to consult the MED

entry, whilst the DMLBS, in turn, references the Anglo-French term, suggesting that the word was borrowed into Latin through Anglo-French rather than directly from Middle English.

As shown in the table above, not all words are shared among the three languages. For instance, the English term *acremen*, denoting a class of tenants in medieval England, is recorded in the MED and DMLBS but not in the AND. This is because the word is attested in English- and Latin-matrix texts but not in documents with an Anglo-French matrix. The headword recorded in the DMLBS is the Latinised form *acremannus*, whose English origin is indicated by the presence of the Old English form *æcerman* at the beginning of the entry.

Mixed-language quotations are occasionally signalled in the MED by means of square brackets and presented as additional illustrative material, as explained by the editor:

Occasionally non-Middle English quotations (Latin, Old French, Old English, Middle Scots, etc.) are used for various illustrative purposes, e.g., for Latin or Old French, to exhibit the meaning or construction of the foreign word taken over into Middle English. (Lewis 2007: 19)

For instance, the MED entry for *croft* includes five Latin quotations below the sense ‘small piece of ground used for agricultural purposes (usually enclosed, often adjoining a house)’ with dates ranging from 1160 to 1400. All these are provided within brackets at the beginning of the block of quotations, separated from the English-matrix ones. However, this is not always the case.

If we consult the MED entry for *ridding(e)*, we find that square brackets are not used consistently and that different dictionaries show a different picture. *Ridding(e)* is a Middle English gerund form used to refer firstly to actions of clearing or cleaning of a site and, secondly, to an assart, that is an area of ‘forest or waste land converted to arable use by clearing trees, bushes, etc.’ OED. It is recorded as *riddinga* in the DMLBS, but only in the sense referring to an area of cleared land, which seems to suggest that the senses referring to clearing and cleaning are not attested in Medieval Latin. Nevertheless, the MED records several Latin attestations of the word, i.e. quotations in which the word is embedded in a Latin matrix, in the following senses: ‘clearing away of debris, scaffolding, etc.’, ‘clearing of a house, garden, pool, stables, etc.’, and ‘clearing of a site, esp. in preparation for construction’. None of these is marked by square brackets.

Interestingly enough, among these Latin quotations, the ones drawn from the *Account Rolls of the Abbey of Durham* are distinguished by the presence of a French grammatical function word preceding the English word, as illustrated by

the example below. This is a phenomenon which, according to Ingham (2009), represents evidence of the use of French as a medium of communication among estate professionals.

- (6) Joh'i Thomson pro *le riddyng*e et deposicione finali omnium murorum, lapidum, et lez butteressez veteris ecclesie Beate Marie Magd (1449)

The problems related to the inconsistency through which mixed-language attestations are signalled in historical dictionaries is discussed by Wright (2011, 2013b). In particular, she highlights errors and inconsistencies concerning first attestation dates, demonstrating how the DMLBS often provides antedatings of usage of English words.

If we look up the word *selfode*, an English term which denoted a specific kind of land holding and the class of tenants occupying it, we find that it is recorded in both the MED and DMLBS by means of Latin-matrix quotations only. In addition, the two earliest attestations provided by the OED are Latin quotations. Nevertheless, if we take only the MED and OED into account, we are led to believe that the word was not attested before the end of the thirteenth century in the sense relating to the type of tenant, since the earliest attestations appear to be the 1271 quotation (*Inquisition Post Mortem*) recorded in the OED and the 1281–1282 one (*Calendar of Documents Relating to Scotland Preserved in Her Majesty's Public Record Office*) in the MED. In fact, the DMLBS provides an earlier attestation, from 1209 (*Cartulary of Cirencester Abbey*).

Due to the status of Latin, the main language of record at the time, it is not unusual to find that certain English words were attested in Latin documents before their first appearance in English-matrix texts, hence consultation of the MED alone might lead to incorrect interpretation of attestations of usage. As illustrated in the example above, the first attestation of the English term *selfode* (1209) is found in the DMLBS but not in the MED or OED, which record later attestations.

The commonalities, differences, and inconsistencies discussed above show that taking into consideration only one of the languages of post-Conquest England is likely to be misleading since “monolingual lexicography in a multilingual environment does not work” (Trotter 2000b: 4). In addition, the presence of the same term, albeit in different forms, in the three historical dictionaries raises questions about the decisions made by lexicographers in terms of assigning words to languages.

Is *acremannus* a Latin word simply because it occurs in a Latin text with a Latin inflection notwithstanding its non-Latin etymology? If this is the case, how

should we treat nonce words or hapaxes such as *mille*, an English word attested only once in an Anglo-French document?

In general, “it is very difficult to assess the degree of assimilation of a borrowed word to a given language at a given point in time” (Wright 2000: 155), even with regard to contemporary phenomena. When it comes to lexical borrowing in the medieval period, the scarcity of surviving evidence and the presence of the same term in two or all the three historical dictionaries makes such assessment even more difficult to undertake. The inclusion of a foreign word in a dictionary would normally suggest its lexicalisation in the language in question; however, in the case of historical dictionaries, editorial policies differ from dictionary to dictionary, as well as from editor to editor, and are not always consistently applied, especially when it comes to mixed-language attestations, as previously illustrated. Evidently, “[t]he mixed-language text-type is difficult for monolingual dictionaries to handle” (Wright 2013a: 131). After all, lexicographers and scholars interested in the history of English have traditionally focused on monolingual literary texts, disregarding the fact that documents drawn up in everyday situations in medieval England display “varying proportions of Medieval Latin, Anglo-Norman and Middle English” (Wright 2020a: 5).

Since it is not an easy task to decide whether a certain word should be labelled as English, Latin, or Anglo-French when it comes to words attested in mixed-language texts, the overlapping use of the medieval languages represents a major issue from the point of view of lexicography.

### **3 A dictionary of the languages of medieval England**

Even though we cannot confidently assess the medieval perception of language boundaries, we can certainly observe that the modern conceptions which take the “separatedness of languages” for granted (Trotter 2009: 155) cannot be applied to the languages of medieval England, which manifested themselves as a continuum. The inconsistencies observed in historical dictionaries call into question the reliability of monolingual investigations in the light of mixed-language evidence, suggesting that “the different medieval languages cannot be studied in isolation” (Schendl 2015: 15).

Since we are necessarily dependent on historical dictionaries, research in the field of English historical linguistics, lexicology, and semantics would certainly benefit from a consistently designed resource assembling the lexicon(s) of the three languages of medieval England. A dictionary of the medieval languages would enhance users’ awareness of the multilingual environment of the period

and the “symbiotic relationship” by means of which English, Anglo-French, and Latin were inextricably interconnected (Turville-Petre 1996: 181).

Notwithstanding the validity of these assumptions, such work represents a major lexicographic challenge. From a methodological and practical perspective, the following questions arise:

1. Which lexicographic approach would be the most appropriate to tackle the multilingual environment of post-Conquest England?
2. Would it be possible for lexicographers to reproduce the medieval fuzziness of language boundaries?

So far, only two resources treating the lexis of medieval England from a multilingual perspective have been compiled and made available. These are the *Lexis of Cloth and Clothing Project [LexP] n.d.* and the *Bilingual Thesaurus of Everyday Life in Medieval England ([BTh] n.d.)*, two databases focusing on specific lexical domains.

The LexP assembles the vocabulary in use in Britain between 700 and 1450 to refer to dress and textiles. In addition, it gathers non-textual evidence provided by surviving textiles, accessories, and art from the medieval period. The lexical data are drawn from historical dictionaries and displayed in alphabetical order, implying a semasiological approach aimed at investigating meaning. At the same time, however, the presence of images prompts an onomasiological approach in that users are urged to investigate the ways in which the visual item was described in language. Indeed, the compilers of the LexP aimed at creating “a tool which weaves together dictionary, encyclopedia and bibliography” (Rutten 2008: 772). Moreover, in addition to definitions, citations, and etymological information, each entry offers the different forms in which the word is attested in the medieval languages. In fact, the project does not focus solely on the three languages of literacy of post-Conquest England but it records vocabulary in Old and Middle English, Anglo-French, Medieval Latin, Anglo-Norse, and Celtic languages such as Welsh, Old Irish, etc. As a result, we can indeed refer to it as a multilingual resource. For instance, the terms *baldekin* (Anglo-French), *baldekinus* (Latin), *baudekin* (Middle English), *boytkin* (Cornish), *baudkin* (Older Scots), denoting a specific kind of textile used for drapes and canopies, are all recorded in the entry below the Present-Day English headword *baldachin*. Users can either browse the list of headwords or search for a specific form (e.g. *baldekin*) in the database.

The [BTh] (n.d.) organises the Middle English and Anglo-French lexis relating to medieval professions and everyday life activities in conceptual groupings, i.e.

by means of an onomasiological approach. More specifically, it focuses on the vocabulary of seven occupational domains of late medieval England (Building, Domestic Activities, Farming, Food Preparation, Manufacture, Trade, Travel by Water) in use between 1150 and 1450. The lexical data come from the MED and AND and the conceptual classification devised for the project is based on the category structures of the *Historical Thesaurus of the Oxford English Dictionary* ([HTOED] *n.d.*), occasionally modified because the editors aimed to assemble associative fields rather than build a typically semantic hierarchy.

This resource represents one of the first attempts “to capture the characteristics of two different languages in overlapping use during the same period” (Sylvester et al. 2017: 318). Firstly, the macrostructure of the [BTh] (*n.d.*) allows us to gain an overview of the terms that were used as (near) synonyms<sup>2</sup> or cohyponyms in post-Conquest England: Middle English and Anglo-French words are grouped in alphabetical order below a Present-Day English heading expressing the concept they refer to. In this way, referents become the major focus of investigation, making language boundaries less prominent. However, the two languages are treated as distinct entities in that lexical items are presented in two separate sections, below the headings “Anglo Norman”<sup>3</sup> and “Middle English”. This inevitably conveys a modern-day perspective which might result in users not being aware of the overlapping condition of the medieval languages.

Secondly, attention to the issues related to mixed-language attestations and shared lexis clearly emerges in the dictionary microstructure. The matrix language of the documents in which the word in question is recorded is signalled by means of labels (“French”, “Latin”, etc.) and referred to as “language of citation”. This appears to be an effective solution to make users aware of medieval language-mixing practices. Moreover, by means of labels (“Old English”, “Old French”, “Old Norse”, etc.) derived from the etymological information found in the MED, the Middle English words are provided with information about their language of origin. Unfortunately, the same does not apply to the Anglo-French lexical items, for which only the language of citation is signalled. For instance, the term *croft* appears in both sections of the category “small holding or croft”

---

<sup>2</sup>When the term *synonym* is used, a caveat is in order. According to Samuels (1972), “language possesses no pure synonyms”. Synonymy produced by the introduction of borrowings into the language is expected to exist only temporarily: after a period of time in which variants coexist, “either one of them will become less and less selected and eventually discarded, or a difference of meaning, connotation, nuance or register will arise to distinguish them” (1972, 65).

<sup>3</sup>The language is referred to as “Anglo Norman” in the [BTh] (*n.d.*), although some scholars favour the use of the more generic term “Anglo-French”, which better reflects the features of the language (see Durkin 2014: 230).

because it is attested in both Middle English and Anglo-French citations (and is thus recorded in both the MED and AND), as shown in the example below. Even though the link can easily be inferred because of the presence of the term in both sections, its Old English origin is indicated in the Middle English section only.

SMALL HOLDING OR CROFT (N.)

ANGLO NORMAN:

**Croft** 1200–1421

Languages of citation: French, ?Latin

**Fin** 1200

Languages of citation: French, ?Latin

MIDDLE ENGLISH:

**Croft** 1160–1450

Language of origin: Old English

Languages of citation: Latin, English

**Curtilage** 1439–1450

Language of origin: Old French

Language of citation: English

**Graunge** 1440–1447

Language of origin: Old French

Language of citation: English

Nevertheless, the onomasiological arrangement of the medieval lexis devised by the [BTh] (n.d.) editors is effectively able to shift the focus away from language boundaries, since referents rather than word-forms are used as the starting point of investigation. This helps us to address the question relating to the most appropriate lexicographical approach to be adopted in a dictionary of the languages of medieval England: in a note to an essay discussing language contact in the medieval period, Trotter observes that “[t]here is a strong case for an onomasiological treatment of later medieval England, which would alone be able really to show how the three languages (Anglo-French, Middle English and Latin) functioned” (Trotter 1996: 35).

Moreover, by means of a new conceptual classification which takes into account the cognitive experience of the Middle Ages, the [BTh] (n.d.) brings users

closer to the environment of medieval England from both a linguistic and cognitive perspective, without imposing “a modern day world view onto the medieval data” (Sylvester et al. 2017: 310). Consequently, it could be considered a model for the creation of a dictionary, or rather, a thesaurus of the languages of medieval England, which would offer users the possibility to browse medieval concepts to find out how they were referred to in the three languages during the late medieval period.

To this end, it would be advisable to avoid internal subdivisions which might imply a separation of the medieval languages, in order to obtain a multilingual lexicographic resource challenging the monolingual lexicographic tradition from which the historical dictionaries now available originated. In addition, on the model of the [HTOED] n.d., (near) synonyms and/or co-hyponyms in the three languages should be arranged chronologically: in this way, users would be able to trace the diachronic development of the semantic space and see when lexical items were adopted and when they fell out of use.

Nevertheless, the question of language labelling still remains unsolved. Nowadays, we are used to a concept of language boundaries that differs from the medieval one and to a lexicographic tradition that handles languages as separate entities. As a result, in order to meet modern users’ needs, each term in a potential thesaurus of medieval vocabulary would still need to be assigned to one of the three languages. As previously explained, data from mixed-language texts inevitably pose problems in terms of language labels, making the designing of the tool proposed here a challenging, controversial task. So far, the solution adopted by the [BTh] (n.d.) editors, who opted for signalling the matrix language(s) of the citations in which every lexical item appears, seems the most appropriate and effective for addressing the problem of language mixing in medieval England.

Even though the creation of such a long-needed work seems to still lie in the future, existing resources such as the LexP and [BTh] (n.d.) demonstrate that, by abandoning the monolingual approach of traditional historical lexicography, it is possible to convey a perspective that is closer to the one of medieval speakers, getting away from the modern one which stems from “the ideology of modern nationalism and the comforting doctrines of monolingual simplicity” (Trotter 2000a: 3). Obviously, were such an ambitious project to be initiated, issues of a different nature would arise. In particular, the problem of the long-term survival of online lexicographic works described by Preston-Kendal in Chapter 3 will need to be addressed in order to ensure the preservation and accessibility of the achieved results. However, this is beyond the scope of this paper.

## 4 Conclusions

The inconsistencies observed in the treatment of shared lexis and mixed-language attestations in the existing historical dictionaries, i.e. the MED, AND, and DMLBS, have shown that “a perspective drawn from one language alone [...] is manifestly inadequate” (Trotter 1996: 31). Consequently, this paper advocates the compilation of a multilingual dictionary of the languages of medieval England, which would bridge the gap between existing monolingual dictionaries by providing a more accurate perspective of the linguistic situation of the time. However, the fuzziness of boundaries typifying the medieval languages and their being “in continuous, concurrent use throughout the Middle Ages and in constant contact” (Hunt 2011: 68) makes them extremely difficult to reproduce in lexicographic terms. Further research is needed to answer the question of how to handle shared lexis in a multilingual context and reconcile the needs of modern dictionary users with the actual features of the medieval languages.

Even though these problems are still unresolved and represent a major obstacle to the realisation of a multilingual lexicographic resource, the advantages offered by an onomasiological approach in the handling of multilingual data should be taken into account. The fact that language boundaries are perceived as less prominent when lexical investigation starts from the content side rather than from the formal one will possibly prove beneficial in the making of such a long-needed work, at least in terms of overcoming the modern urge to treat languages as distinct entities in favour of conceptions of language boundaries in line with the medieval ones. As explained in this paper, in order to achieve better results in this sense, the multilingual lexical material should be presented in chronological order in each conceptual group without being separated according to the language.

In sum, the existing historical dictionaries treating the languages of medieval England, on which historical linguistic studies necessarily rely, present issues and inconsistencies that result in an imprecise overview of the linguistic environment of medieval England. Hence, a new, multilingual lexicographic resource is an important desideratum for historical linguists, notwithstanding the issues and implications involved in the initiation of such an ambitious project. In light of the recent findings concerning the status of the languages of medieval England, we necessarily need to abandon the traditional monolingual approach, which “is neither appropriate nor adequate for the investigation of language use in a society where multilingualism was endemic and where, for the educated at least, monolingualism was the exception and not the norm” (Trotter 2000a: 3).

## References

- [BTh]. N.d. *Bilingual Thesaurus of Everyday Life in Medieval England*. <https://thesaurus.ac.uk/bth/>.
- [HTOED]. N.d. *Historical Thesaurus of the OED*. <https://www.oed.com/thesaurus/start>.
- [LexP]. N.d. *The Lexis of Cloth and Clothing Project*. <http://lexisproject.arts.manchester.ac.uk/research/index.html>.
- Ashdowne, Richard, David Howlett & Ronald Latham (eds.). 2018. *Dictionary of medieval Latin from British sources*, 3 vols. Oxford: Oxford University Press.
- Durkin, Philip. 2014. *Borrowed words: A history of loanwords in English*. Oxford: Oxford University Press.
- Hunt, Tony. 2011. The languages of medieval England. In Michael Baldzuhn & Christine Putzo (eds.), *Mehrsprachigkeit im Mittelalter*, 59–68. Berlin: De Gruyter. DOI: 10.1515/9783110254815.59. (15 May, 2023).
- Ingham, Richard. 2009. Mixing languages on the manor. *Medium Ævum* 78(1). 80–97. DOI: 10.2307/43632800.
- Ingham, Richard. 2010. *The Anglo-Norman language and its contexts*. Woodbridge: Boydell.
- Jefferson, Judith & Ad Putter. 2013. *Multilingualism in medieval Britain (c. 1066–1520)*. Turnhout: Brepols.
- Lewis, Robert E. 2007. *Middle English dictionary: Plan and bibliography*. Ann Arbor: University of Michigan Press.
- Lewis, Robert E. (ed.). 1952–2001. *Middle English Dictionary*. <https://quod.lib.umich.edu/m/middle-english-dictionary/dictionary>.
- Rothwell, William. 1998. Arrivals and departures: The adoption of French terminology into Middle English. *English Studies* 79(2). 144–165.
- Rutten, Stuart. 2008. On the lexis of cloth and clothing project. In Elisenda Bernal & Janet DeCesaris (eds.), *Proceedings of the XIII EURALEX international congress (barcelona, 15-19 July 2008)*, 771–778. Barcelona: IULA, Documenta Universitaria.
- Samuels, Michael Louis. 1972. *Linguistic evolution: With special reference to English*. Cambridge: Cambridge University Press.
- Schendl, Herbert. 2015. Language contact and code-switching in multilingual late medieval England. In Juan Camilo Conde-Silvestre & Javier Calle-Martín (eds.), *Approaches to Middle English: Variation, contact and change*, 15–34. Frankfurt am Main: Peter Lang.

- Schendl, Herbert & Laura Wright. 2011. Code-switching in early English: Historical background and methodological and theoretical issues. In *Code-switching in early English*, 15–45. Berlin: de Gruyter.
- Sylvester, Louise, Imogen Marcus & Richard Ingham. 2017. A bilingual thesaurus of everyday life in medieval England: Some issues at the interface of semantics and lexicography. *International Journal of Lexicography* 30(3). 309–321. (15 May, 2023).
- Townend, Matthew. 2006. Contacts and conflicts: Latin, Norse, and French. In Lynda Mugglestone (ed.), *The Oxford history of English*, 61–85. Oxford: Oxford University Press.
- Trotter, David. 1996. Language contact and lexicography: The case of Anglo-Norman. In Hans Frede Nielsen & Lene Schøsler (eds.), *The origins and development of emigrant languages. Proceedings from the second rasmus rask colloquium, odense university, november 1994*, 21–39. Odense: Odense University Press.
- Trotter, David. 2000a. Introduction. In *Multilingualism in later medieval Britain*, 1–5. Cambridge: D. S. Brewer.
- Trotter, David. 2000b. *Multilingualism in later medieval Britain*. Cambridge: D. S. Brewer.
- Trotter, David. 2006. Language contact, multilingualism, and the evidence problem. In Ursula Schaefer (ed.), *The beginnings of standardization: Language and culture in fourteenth-century England*, 73–90. Frankfurt am Main: Peter Lang.
- Trotter, David. 2009. “Stuffed Latin”: Vernacular evidence in Latin documents. In Jocelyn Wogan-Browne (ed.), *Language and culture in medieval Britain: The French of England, c.1100–c.1500*, 153–163. Woodbridge: Boydell & Brewer.
- Trotter, David. 2011. Death, taxes and property: Some code-switching evidence from Dover, Southampton, and York. In Herbert Schendl & Laura Wright (eds.), *Code-switching in early English*, 155–89. Berlin: De Gruyter Mouton.
- Turville-Petre, Thorlac. 1996. *England the nation: Language, literature, and national identity, 1290–1340*. Oxford: Clarendon Press.
- Wogan-Browne, Jocelyn, Carolyn Collette, Marianne Kowaleski, Linne Mooney, Ad Putter & David Trotter. 2009. *Language and culture in medieval Britain: The French of England, c.1100–c.1500*. York: York Medieval Press.
- Wright, Laura. 2000. Bills, accounts, inventories: Everyday trilingual activities in the business world of later medieval England. In David Trotter (ed.), *Multilingualism in later medieval Britain*, 149–156. Cambridge: D. S. Brewer.
- Wright, Laura. 2002. Code-intermediate phenomena in medieval mixed-language business texts. *Language Sciences* 24(3). 471–489. DOI: 10.1016/S0388-0001(01)00045-6.

- Wright, Laura. 2011. On variation in medieval mixed-language business writing. In Herbert Schendl & Laura Wright (eds.), *Code-switching in early English*, 191–218. Berlin: de Gruyter.
- Wright, Laura. 2013a. Mixed-language accounts as sources for linguistic analysis. In Judith Jefferson & Ad Putter (eds.), *Multilingualism in medieval Britain (c. 1066–1520)*, 123–136. Turnhout: Brepols.
- Wright, Laura. 2013b. On historical dictionaries and language boundaries. In Liliana Sikorska & Marcin Krygier (eds.), *Evur happie & glorious, ffor I hafe at will grete riches*, 11–26. Frankfurt am Main: Peter Lang.
- Wright, Laura. 2020a. Introduction. In Laura Wright (ed.), *The multilingual origins of Standard English*, 3–15. Berlin: De Gruyter Mouton.
- Wright, Laura. 2020b. *The multilingual origins of Standard English*. Berlin: De Gruyter.

# Chapter 5

## The treatment of English high-frequency verbs in the *Promptorium Parvulorum* (1440)

Kusujiro Miyoshi

Soka Women's College in Tokyo, Japan

This dictionary has long attracted interest as the first English-Latin dictionary, but, as far as I am aware, its treatment of English high-frequency verbs has been little investigated. My intention is to fill this void. My analytic procedure will be to adopt a sampling method, selecting 128 entries related to 11 verbs, namely, *cast*, *come*, *fall*, *get*, *go*, *make*, *put*, *set*, *take*, *throw* and *turn*. I will then analyze how this dictionary treats individual verbs, phrasal verbs and other relevant phrases. If a dictionary published more than three hundred years before Johnson's *Dictionary* does in fact treat English high-frequency verbs in detail, one generally accepted notion about the history of English lexicography may have to be revised.

### 1 Introduction

I want to begin my paper by posing this question: "What is the first English dictionary that treated quite a few high-frequency verbs, distinguishing their senses and indicating their conjugated forms, as well as listing their relevant idiomatic phrases, especially phrasal verbs?". To my understanding, it has generally been believed to be Samuel Johnson's *Dictionary of the English Language* of 1755. If so, is this answer in accord with the facts? Anne McDermott and Rosamund Moon once remarked that [Johnson] was the first to deal systematically with such fundamental features of the English language as phrasal verbs, common polysemous verbs such as *take* [...] participial adjectives, verbal nouns, and many others. (McDermott & Moon 2005: 153) If we disregard one word "systematically" in this



Kusujiro Miyoshi. 2025. The treatment of English high-frequency verbs in the *Promptorium Parvulorum* (1440). In Geoffrey Williams, Mathilde Le Meur & Andrés Echavarría Peláez (eds.), *West meets East: Papers in historical lexicography and lexicology from across the globe*, 59–76. Berlin: Language Science Press. DOI: 10.5281/zenodo.15394477



account, a concept they do not expand on in their paper, this statement is quite off the mark.

In fact, there are two ways to define an English dictionary. One is to regard it as a monolingual dictionary, which may justify seeing Johnson's Dictionary as the first that detailed the senses and usage of English high-frequency verbs.

However, there is another way. For one thing, we know Henry Wheatley's pioneering survey of the history of English lexicography "Chronological notices of the dictionaries of the English language" (Wheatley (1865)). discusses both English-Latin dictionaries and English monolingual dictionaries. And we also know Gabriele Stein's monumental philological study on English lexicography entitled *The English Dictionary Before Cawdrey* (Stein 1985). Stein mainly treats historic bilingual and polyglot dictionaries published in England before 1604.

These examples show that, for Wheatley and Stein, the category of English dictionaries includes not only monolingual ones but also those which show how English words and phrases can be translated into and explained in other languages. Such a situation also applies to the *Oxford History of English Lexicography* (Cowie 2009) edited by A. Cowie and the first and second volumes of the *Ashgate Critical Essays on Early English Lexicographers* (Franzen 2012) edited by Christine Franzen.

Taking the second view, which is more inclusive, makes it far from correct to regard Johnson's Dictionary as the answer to the initial question. According to my research of the available dictionaries today, it is the *Promptorium Parvulorum* (1440) that includes all the features mentioned above. This first English-Latin dictionary was issued at Norfolk in the Middle English period by an anonymous author known as Galfredus Grammaticus, a monk and grammarian, according to R. C. Alston Alston (1968). The contents of the dictionary are available in its modern reproduced version, edited by A. Mayhew and published by the Early English Text Society in 1908.

The *Promptorium*, compiled more than three hundred years before Johnson's Dictionary, has long attracted scholars as it is the first English-Latin dictionary, being discussed quite often in their various research papers and books, including Wheatley (1865), Starnes & Noyes (1946), Wells (1973), Hüllen (2006), Sauer (2009) and Bately (2009), with the culmination of Stein (1985) and Stein (1997). However, it seems that research specializing in English high-frequency verbs has yet to be done. I hope to contribute to meeting this need.

My sampling method selects 128 entries representing 11 verbs. Specifically, the *Promptorium* provides the number of entries noted in parentheses for each verb in my selection like this: *cast* (10), *come* (5), *fall* (6), *get* (4), *go* (15), *make* (47), *put* (14), *set* (10), *take* (6), *throw* (6) and *turn* (5).

## 5 English high-frequency verbs in the *Promptorium Parvulorum*

There are two further notes to be added. One concerns the spellings in the *Promptorium* that often seem to be archaic and dialectal to our modern eye. The author of the dictionary thus often renders the infinitives of verbs according to contemporaneous morphology, such that the modern *get*, *put*, *set* and *take*, for instance, appear as *getyn*, *puttyn*, *settyn* and *takyn*, applying the infinitive endings of his region in the period. In this regard, we may also have to remember that the use of *to* was yet to be fully consolidated for indicating an infinitive at his time, except for the case where the notion of a purpose is expressed.

In relation to such features of Middle English, Gloria Mambelli's paper "A dictionary of the languages of medieval England", included in this book, is highly useful in determining how the language has been and will be dealt with in relevant dictionaries.

The other note concerns the lemmatization of relevant entries in the dictionary, to which I will attach considerable importance. Regarding this point, Stein, based on her investigation made from an overall viewpoint, remarks that:

It looks as if the author [of the *Promptorium*] explains the English words before he gives the Latin translation. [...] This unusual lemmatization derives obviously from the practice of changing Latin-English entries into English-Latin ones, The English lemma is thus basically still an explanation of a former Latin-English dictionary. (Stein 1985: 98–99)

Assuming that this is also applicable to the treatment of English high-frequency verbs in the dictionary, including distinguishing their senses and providing entries for their conjugated forms and their relevant idiomatic phrases, this will still not affect the fact that the *Promptorium* is the earliest existent dictionary that lists relevant words and phrases, which I suggest should be given its due weight.

Given our discussion so far, I will first survey the treatment of the infinitival and conjugated forms of English high-frequency verbs in the *Promptorium* (Section 2). This will constitute my preliminary analysis. I will then detail how the dictionary treats individual verbs (Section 3), phrasal verbs (Section 4) and other types of phrases (Section 5). This allocation of sections is based on three points which I regard as characterizing English high-frequency verbs: (1) having irregular conjugations; (2) having multiple senses; and (3) forming quite a few idiomatic phrases.

## 2 *The Promptorium's* treatment of the infinitival and conjugated forms of English high-frequency verbs

In analyzing how the *Promptorium* handles the infinitival and conjugated forms of these verbs, and their related phrases, I will consider three viewpoints: (1) the treatment of their infinitival and gerundive forms; (2) that of their past-participial forms; and (3) the provision of an overview. I will also indicate the basic structure of entries in the dictionary related to the high-frequency verbs.

### 2.1 The treatment of infinitival and gerundive forms

The reason why I discuss the treatment of the infinitival and gerundive forms together here is that, within the range of my study, there are four cases where the *Promptorium* provides two entries for the same verb, one for the infinitival form and one for the gerundive form, to which I also want to refer, recognizing that this fact will benefit my later analysis.

Two cases out of the four are as follows:

**Puttyn**, or leynn: *Pono, -is, -sui, -re; 3com., act.: Colloco, -as, -aui, -re; prime con., act.*

**Puttynge**, or leyynge: *Possicio, -nis; Collacio, -nis; omnia fem., 3 Ded.*

---

**Throwyn**, or castyn: *Iacto, -as; prime con., act; Projicio, -is, eci, -re; 3 con., act.*

**Throwyng**, or castyng: *Iactura, -e; fem., prime; Iactas, -tus; masc., 4.*

Here, Latin words corresponding to the entry-words ending with “-yn”, shows that they are indicative present forms, meaning that the verbs *puttyn* and *throwyn* are treated as infinitives. The head-words with “-yng”, they are gerunds, not present participles, as their corresponding Latin words *possicio* and *iactura* are, respectively, nominal forms.

As to other two cases, they are the followings, which show the same situation as above:

**Settyn** and plantyn: *Planto, -as, aui, -re; prime con., neut.*

**Settyng**, or plantyng: *Plantacio, -nis; fem. 3.*

---

**Takyn**, or resevyn: *Accipio, -is, cepi, -re, -ceptum; Capio, -is, cepi, -re, captum [...].*

**Takyngge**, or resevynge: *Accepicio, -nis; Suscepicio, -nis; omnia fem., 3: Caplura, -re; fem., prime.*

Of the 128 entries in my analysis, 42 concern words with “-yn” and 19 concern the words with “-ynge”. However, this does not mean that they are limited to 42 with “-yn”. When relevant Latin words are browsed through, it we find that the *Promptorium* provides 102 entries for infinitival forms of verbs and verbal phrases, with the use of spellings like *cum* for *come*, *goon/gon* for *go* and *make*, as well as *makyn* for *make*. The entries related number only 19. The *Promptorium* seems not to treat present-participial forms, which may be confused with gerundive forms, at least within the limits of my scope study, including when the Latin words listed are considered.

## 2.2 The treatment of past-participial forms

For the treatment of English past-participial forms, the *Promptorium* usually provides the perfect passive participle and adjectival forms of Latin words, as the following three entries show:

**Cast**, or castyde: *Iactus, -ta; Proiectus, -ta, -tum.*

---

**Sette**, or putte: *Positus, -ta, -um; Collocatus, -a, -um.*

**Put togedyr**, or onyd: *Contiguus, -a, -um.*

The author of the *Promptorium*, thus provides 7 entries on English past-participial words and phrases within my data. The Promptorian does not deal with preterit forms, which are equivalent to past-participial forms in our modern English, as in the cases of *cast*, *put* and *set*.

## 2.3 The provision of an overview

How often does the *Promptorium* provide entries on infinitival and conjugated forms of verbs and verbal phrases for each of the 11 verbs I cover? Table 1 illustrates this ;hereafter, the abbreviations “inf.”, “ger.” and “pp.” stand for, “infinitival”, “gerundive” and “past-participial”, respectively.

The distribution of the three forms being quite uneven among entries on relevant verbs, it is hardly possible from this table alone to tell what principle(s) the

Table 1: The number of entries for infinitival and conjugated forms of verbs and verbal phrases in the *Promptorium*

	inf.		ger.		pp.		total
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
<i>cast</i>	7	70.0	1	10.0	2	20.0	10
<i>come</i>	5	100.0	0	0.0	0	0.0	5
<i>fall</i>	3	50.0	2	33.3	1	17.7	6
<i>get</i>	2	50.0	2	50.0	0	0.0	4
<i>go</i>	15	100.0	0	0.0	0	0.0	15
<i>make</i>	47	100.0	0	0.0	0	0.0	47
<i>put</i>	9	64.3	2	14.3	3	21.4	14
<i>set</i>	6	60.0	3	30.0	1	10.0	10
<i>take</i>	5	83.3	1	16.7	0	0.0	6
<i>throw</i>	3	50.0	3	50.0	0	0.0	6
<i>turn</i>	0	0.0	5	100.0	0	0.0	5
total	102	79.7	19	14.8	7	5.5	128

author of the dictionary might have employed in providing such entries. However, it may at least be said that the table presents an overview of the historically most primitive treatment of English high-frequency verbs in lexicography.

### 3 The *Promptorium*'s treatment of individual verbs

Having surveyed the treatment of the high-frequency verbs in the *Promptorium* in the previous section, this section, aims to clarify how the dictionary deals with the relevant verbs individually from two points of view: (1) the provision of entries for their infinitival and conjugated forms; and (2) the discrimination and indication of their senses in the English language.

For this purpose, I will collate Tables 2 and 3 which my investigation has produced; the former concerns entries whose head-words include the senses of the verbs and the latter entries whose head-words do not include them, in both of which I arrange the head-words in accordance to their infinitival, gerundive and past-participial forms, indicating, for reference, Latin word/words the author provides for each.

5 English high-frequency verbs in the *Promptorium Parvulorum*

Table 2: Entries on verbs in the *Promptorium* where their senses are “specified” in the head-words

“inf.” (Latin)	“ger.” (Latin)	“pp.” (Latin)
“Castyn, or brakyn” ( <i>vomo</i> )		
“Castyn, or trowyn” ( <i>iacto, jacio</i> )		
“ffallyn, or ouer trowyn” ( <i>cado, ruo</i> )		
“ffalyn, or happyn” ( <i>accidit, eacenet</i> )		
“Gettyn, or wynnyn” ( <i>lucror, obtimeo</i> )	“Gettyng, or havying be wynnyng” ( <i>lucrum,</i> <i>adquisicio</i> )	
“Puttyn, or leyn” ( <i>pono,</i> <i>colloco</i> )	“Puttynge, or leyyng” ( <i>posicio, collacio</i> )	“Put, or leyd” ( <i>positus</i> )
	“Puttynge, or schowyng” ( <i>pulsus</i> )	
“Settyn and plantyn” ( <i>planto</i> )	“Settynge, or plangyng” ( <i>plantacio</i> )	
“Settyn and ordeynyn” ( <i>statuo</i> )		
“Settyn, puttyn, or leyn” ( <i>pono, colloco</i> )	“Settynge, or puttyng” ( <i>plantacio</i> )	“Sette, or putte” ( <i>positus</i> )
“Takyn, or resevyn” ( <i>accipio, capio,</i> <i>apprehendo, sumo, tollo</i> )	“Takyng, or resevyng” ( <i>accepio, omnia, captura</i> )	
“Throwyn, or castyn” ( <i>iacto, projicio</i> )	“Throwyng, or castyng” ( <i>iactura, iactus</i> )	

Table 3: Entries for verbs in the *Promptorium* where their senses are “not specified” in the head-words

“inf.” (Latin)	“ger.” (Latin)	“pp.” (Latin)
		“Cast, or castyde” ( <i>iactus</i> )
“Cum, or come” ( <i>venio</i> )		
	“ffallyng” ( <i>defectus</i> )	“ffalle” ( <i>casus</i> )
“Goon” ( <i>pergo, vado, eo, gradior</i> )		
“Makyn, or make” ( <i>facio, compono, psalmo</i> )		

### 3.1 The provision for entries on infinitival and conjugated verb forms

Out of the 26 entries provided for the 11 individual verbs, 14 are for infinitival forms, 8 are for gerundive and 4 are past-participial, the detail of which is as indicated in Table 4.

Table 4: The number of entries for infinitival and conjugated forms of individual verbs in the *Promptorium*

	<i>cast</i>	<i>come</i>	<i>fall</i>	<i>get</i>	<i>go</i>	<i>make</i>
inf.	2	1	2	1	1	1
ger.	—	—	1	1	—	—
pp.	1	—	1	—	—	—
total	3	1	4	2	1	1
	<i>put</i>	<i>set</i>	<i>take</i>	<i>throw</i>	<i>turn</i>	total
inf.	1	3	1	1	—	14
ger.	2	2	1	1	—	8
pp.	1	1	—	—	—	4
total	4	6	2	2	—	26

### 3.2 The discrimination and indication of the senses of the individual verbs and conjugated forms of verbs

What is important here is why the *Promptorium* provides such a number of entries for 11 individual verbs. Examining this more closely, we see that in 20 entries out of the 26 head-words, which make up 76.9%, the dictionary indicates the senses of the verbs, as reflected in Table 2. shows the situation numerically.

Table 5: the number of entries on infinitival and conjugated forms of individual verbs with the indication of their senses in the *Promptorium*

	<i>cast</i>	<i>come</i>	<i>fall</i>	<i>get</i>	<i>go</i>	<i>make</i>
inf.	2	—	2	1	—	—
ger.	—	—	—	1	—	—
pp.	—	—	—	—	—	—
total	2	—	2	1	—	—
	<i>put</i>	<i>set</i>	<i>take</i>	<i>throw</i>	<i>turn</i>	total
inf.	1	3	1	1	—	11
ger.	2	2	1	1	—	7
pp.	1	1	—	—	—	2
total	5	6	2	2	—	20

The author of the *Promptorium* provides more than one entry of the infinitival forms of verbs in Table 5, *cast*, *fall* and *set*, indicating their senses in English. What we see here is the first discrimination of the senses of English high-frequency verbs in a dictionary.

Besides, considering the number of entries for conjugated verb forms in Table 5, the *Promptorium* provides 9 entries for forms whose head-words include the senses of verbs. Returning to Table 2, this shows 2 cases where entries are provided for infinitival, gerundive and past-participial forms whose senses are the same, under the head-words:

“**Puttyn**, or leyn”, “**Puttynge**, or leyynge” and “**Put**, or leyd”.

“**Settyn**, puttyn, or leyn”, “**Settynge**, or puttynge” and “**Sette**, or putte”.

Table 2 also shows 4 cases where entries are provided for infinitival and gerundive forms whose senses are the same: “**Gettyn**, or wynnyn” and “**Gettyng**, or

havying be wynnyng”; “Settyn and plantyn” and “Settynge, or plangynge”; “Takyn, or resevyn” and “Takyng, or resevyng”; and “Throwyn, or castyn” and “Throwyng, or castynge”.

#### 4 The *Promptorium*'s treatment of phrasal verbs

Having analyzed entries for the 11 verbs in the *Promptorium* individually, I will now deal with those on idiomatic phrases related to these verbs in the following two sections, beginning with its treatment of phrasal verbs.

There are quite a few concepts of what a phrasal verb is. In this section, I define it as the following, lumping together what are called both literal meanings and figurative meanings

- (i) that regarded as a transitive verb: an intransitive verb + a preposition(s)/a transitive verb + an adverbial particle(s).
- (ii) that regarded as an intransitive verb: an intransitive verb + an adverbial particle(s).

Based on this definition, *Tables 6 and 7*, show entries for phrasal verbs, like those in the analysis of entries for individual verbs in the previous section.

*Tables 8 and 9*, on the number of entries have been produced in accordance with *Tables 6 and 7*, since adopting this method reveals the treatment of phrasal verbs in the *Promptorium*.

These numerical tables, show two notable facts; that is, the tendencies apparent in *Tables 8 and 9* are in stark contrast with those in the case of individual verbs in two respects. One is that the *Promptorium* only provides a small number of entries for the conjugated forms of phrasal verbs; in fact, 7 (1+4+2) entries out of 27 (7+20), which make up 25.9%, as compared to the entries for individual verbs where 12 entries out of 26 are given for their conjugated forms, accounting for 46.2%. The other is that its provision of entries for phrasal verbs whose head-words include their senses are also small in number. The dictionary gives 20 entries for individual verbs whose head-words include their senses out of 26, accounting for 76.9%, but provides only 7 entries for phrasal verbs whose head-words include their senses out of 27 (7+20), which make up 26.9%.

These two facts may give the impression that the author's treatment of phrasal verbs was quite slovenly and far from satisfactory. Still, he actually considers 27 such idioms related to the 11 verbs, which can hardly be thought small in number for the first dictionary that lists these idioms. How then can we interpret

## 5 English high-frequency verbs in the *Promptorium Parvulorum*

Table 6: Entries on phrasal verbs in the *Promptorium* where their senses are “specified” in the head-word

“inf.” (Latin)	“ger.” (Latin)	“pp.” (Latin)
“Cum aftyr, or folow” ( <i>succedo, sequor</i> )		
“Goyn abowten, or wynlyllyn” ( <i>circumeo</i> )		
“Go behynd, or folow” ( <i>sequor, descenao</i> )		
“Goo to, or begyn a dede” ( <i>aggredior</i> )		
“Puttyn a-wey, or refusyn” ( <i>repudio, refuto</i> )		
“Puttyn forth, as man doth his hand or oder lyke” ( <i>porrigo, extendo</i> )		
		“Put togyder, or onyd” ( <i>contiguus</i> )

these contradictory situations? There are two answers to this question. One is the extension of Gabriele Stein’s analysis alluded to in Section 1, that is, the author of the *Promptorium* changed the Latin-English entries of his sources into English-Latin ones, the English lemma still being basically an explanation of a former Latin-English dictionary, which applied to his treatment of individual verbs and phrasal verbs, as well. However, does this answer clarify the reason for the stark contrast between his treatment of individual verbs and that of phrasal verbs which I have shown, as well as the reason for his treatment of quite a few phrasal verbs?

Another is that the author of the *Promptorium* basically treats usual phrasal verbs at his time whose senses are not necessary to be explained in head-words in the relevant entries. From this viewpoint, the reason why the indications of the senses in head-words, as seen in “Cum aftyr, or folow”, “Goo to, or begyn a dede”

Table 7: Entries for phrasal verbs in the *Promptorium* where their senses are “not specified” in the head-words

“inf.” (Latin)	“ger.” (Latin)	“pp.” (Latin)
“Castyn a-wey” ( <i>abico, proicio</i> )	“Castyng downe or a-wey” ( <i>proieccio</i> )	“Cast downe” ( <i>deiecio, prosterno</i> )
“Cum dowyn” ( <i>descendo</i> )		
“Cum in” ( <i>ingredior, introeo, ineo</i> )		
“Cum to” ( <i>advenio</i> )	“ffallyng dowyn” ( <i>epilencia</i> )	
“Goone a-forne” ( <i>precedo</i> )		
“Goon aftyr” ( <i>sucedo</i> )		
“Goon a-wey” ( <i>recedo, descedo</i> )		
“Goo forthe” ( <i>procedo</i> )		
“Gon owte” ( <i>exeo, egredior</i> )		
“Puttyn aftyr” ( <i>propono</i> )		
“Puttyn a-fore” ( <i>prepono</i> )		
“Puttyn a-wey” ( <i>depono, expello, depello</i> )		
“inf.” (Latin)	“ger.” (Latin)	“pp.” (Latin)
	“Puttyn owt, or away” ( <i>eruo</i> )	
		“Putt to anythyng” ( <i>appono</i> )
“Takyn a-wey by strength or violens” ( <i>extorqueo</i> )	“Throwyng downe fro hey place” ( <i>precipicium</i> )	
	“Turnyng abowte” ( <i>versio, giriversio</i> )	

5 English high-frequency verbs in the *Promptorium Parvulorum*

Table 8: The number of entries for phrasal verbs “with” the indication of their senses in the *Promptorium*

	<i>cast</i>	<i>come</i>	<i>fall</i>	<i>get</i>	<i>go</i>	<i>make</i>
inf.	—	1	—	—	3	—
ger.	—	—	—	—	—	—
pp.	—	—	—	—	—	—
<b>total</b>	—	1	—	—	3	—
	<i>put</i>	<i>set</i>	<i>take</i>	<i>throw</i>	<i>turn</i>	<b>total</b>
inf.	2	—	—	—	—	6
ger.	—	—	—	—	—	—
pp.	1	—	—	—	—	1
<b>total</b>	3	—	—	—	—	7

Table 9: The number of entries on phrasal verbs “without” the indication of their senses in the *Promptorium*

	<i>cast</i>	<i>come</i>	<i>fall</i>	<i>get</i>	<i>go</i>	<i>make</i>
<b>inf.</b>	1	3	—	—	5	—
<b>ger.</b>	1	—	1	—	—	—
<b>pp.</b>	1	—	—	—	—	—
<b>total</b>	3	3	1	—	5	—
	<i>put</i>	<i>set</i>	<i>take</i>	<i>throw</i>	<i>turn</i>	<b>total</b>
<b>inf.</b>	4	—	1	—	—	14
<b>ger.</b>	—	—	—	1	1	4
<b>pp.</b>	1	—	—	—	—	2
<b>total</b>	5	—	1	1	1	20

and “**Puttyn a-wey**, or *refusyn*”, signify that the relevant phrasal verbs concern those with figurative, not literal, meanings, seems to become understandable, the indications being a special type of annotation to avoid confusing readers. If this is the case, it is no wonder that the author intended to list as many phrasal verbs as possible in the dictionary.

## 5 Its treatment of other notable phrases

The term “notable phrases” in the title of this section, means phrases which are thought to be common at the time of the author of the *Promptorium*. This means that the head-words of entries like the following, which only explain the senses of the Latin words provided, are totally irrelevant to such phrases:

Turnyng fro bad to goode: *Conversio*, -is; fem., 3.

Turnyng fro gode to bad: *Peruersio*, -is; fem., 3.

Turnyng of divers ways: *Diverticu lum*, -i; neut., 2 [...]: *Diversiclinicum*, -ij; neut., 2 [...].

From this point of view, entries concerning *make*, which constitute 47 of the 128 entries related to the 11 verbs considered here, which make up 36.7%, are very special. Of these 47, 45, most of which concern *make* as a causative verb, seem to be provided for standard phrases at the time of the author of the dictionary. For this reason, I will mainly investigate entries for idiomatic phrases related to *make*, pointing out remaining entries on such phrases by way of addendum at the end.

### 5.1 Notable phrases related to *make*

The head-words of the 45 entries with Latin words provided are as indicated in *Table 10*; the abbreviation “RAEI” in the parentheses standing for “reference to another entry indicated”.

This table, shows two notable facts. One is that all head-words of the 45 entries are infinitival forms and the other is that none of the head-words include the senses of relevant phrases, which may strengthen the possibility that the author took trouble of giving such indications only when necessary, far from simply inverting Latin head-words and English explanations in Latin-English dictionaries which may have been his sources, which I have suggested in the previous section.

5 English high-frequency verbs in the *Promptorium Parvulorum*

Table 10: Notable head-words in entries concerning *make* in the *Promptorium*

1	“Makyn able” ( <i>abilito</i> )	2	“Makyn a seth” ( <i>satisfacio</i> )	3	“Makyn better” ( <i>melioro</i> )
4	“Makyn bytter” ( <i>exacerbo, similiter</i> )	5	“Makyn blake” ( <i>denigro</i> )	6	“Makyn blunte” ( <i>obtundo</i> )
7	“Makyn clene” ( <i>mundo, purifico</i> )	8	“Makyn cumnawnte” ( <i>pango</i> )	9	“Makyn drunkyn” ( <i>inebrio</i> )
10	“Make dul” ( <i>hebito, obtundo</i> )	11	“Makyn evyn” ( <i>equo</i> )	12	“Make fette, or fatte” ( <i>inpinguo, sagino</i> )
13	“Make fowle” ( <i>deturpo, sordido</i> )	14	“Make gay” ( <i>orno</i> )	15	“Make fre” ( <i>manumitto</i> )
16	“Makyn harde” ( <i>induro</i> )	17	“Makyn hevy in herte or sory” ( <i>contrusto, mestico</i> )	18	“Makyn hevy in wytte” ( <i>grauo</i> )
19	“Make Ioy” (RAEI)	20	“Make knowyn” ( <i>manifesto, natifico</i> )	21	“Make large” ( <i>amplio</i> )
22	“Make lawful” ( <i>legitimo</i> )	23	“Make lene” ( <i>macero</i> )	24	“Make lesse” ( <i>minoro</i> )
25	“Make meynd” ( <i>commemoro</i> )	26	“Make wonderful, or mervelyows” ( <i>mirifico</i> )	27	“Make mery and gladyn oder mene” ( <i>letifico</i> )
28	“Make myry, or be mery in hert or chere” ( <i>letor, iocundor</i> )	29	“Makyn more” ( <i>maioro</i> )	30	“Make nesch” ( <i>mollifico</i> )
31	“Makyn perfyte” ( <i>perficio</i> )	32	“Make playne” ( <i>plano</i> )	33	“Make pleyne” ( <i>conqueor</i> )
34	“Make plentyows” ( <i>ffecundo</i> )	35	“Make qveynte, or wonderful” ( <i>mirifico</i> )	36	“Make redy” ( <i>paro</i> )
37	“Make rych” ( <i>dito</i> )	38	“Make paste” ( <i>intero</i> )	39	“Make sacrifice” ( <i>sacrifico</i> )
40	“Make sekyr in grawnte” ( <i>ratifico, confirmo</i> )	41	“Make sythy” ( <i>elucido</i> )	42	“Make solemnyte” ( <i>solemnyzo</i> )
43	“Make tokyn to a noder” ( <i>nuo, annuo</i> )	44	“Make wery” ( <i>ffatigo, lasso</i> )	45	“Make with chylde” ( <i>impregno</i> )

Incidentally, it may be interesting to note that there are two cases in which the author distinguishes the use of phrases: (1) “**Makyn hevyn** in herte or sory” (No. 17) and “**Makyn hevyn** in wytte” (No. 18); and “**Make mery** and gladyn oder mene” (No. 27) and “**Make myry**, or be mery in hert or chere” (No. 28). The senses of *makyn hevyn* and *make mery/myry* being not indicated, this may be regarded as the reflection of the author’s consciousness of the difference of semantic domains between the English phrases and Latin words provided for each. If this is the case, it may indicate his refined sense of language.

## 5.2 Remaining notable phrases

Table 11 shows the situation of the head-words in 10 remaining entries which are thought to concern usual phrases at the author’s time.

Table 11: Remaining entries on notable phrases in the *Promptorium*

“inf.” (Latin)	“ger.” (Latin)	“pp.” (Latin)
		“Cast werk, or disposin” ( <i>dispono</i> ) “Cast lott” ( <i>sortior</i> )
“Getyn by prayers” ( <i>impetro</i> )		
“Goon in-to a place” ( <i>introeo, ingredior</i> )		
“Goon on forte” ( <i>pedito</i> )		
“Go slowly” ( <i>lento</i> )		
“Go wrong” ( <i>devio, deliro</i> )		
“Settyn at noʒte, or disposin” ( <i>vilipendo, ffloccipendo, nichilpendo</i> )		
“Settyn in yese and reste” ( <i>quieto</i> )		
	“Takyn on honde” ( <i>manucapio</i> )	

This table shows that the author only allocates senses to the phrases *cast werk* and *setty n at nozt*, providing two head-words of conjugated forms, other head-words being infinitival forms without the indications of senses.

## 6 Conclusion

Having explored the treatment of infinitival and conjugated forms of English high-frequency verbs, as well as that of phrasal verbs and other idiomatic phrases, in the *Promptorium*, we may now safely say that it is the earliest known dictionary that substantially exhibits the features I mentioned at the beginning of this chapter, so that “the *Promptorium*, which was compiled as a ‘storehouse’ for children in the 1440’s, is the first dictionary that treated quite a few high-frequency verbs, distinguishing their senses and indicating their conjugated forms, with listing their relevant idiomatic phrases, especially phrasal verbs.

It is not that authorities thus far have not approached this point. In her meticulous analysis of her 1997 book, Gabriele Stein investigates its head-words from a comprehensive and overall perspective, dealing with their prepositional phrases, past participles, infinitives and infinitive clauses, *etc.* and has clarified the whole picture of the dictionary. I consider that after a holistic analysis a specialized analysis should follow. From this viewpoint, I aimed to conduct a specific analysis of the treatment of English high-frequency verbs in the dictionary which Stein seems not to have done in relation to the history of lexicography.

One more point. If there is a grave misunderstanding that Samuel Johnson was the first lexicographer who detailed the high-frequency verbs, it may be a consequence of exaggerating the movement to ascertain the English language which arose at the beginning of the Late Modern English period. I do not deny that the movement was historically large in scale and that it actually induced Johnson to compile his *Dictionary*. At the same time, however, we should not disregard the probability that there had been lexicographers of bilingual and polyglot dictionaries before him who had greatly troubled themselves over the treatment of the high-frequency verbs. I claim that whoever issued the *Promptorium* more than three hundred years before Johnson’s *Dictionary*, or more than one hundred years before the birth of William Shakespeare, was one of such lexicographers.

## Acknowledgments

I want to express my sincere gratitude to Professor Emeritus John Considine (the University of Alberta in Canada) for his invaluable and inspiring suggestions in writing this paper. Remaining errors, if any, are my own. Since my presentation at ICHLL (International Conference for Historical Lexicography and Lexicology)

in 2022, on which this chapter is based, Professor Emeritus Jean-Louis Duchet kindly informed me that the *Promptorium* also lists verbs and verbal phrases which were specific at the time, such as *delvyn owt off þe earth*, which may broaden the horizon of my future research; I also want to extend to him my sincere gratitude.

## References

- Alston, R. C. 1968. *Promptorium parvulorum 1499* (English Linguistics 1500-1800 108). Menston: The Scolar Press Limited.
- Bately, Janet. 2009. Bilingual and multilingual dictionaries of the renaissance and early seventeenth century. In Anthony P. Cowie (ed.), *The Oxford history of English lexicography*, vol. 1, 41–64. Oxford & New York: Oxford University Press.
- Cowie, A. P. (ed.). 2009. *The Oxford history of English lexicography*, vol. 1. Oxford & New York: Oxford University Press.
- Franzen, Christine (ed.). 2012. *Ashgate Critical Essays on Early English Lexicography*, vol. 2: Middle English. Farnham & Burlington: Ashgate.
- Hüllen, Werner. 2006. *English Dictionaries, 800-1700: The Topical Tradition*. Oxford & New York: Oxford University Press.
- McDermott, Anne & Rosamund Moon. 2005. Introduction: Johnson in context. *International Journal of Lexicography* 18(2). 153–156.
- Sauer, Hans. 2009. Glosses, Glossaries and Dictionaries in the Medieval Period. In Anthony P. Cowie (ed.), *The Oxford History of English Lexicography*, vol. 1, 17–40. Oxford & New York: Oxford University Press.
- Starnes, DeWitt Talmage & Gertrude Elizabeth Noyes. 1946. *The English Dictionary from Cawdrey to Johnson, 1604-1755*. Chapel Hill: University of North Carolina Press.
- Stein, Gabriele. 1985. *The English Dictionary Before Cawdrey*. Tübingen: Max Niemeyer Verlag.
- Stein, Gabriele. 1997. *John palsgrave as renaissance linguist: A Pioneer in Vernacular Language Description*. Oxford & New York: Oxford University Press.
- Wells, Ronald A. 1973. *Dictionaries and the Authoritarian Tradition: A study in English Usage and Lexicography*. The Hague & Paris: Mouton & Co. N.V., Publishers.
- Wheatley, Henry B. 1865. Chronological Notices of the Dictionaries of the English Language. *Transactions of the Philological Society* 10(1). 218–293. DOI: 10.1111/j.1467-968X.1865.tb00472.x.


## Chapter 6

# Disattributing the *Encyclopédie* article on *définition en logique* from Jean-Henri-Samuel Formey

Alexander K. Bocast

Berkeley Bridge Press, Anacortes, Washington, USA

I argue that Denis Diderot's attribution of the article *définition en logique* to Jean-Henri-Samuel Formey in the *Encyclopédie* entry for DÉFINITION cannot be justified. First, I lay the groundwork for the argument by analyzing the sources, roles, and attributions for each paragraph in the whole entry on DÉFINITION. The origins of 40 of the 42 paragraphs in DÉFINITION are unambiguously identified; these 40 paragraphs are either original contributions or traceable to specific sources. These data demonstrate that the *Encyclopédie* attributions for *définition en mathématiques* to Jean le Rond d'Alembert and *définition en rhétorique* to Edme-François Mallet and Jean François Marmontel are warranted. However, no paragraphs or passages in *définition en logique* can be traced to Formey as *author*. Second, having found no reason to consider Formey the author of *définition en logique*, I ask whether Formey could have translated, compiled, or edited any part of this article. Regarding Formey as translator, having analyzed Formey's bilingual edition of Savile's *Advice to a Daughter*, I find no evidence that Formey could have translated the English sources found in the article. Regarding Formey as compiler, I show that the treatment of definition in *définition en logique* is not consistent with Formey's writings on *definition*. When Formey wrote about *definition*, he consistently used Christian Wolff's philosophical terms. However, three paragraphs in *définition en logique* ridicule precisely those terms, which are not found elsewhere in the article. There is no reason to believe that Formey would have compiled an article from sources hostile to his own views. Regarding Formey as editor, I find no evidence that he had any editorial role in the preparation of *définition en logique* or in the larger production of the *Encyclopédie* itself. Instead, Étienne Bonnot de Condillac and d'Alembert are the editors most likely to have touched *définition en logique*. Thus, my argument stands that Formey should be *disattributed* from the article *définition en logique*.

Alexander K. Bocast. 2025. Disattributing the *Encyclopédie* article on *définition en logique* from Jean-Henri-Samuel Formey. In Geoffrey Williams, Mathilde Le Meur & Andrés Echavarría Peláez (eds.), *West meets East: Papers in historical lexicography and lexicology from across the globe*, 77–98. Berlin: Language Science Press. DOI: 10.5281/zenodo.15394479 © 



## 1 Introduction and plan of analysis

While translating the *Encyclopédie*'s DÉFINITION from French into English for a project on Gergonne's theory of *definition*,<sup>1</sup> I found this anomaly: although the article *définition en logique* is attributed to Jean-Henri-Samuel Formey, I could find no trace of Formey in the article itself. To understand this anomaly, I ask what relation, if any, did Formey have to this article and, if so, would this relation merit this attribution? Few detailed analyses of individual *Encyclopédie* articles have been published,<sup>2</sup> so this information may also contribute to other studies of the *Encyclopédie*.

First, I look at the textual sources for the *Encyclopédie*'s entry for DÉFINITION and the articles it contains. With this detailed information in hand, gathered at the level of individual paragraphs, I look at relationships, if any, between attributions made within DÉFINITION by the editors of the *Encyclopédie* and the articles so attributed.

Second, I ask whether the attribution of the article on *définition en logique* to Formey is warranted. Attributions in the *Encyclopédie* tend to be ambiguous; they lump authors, compilers, translators, and editors together. They generally fail to distinguish whether an attributed person provided an original composition, assembled material from existing texts, translated an existing text, or edited an article supplied by other contributors. Would it be reasonable to assert that Formey is associated with this article in any of those roles?

## 2 Analysis of the whole entry DÉFINITION

### 2.1 Overview of DÉFINITION

The *Encyclopédie* entry DÉFINITION is substantial; it weighs in at seven and a half columns. It contains three named articles: the first treats *définition en logique*; the second discusses *définition en mathématiques*; and the third is focused on *définition en rhétorique*. The longest of these articles, by far, is the article on *définition en logique*, at almost six columns.

---

<sup>1</sup>See Diderot and d'Alembert (1751-1772: s.v. *définition*). This entry appears in the 4th volume, published in 1754. For my translation of this entry, see its three parts (Formey 2016, d'Alembert 2016, Mallet & Marmontel 2016) at *The Encyclopedia of Diderot & d'Alembert Collaborative Translation Project* <https://quod.lib.umich.edu/d/did/>. For Gergonne's thoughts on definition, see Gergonne (1818).

<sup>2</sup>One such study is Coste (2009).

The article on *définition en logique* is attributed as an ‘article of monsieur Formey’,<sup>3</sup> implying that Formey wrote that article. The article on *définition en mathématiques* is attributed to Jean le Rond d’Alembert, and I suppose that, as an editor of the *Encyclopédie*, d’Alembert would know what d’Alembert wrote. The third article, on *définition en rhétorique*, is attributed in part to Edme-François Mallet and in part to Jean François Marmontel.

Now to the sources of these texts. The first ten paragraphs of *définition en logique* and the first four paragraphs of *définition en rhétorique* are taken directly from the entry DEFINITION in Ephraim Chambers’s *Cyclopædia*, in the order that those paragraphs are found in Chambers’s entry. These two articles are also presented in the same order, that is, *definition in logic* precedes *definition in rhetoric* as *définition en logique* comes before *définition en rhétorique*. However, Chambers’s account of DEFINITION does not contain an article that corresponds to the *Encyclopédie* article on *définition en mathématiques*.<sup>4</sup>

## 2.2 Définition en logique

The article *définition en logique* comprises 32 paragraphs. The origin of 30 of these paragraphs is definitely known. The source of one paragraph can be confidently deduced. One paragraph is merely a one-sentence editorial segue, without substance. Table 1 summarizes the sources, roles, and attributions for each paragraph in *définition en logique*. It shows that Diderot’s attribution of this article to Formey has no grounds in any role related to the content of this article.

In all, eleven of the 32 paragraphs of *définition en logique* are taken from Chambers; this is more than one-third of the paragraphs in the *Encyclopédie* article. As well, four of the five paragraphs in *définition en rhétorique* were taken from Chambers. This preponderance of Chambers in their corresponding *Encyclopédie* articles strongly suggests that those paragraphs are based upon the translation delivered by John Mills and Gottfried Sellius to the unscrupulous publisher André le Breton in 1745. I will return to this when I ask whether Formey could himself have translated Chambers’s articles from English into French.

The next major source for *définition en logique* is the writings of Étienne Bonnot de Condillac, specifically his *Essai sur l’origine des connoissances humaines* (Condillac 1746).<sup>5</sup> The 1746 date of publication of Condillac’s *Essai* is, I think, sig-

<sup>3</sup>‘Article de M. Formey.’

<sup>4</sup>For a detailed analysis of Chambers’s entry on *definition* and his sources, see Bocast (2016).

<sup>5</sup>A few years later, Thomas Nugent translated Condillac’s *Essai sur l’origine des connoissances humaines* into English. Nugent emphasized Condillac’s relation to John Locke when he titled his translation *An Essay on the Origin of Human Knowledge: Being a Supplement to Mr. Locke’s Essay on the Human Understanding* (Condillac 1756).

Table 1: Sources and roles for the *Encyclopédie* article “définition en logique”.

définition en logique		editors		paragraphs	<i>Encyclopédie</i>
Author	Compiler	first	final	count=32	source
	Chambers ; translated by Mills & Sellius			¶¶ 1–10 ¶¶ 22	Cyclopædia (Chambers 1741: s.v. definition)
Buffier	Condillac			¶¶ 11–13 ¶¶ 20–21	Traité des premières veritez (Buffier 1724: pt. 2, ch. 7)
Condillac	Condillac			¶¶ 14–19 ¶¶ 24 ¶¶ 26–27  ¶¶ 23	Essai sur l’origine des connoissances humaines (Condillac 1746: pt. 1, §§ 3 & 5; pt. 2, § 1) one sentence segue; editorial glue
Locke; translated by Coste	Condillac	Condillac	d’Alembert	¶¶ 25	Essai philosophique concernant l’entendement humain (Locke 1700b: bk. 3, ch. 4)
Condillac				¶¶ 28	presumably Condillac, but unresolved
Port- Royal	Condillac			¶¶ 29–32	La Logique, 6th ed. (Arnauld & Nicole 1712: pt. 1, ch. 12 & 13

nificant. Condillac was at the time in Paris, dining weekly with Denis Diderot and Jean Jacques Rousseau, and contributed somewhat to the early editing of the *Encyclopédie* (d'Alembert 1995: xiii–xiv; noted in Schwab's introduction). I believe that both the Mills/Sellius translation *and* a treatment of *définition* from Formey fell into Condillac's hands for editing.

At that time, two intellectual movements dominated the Enlightenment of western Europe. One centered on the Englishmen Isaac Newton and John Locke; and the other followed the Germans Gottfried Wilhelm Leibniz and Christian Wolff. (Remember that Newton and Leibniz both claimed to have invented the calculus.) Condillac favored Newton and Locke, while Formey was a quite orthodox proponent of Wolff's philosophy. Again, I will come back to Condillac's contribution when I ask whether the *Encyclopédie* article on *définition en logique* is consistent with Formey's commitment to Wolff.

Next comes Claude Buffier, with extracts taken from his *Traité des premières veritez, et de la source de nos jugemens* (Buffier 1724). His five paragraphs bracket the longest run of paragraphs taken from Condillac's *Essai*. Buffier was a Jesuit, anathema to the Huguenot Formey, who took a psychological approach to epistemology that was not inconsistent with Condillac's primacy of sensation as the origin of knowledge.

Buffered between paragraphs traced to Condillac, there is a paragraph taken from John Locke's *Essay Concerning Human Understanding* (Locke 1700a). In consultation with Locke, Pierre Coste translated Locke's *Essay* into French as the *Essai philosophique concernant l'entendement humain* (Locke 1700b). The long paragraph found in the *Encyclopédie* article is taken directly from Coste's translation.

Of course, if a Francophone intellectual discussed *définition* in the 18th century, there was sure to be some recourse to the *Port Royal Logic*, that is, to the *La Logique ou l'Art de penser* of Antoine Arnauld et Pierre Nicole (Arnauld & Nicole 1712). The final four paragraphs of the article on *définition en logique* are taken from part 1, chapters 12 and 13, it seems of the 6th edition of 1712.

There are only two other paragraphs in the article *définition en logique* to mention. One consists of a single short sentence: 'In general, we can make the following *observations* about definition'.<sup>6</sup> It is followed by six ordinarily-identified observations. The first two were taken from Condillac, the third is of unknown provenance, and those remaining are found in the *Port-Royal Logic*. That third observation is given by the 28th paragraph; it consists of a single sentence, but is of respectable length: 'Third, definitions never help us know the nature of substances, but only essences that we conflate with notions that we form of

---

<sup>6</sup>On peut faire sur la définition en général les réflexions suivantes.'

things, notions based on *archetypal* ideas and not on actually existing real things such as substances.<sup>7</sup> The word *archetypal* certainly suggests Condillac as *author*, just as the ordinal ordering of these observations suggests Condillac as *editor*. However, I have not traced this fragment to some specific passage in Condillac; it may have been constructed from Condillac's words by another *Encyclopédie* editor.

This accounts for the whole of the text of the article on *définition en logique*, excepting of course a number of details and transitions within paragraphs. Of the 32 paragraphs of the article on *définition en logique*, 30 paragraphs can be definitively traced to specific passages in the works of Chambers, Buffier, Condillac, Coste's Locke, and the *Port-Royal Logic*. No passages can be traced to the pen of Formey.

The article on *définition en mathématiques* is not related to Chambers's DÉFINITION, which does not contain an article on *definition in mathematics*. It is attributed to d'Alembert, perhaps by d'Alembert, and I have found no reasons to doubt that assertion. This attribution assures a reader that d'Alembert wrote *définition en mathématiques*. Further, *définition en mathématiques* contains references to the physically-preceding *définition en logique*; to me, these references argue that *définition en mathématiques* was prepared after or alongside *définition en logique*. Throughout the *Encyclopédie*, signatures of d'Alembert are frequently associated with attributions to Formey;<sup>8</sup> it is reasonable to suppose that d'Alembert also played some editorial role in the production of *définition en logique*. The attribution to d'Alembert recognizes his role as *author* of the article.

### 2.3 Définition en rhétorique

Now I can briefly consider the article on *définition en rhétorique*. The first four paragraphs account for the whole of Chambers's corresponding *definition in rhetoric* as translated by Mills & Sellius;<sup>9</sup> I assume their translation was fairly literal. Diderot presumably tasked Edme-François Mallet 'to review, correct, and augment' the article seeded by that translation.<sup>10</sup> I deduce that Mallet completed

---

<sup>7</sup>3°. Les définitions ne nous aident jamais à connoître la nature des substances, mais seulement les essences qui se confondent avec les notions que nous nous faisons des choses ; notions fondées sur des idées archétypes, & non pas d'après des modeles réellement existans, ainsi que sont les substances."

<sup>8</sup>For an example, see Coste (2009).

<sup>9</sup>Chambers's sources for these paragraphs are examined by Bocast (2016: 203-225).

<sup>10</sup>"Mais en présentant à chacun un rouleau de papiers, qu'il ne s'agissoit que de revoir, corriger, augmenter". (Diderot & d'Alembert 1751: s.v. encyclopédie)

this task because Diderot attributed these paragraphs to him.<sup>11</sup> Subsequently, Diderot asked Marmontel to review and improve articles prepared by Mallet, who was prolific but may have been sometimes overly hasty.<sup>12</sup> Diderot's attribution "Article de M. Marmontel" at the end of the article appears to confirm Marmontel's editorial review of Mallet's editorial review.

As printed, the text given to Mallet has been carefully improved and expanded, but the article does not venture further than Chambers's original paragraphs. Cumulative changes to Chambers's article by Mills and Sellius, Mallet, and Marmontel cannot be sorted out, but whatever changes Marmontel might have made did not displace the attribution to Mallet at the end of the fourth paragraph. The attribution to Marmontel is found later in the text, at the end of the fifth and final paragraph.

However, beyond physical proximity, the fifth paragraph is wholly unrelated to the preceding article *définition en rhétorique*. This divide and this paragraph are seen again in Panckoucke's *Encyclopédie méthodique* (Marmontel 1782), and Marmontel definitely claimed this text as his own in his *Oeuvres complètes* (Marmontel 1787: 7, 368-369). It seems that Marmontel appended this paragraph to Mallet's manuscript before returning it to Diderot; this new material addresses an aspect of 'philosophical definitions', specifically a metaphysical role in unifying intersubjective understandings of things held to be ontologically real. Mallet's *définition en rhétorique* and Marmontel's *définitions philosophiques* are distinguished by their respective use of internal references. Following Chambers's *definition in rhetoric*, Mallet's *définition en rhétorique* contains no pointers to other entries. In contrast, Marmontel supplied explicit references to entries for *idea*, *truth*, *evidence*, and *error*, entries primarily tied to philosophy rather than rhetoric.<sup>13</sup> Table 3 summarizes the sources, roles, and attributions for each paragraph in *définition en rhétorique*. It shows an attribution to Mallet that recognizes his role as primary editor of the paragraphs constituting the article *définition en rhétorique* and an attribution to Marmontel that recognizes his role as author of the final paragraph of the entry on *définition*.

Thus, Diderot's attribution of *définition en rhétorique*, especially given internal evidence and retrospective testimony from the *Encyclopédie méthodique* and Marmontel's *Œuvres complètes*, should be considered in light of these inter-

---

<sup>11</sup>The attribution symbol for Mallet is (G). It is placed at the end of the fourth paragraph in the article.

<sup>12</sup>"Beginning in volume iii, Marmontel was assigned . . . to revise before publication other articles that Mallet had drafted." (Kafker & Kafker 1988: 240).

<sup>13</sup>That is, references to the entries IDÉE, VÉRITÉ, ÉVIDENCE, and ERREUR.

Table 2: Sources, roles, and attributions for *définition en rhétorique*

<i>définition en rhétorique</i>		editors		paragraphs count=5	<i>Encyclopédie</i>	
Author	Compiler	first	final	attribution	source	
de Colonia; translated by Chambers	Chambers; translated by Mills & Sellius	Mallet	Marmontel	Mallet	¶¶ 1-4	Chambers (1741): s.v. definition
Marmontel		Marmontel	Marmontel	Marmontel	¶¶ 5	Marmontel

woven roles: *author*, Dominique de Colonia;<sup>14</sup> *translator & compiler*, Ephraim Chambers; *translator*, John Mills & Gottfried Sellius; *primary Encyclopédie editor*, Edme-François Mallet; and *secondary Encyclopédie editor*, Jean François Marmontel. Diderot’s attribution of *définition en rhétorique* should be read as affirming Mallet’s role as its *primary editor*.

Marmontel’s lone paragraph is wholly inadequate to encompass *definition* in philosophy.<sup>15</sup> I suspect that Diderot simply did not notice that Marmontel had gone beyond his remit to review and revise Mallet’s article. Thus, Diderot did not insert the typographic and syntactic cues that typically introduce an article and a new addressee within an *Encyclopédie* entry.<sup>16</sup> I think he simply added the attribution “*Article de M. Marmontel*” to the end of the manuscript before sending it off to the printers. Whether I have correctly intuited this story, sufficient evidence indicates that the intent of Diderot’s attribution to Marmontel was to recognize his contribution as the secondary *Encyclopédie* editor of *définition en*

<sup>14</sup>Chambers’s *definition in rhetoric* is taken almost entirely from Dominique de Colonia (1710). For a detailed look at de Colonia as a source for Chambers, see Bocast (2016: 206-228).

<sup>15</sup>Indeed, Marmontel’s expansion of this paragraph in the *Encyclopédie méthodique* is more than five columns long (Marmontel 1782). There Marmontel treats *definition* in legal argumentation, in laudatory speech and elegies, in poetry, and, finally, in philosophy.

<sup>16</sup>As printed, the article structure of the entry *définition* is revealed by typographic and syntactic cues, except for the final paragraph:

DÉFINITION, *en logique*, est . . . *Article de M. FORMEY.*

DÉFINITION, *en mathématiques*, c’est . . . (O)

DÉFINITION, *en rhétorique*, c’est . . . (G)

A l’égard des *définitions* philosophiques, elles sont . . . *Article de M. MARMONTEL.*

A sequence of consistent cues might have been:

DEFINITION, *en logique*, est . . . *Article de M. FORMEY.*

DÉFINITION, *en mathématiques*, c’est . . . (O)

DÉFINITION, *en rhétorique*, c’est . . . (G)

DÉFINITION, *en philosophie*, c’est . . . *Article de M. MARMONTEL.*

These cues mark four articles in this entry rather than the customary three.

*rhétorique*; instead, I believe this attribution should be read to acknowledge Marmontel as author of the last paragraph rather than as editor of the preceding four paragraphs.

### 3 Analysis of *définition en logique*

#### 3.1 Formey considered as author

Given this analysis of textual sources, the notion that Formey might have authored *définition en logique* is quickly disposed of. I have identified specific authors and sources for 30 of the 32 paragraphs in the article, and putative editors for the remaining two. None of these 32 paragraphs can be traced to Formey's writings. Thus, Formey is clearly *not* the *author* of this article.

#### 3.2 Formey considered as translator

##### 3.2.1 ... regarding Chambers's DEFINITION

Because a third of the article on *définition en logique* is taken from English sources, I ask whether Formey was himself capable of translating such material from English to French. Formey was of course intimate with French, and as Secretary of the Berlin Academy he translated works into French from German and from Latin.<sup>17</sup> However, I have no evidence that Formey conversed or wrote in English. He may have had some ability to read English in correspondence, but few letters directed to him as Secretary of the Academy were written in English.<sup>18</sup>

Formey's *Conseils pour former une bibliothèque* is indicative. In the first edition of 1746, he wrote 'but of more value than these [other dictionaries] for those who

<sup>17</sup>“Er übernahm die Verpflichtung, deutsch oder lateinisch abgefaßte wissenschaftliche Referate für die Sammelbände der Akademie zu übersetzen”. (Krauss 1963: 55).

Thomson (1980) related that “De 1745 à 1747, il servit comme traducteur du roi dans le Département des Affaires étrangères.” This was echoed by Kafker & Kafker (1988: 140): “from 1745 to 1747, [Formey] served as the royal translator for the Prussian Department of Foreign Affairs.”

Grosse notes: “Formey was engaged in the translation and publication of other authors' works: beginning with sermons and the writings of the German philosopher Christian Wolff, he later also translated the works of ancient philosophers and the scientific contributions of his German colleagues at the Academy.” (Grosse 2018: 10)

None of these suggest any knowledge of English.

<sup>18</sup>“Les éditeurs du présent livre ont choisi de publier 116 lettres écrites en français (trois seulement sont en anglais) de la correspondance passive de Formey. Leur unité affichée est dans leur commune provenance géographique : elles ont toutes été écrites, entre 1737 et 1788, d'Angleterre à Formey résidant à Berlin.” (Hubac 2007: 137)

understand English is Chambers's *Encyclopaedia*; we continue to look forward to a French edition',<sup>19</sup> where the words "pour ceux" imply that Formey himself is *not* among those who understand English. Here and consistently thereafter in those *Conseils* he misrepresented the *Cyclopædia* as the "*Encyclopaedia*", which indicates that he did not possess a copy of Chambers's work but knew it only by reputation. In the 2nd edition of 1750 – *after* he had delivered his manuscripts to the *Encyclopédie* – Formey added that a translation of the *Encyclopaedia* of Chambers was being printed in Paris under the direction of d'Alembert and Diderot.<sup>20</sup> He made no claim to have a stake or a role in this translation.

### 3.2.2 ... regarding Savile's Ladies-New-Years Gift

Chambers disappeared from the 1755 edition of Formey's *Conseils*; the *Encyclopédie* replaced the *Cyclopaedia*. However, this edition intimates that perhaps Formey could translate from English to French. Formey reported that the Berlin firm of *Haude & Spener* had asked him to prepare a new edition of George Savile's *The Ladies New-Years Gift: or, Advice to a Daughter*.<sup>21</sup> This he did 'by putting to the side of the English text a French translation,<sup>22</sup> which already existed, but

---

<sup>19</sup>«Enfin les Dictionnaires Philosophiques sont des repertoires utiles. Il y a en Latin celui de Chauvin, in folio, & en Allemand celui de Walch, en un gros in 8vo, mais ce qui vaut mieux que tout cela pour ceux qui entendent l'Anglois, c'est l'*Encyclopedie de Chambers*, dont on avoit fait esperer une Edition Française, que je ne sai quels contretens ont derangé.» (Formey 1746: 13).

On p. 56, Formey recommends the "*Encyclopédie de Chambers*" as a work of philosophy.

<sup>20</sup>«Enfin les Dictionnaires Philosophiques sont des répertoires utiles. Il y a en Latin celui de Chauvin, in folio, & en Allemand celui de Walch, en un gros in 8vo, mais ce qui vaut mieux que tout cela pour ceux qui entendent l'Anglois, c'est l'*Encyclopedie de Chambers*, dont la Traduction s'imprime actuellement à Paris, sous la direction de Mrs. d'Alembert & Diderot, avec des additions & des ameliorations, qui feront de cet Ouvrage un vrai Trésor.» (Formey 1750: 18).

On p. 80, Formey again recommends the *Encyclopédie de Chambers* as a work of philosophy.

<sup>21</sup>See (Savile 1688). Three editions were published that year in London and Edinburgh. The 4th edition is the last before Savile's death; it is the edition that was translated into French. Subsequent editions are merely reprints from different publishers; fifteen of these were published in London from 1688 through 1765. This work was also published in the three editions of *Miscellanies*, a collection of Savile's writings; see Savile (1700, 1704, 1717)).

<sup>22</sup>For his English text, Formey drew from the 13<sup>th</sup> (1748) London edition; see Savile (1752: "Avertissement"). I have been unable to consult that edition. However, the 12th edition and the 14th edition seem to be identical (Savile 1741, 1756), and I have found no discrepancies between the English text in Formey's edition and either of those editions.

For his French text, Formey had in hand the 1698 Den Haag translation; see Savile (1698, 1752: "Avertissement"). The 1698 edition is practically identical to the first published French translation (Savile 1692). Indeed, subsequent editions are merely reprints of the 1692 edition; thus, Formey *could* have used any of those four printings; see Savile (1692, 1699, 1748).

which I retouched in several places.<sup>23</sup> This might be read to suggest that Formey drew his changes directly from the English work, but an announcement in the *Bibliothèque impartiale* in 1752 should discourage such a reading: “The *Conseils d’un Père à sa fille* in French and English has just been reprinted by *Haude & Spener*. M. Formey, the editor, has recast the French version which was already out of date.”<sup>24</sup>

Formey went on to say that the 1698 “translation has aged and, moreover, it is not free from mistakes. I have carefully retouched it; and my goal, by setting the English and French texts side by side, is at the same time to provide a small workbook for people who are applying themselves to the study of the English language.”<sup>25</sup> Thus, the Haude & Spener edition is bilingual, setting French and English in parallel columns. The French text, before and after Formey, is more a paraphrastic approximation of the English text than an actual translation; it abounds with departures from Savile’s stated intent.

Well, I looked in detail at Formey’s *Conseils*, both the English and the French, and I compared them with each other, with Savile’s original English text, and with the previous printings of the *Conseils*.<sup>26</sup> I studied punctuation, orthography, verb tense, and end-of-sentence hyphenation;<sup>27</sup> relative density of typographical errors in both English and French texts; alignment of English to French text across page boundaries; widows, orphans, and runts, and the allotment and distribution of whitespace; alignment of lexical cognate terms and terminal punctuation in English and French; textual differences between previous editions of the

<sup>23</sup>“Le Libraire *Spener* m’ayant prié de procurer une nouvelle Edition d’un Ouvrage Anglois du Marquis d’*Halifax*, *Conseils d’un Père à sa Fille*, je le fis en mettant à côté du Texte une Traduction Française, qui existoit déjà, mais que je retouchai en plusieurs endroits. Cela fit un Volume in octavo, qui parut au commencement de 1753.” (Formey 1755: 120)

The resulting book had a different title: *Conseils d’un homme de qualité à sa fille*; see Savile (1752).

<sup>24</sup>“Les *Conseils d’un Père à sa fille* en Anglois & en François, ouvrage du célèbre Lord *Halifax*, viennent d’être réimprimés chez *Haude & Spener*. Mr *Formey*, qui en est l’Editeur & qui a refondu la Version Française & qui étoit déjà surannée”. (Luzac, Elie 1752: 1, 158).

The name of the bookseller *Johann Carl Spener*, of *Haude & Spener*, was misprinted in this passage.

<sup>25</sup>“Cette Traduction avoit vieilli; & n’étoit pas d’ailleurs exemte de défauts. Je l’ai soigneusement retouchée; & mon but, en mettant le Texte à côté, est en même tems de fournir un petit Manuel aux personnes qui s’appliquent à l’étude de la Langue Anglois.” (Savile 1752: “Avertissement”)

<sup>26</sup>I examined three chapters in Formey’s edition: the introduction, the 3rd chapter, and the 9th and final chapter. These sections contain 40 pages; the whole book has just 152 pages. In addition, I looked closely at the first five lines of every fifth page, starting with page 6. Altogether, I have examined about 30% of Formey’s French text and the corresponding texts in Savile’s English and in the previous *Conseils*.

<sup>27</sup>French hyphenation was used in both columns. This leads to such oddities in the English column as *enti-rely*, *li-ne*, *ru-le*, *mista-kes*, *star-ve*, and *ushe-red*.

*Conseils* and Formey's edition, including chapter titling;<sup>28</sup> Formey's substitution of synonyms for their equivalents; the fidelity of the previous French translation to Savile's original text; the changes made in the French text by Formey and how these changes relate to both Savile's text and the previous French translation; whether Formey's changes did fix mistakes or improve the translation;<sup>29</sup> and errors that Formey did *not* correct, among other considerations.<sup>30</sup>

With one possible exception,<sup>31</sup> I found no evidence that Formey consulted Savile's English text. I found no passage in Formey's edition that corrected any of the many infelicities in the original translation of 1698.

---

<sup>28</sup>Formey gives the mysterious title "Discours sur le Prochain" to a chapter titled "Censure" in Savile's English and "Des circonspections que doit apporter une Dame en parlant des gens" in previous *Conseils*.

<sup>29</sup>In at least one case, Formey found a grammatical mistake but his fix degraded the translation. On the first page of Formey's edition (Savile 1752: 1) is the passage "il est occupé d'une Passion *qui lui plait*." The expression found in earlier editions (e.g., Savile 1748: 2) is "*qui le flatte*." This difference may be read as 'that he likes' in contrast to 'that flatters him'. Both misrepresent Savile's intent when he wrote "it is possess'd by a darling Passion." Here, Formey did correctly identify a flaw in the original translation, which should have been *qu'il flatte*. Unfortunately, instead of just correcting the grammatical error, which put the cart before the horse, Formey changed the sense of *darling passion*, which is reflected by *qu'il flatte*, to the weak notion 'a passion that he likes'. This suggests that Formey did not consult the English text.

<sup>30</sup>At one point, Formey fell into an interesting error. An English passage reads "and the old Housekeeper shall make a better figure in the Family"; the parallel French passage is "Et une vieille *Ausquiper* sera meilleure figure dans la famille", with a footnote attached to the word *ausquiper* (Savile 1752: 64). Now, *ausquiper* is a nonce word; it is simply the English word *housekeeper* phonetically spelled for French ears. Nonetheless, Formey mistook *ausquiper* as an ordinary word, although it is not clear whether he took it for an English word or a French word. To the existing footnote he appended a claim that the word *ausquiper* is an equivalent of the German word *Hausgeber*. This again suggests that Formey did not read the English text.

<sup>31</sup>Savile's salutation "Dear Daughter" appears like a title on the introductory page (Savile 1741: 3). Savile's *Conseils* take "Dear Daughter" as a salutation rather than a section title and remove it to the introduction's first sentence: "J'éprouve tous les jours, *Ma chère Fille*, que nos pensées les plus agréables sont, malgré nous, mêlées d'inquiétude", leaving the introduction without a title (Savile 1748: 1). The titles of Savile's (1741) chapters are all formed by one to three nouns assisted by the conjunction *and*. The titles of the *Conseils* chapters are longer phrases, such as "Comment il faut qu'une Dame se conduise sur tout en Conversation" (Savile 1748: table des matieres). Set side by side with Savile's English titles, Formey's French titles are modelled on the *form* of their English companions, but take their *nouns and connective tissue* from the previous French titles (Savile 1752: Table). However, the introductory section did not have a French title (Savile 1748: 1). As the compositor at Haude & Spener set up this page (Savile 1752: 1), an unwanted gap would then appear in the French column opposite the salutation "Dear Daughter" in the English column. To fill this gap, "Ma chere Fille" was promoted from the first sentence in the French column to the title position. *Someone* recognized that it would be appropriate to set "Ma chere Fille" in the French column opposite "Dear Daughter" in the English column. *Someone* recognized that "Ma chere Fille" and "Dear Daughter" are functionally equivalent as salutations, even though they are not quite identical in meaning. This someone may have been Formey. However, this is the *only* change in Formey's edition that seems related to Savile's English original, so my money is on a balance-seeking compositor.

Instead, I found many suggestions that Formey ignored or did not understand the English text, to which he made no changes. The few changes that he made to the French text were editorial: orthography, punctuation, choice of word, choice of tense. These changes were made with respect to the previous French text; they are unrelated to the English text. This is to say, these changes could have been made without having in hand Savile's English text. Thus, Formey's *Conseils* does not support the supposition that Formey could translate from English to French.

### 3.2.3 ... regarding Locke's *Essay*

The one paragraph in *définition en logique* that was taken from Locke's *Essay* was lifted, with some editorial compression, from Coste's translation.<sup>32</sup>

Formey noted three books written by Locke in his *Conseils pour former une bibliothèque*. Coincidentally, but perhaps not, these three books had been translated into French *before* Formey made his recommendations. These books are Mazel's *Du gouvernement civil* (Locke 1691); Coste's *Education des enfants* (Locke 1695); and Coste's *Essai sur l'entendement* (Locke 1700b).<sup>33</sup> It might be thought that, addressed to a French audience, only works in French would be considered by Formey's *Conseils pour former une bibliothèque*; but recall that Formey had also recommended Chambers's *Cyclopædia*, then still an English work. One explanation of Formey's selection of these three books may be that he had read or knew their French incarnations but he did not recommend other works by Locke because Formey could not read English.<sup>34</sup>

It still might be so that Formey copied Coste's translation and included it among the pages received from Formey by the editors of the *Encyclopédie*. However, the passage from Locke is bracketed, fore and aft, by paragraphs from Condillac's *Essai*. This context suggests that Condillac, not Formey, supplied this passage.

Nonetheless, it does seem clear that Formey did not possess or at least did not deploy a competence to translate Chambers for his own article on *définition en logique*. Thus, translation by Formey of Chambers's article on *definition in logic* seems far-fetched.

---

<sup>32</sup>See Locke (1729: 336).

<sup>33</sup>It would have been quite impertinent of Formey to offer his own translation of Locke's *Essay*. Coste had worked closely with Locke, who was himself competent in French, to ensure that his translation (and Locke's text itself) suitably and adequately conveyed Locke's ideas (Soulard 2011).

<sup>34</sup>For another example of Formey copying a French translation of an English original rather than translating the English himself, see Coste (2009: 10).

### 3.3 Formey considered as compiler

#### 3.3.1 Formey's writings on definition

If Formey neither wrote nor translated paragraphs in *définition en logique*, could he have assembled that text? For this, I examine Formey's known writings on *definition*, compare them with the substance of this article, and assess whether Condillac's critique of Wolff is consistent with Formey's known views.

The manuscripts that Formey sold to the *Encyclopédie* editors were eventually returned to him at the Prussian Academy, but they disappeared, were looted, during the Russian occupation of Berlin in the aftermath of the Second World War; they have not been seen again (Marcu 1953: 299). Thus, I have been unable to consult Formey's papers to compare them with the printed *Encyclopédie* article. However, Formey did write on *definition* in other places; these include his *La belle Wolfienne* in 1741,<sup>35</sup> his *Abrégé de toutes les sciences, à l'usage des enfans* (1760);<sup>36</sup> his *Abrégé de toutes les sciences, à l'usage des adolescents* (1764);<sup>37</sup> and his

---

<sup>35</sup>“On ne doit faire entrer dans la Définition, que les Qualitez intrinseques du Sujet, c'est-à-dire, ses Qualitez essentielles & ses Attributs, qui fournissent les **marques**, auxquelles il est reconnoissable. Les **Possibilités** intrinseques peuvent néanmoins y entrer. . . . Cette propriété donc de se fondre, étant un résultat des Qualitez essentielles du Plomb, est une **Possibilité** intrinseque, qui convient à toute masse de ce Métal, & qui peut, par conséquent, entrer dans sa Définition. On conçoit alors la Liquebilité, comme une **Qualité** intrinseque du Plomb, & on la range parmi ses **marques**, comme servant à la distinguer des autres Corps. La **Possibilité** même des Modes & des Relations, étant une suite des Qualitez essentielles, en vertu desquelles un Sujet est susceptible de telles ou telles Déterminations extérieures ; cette **Possibilité**, dis-je peut encore être considérée comme intrinseque au Sujet, & admissible dans sa Définition. . . . Tout ce qui est, étant **possible**, puis qu'il est ; & cette **Possibilité** supposant une raison suffisante pourquoi il est ainsi, & non autrement ; la Définition, qui ne rend pas cette raison, n'est que nominale : celle, qui la rend, est réelle. . . . En effet, elles ne font qu'indiquer l'assemblage d'un certain nombre de Qualitez, dont l'une ne détermine pas nécessairement l'autre, & d'où il ne résulte rien, ni pour, ni contre, la **Possibilité** du Défini.” (Formey 1741: 1-120 et seqq).

<sup>36</sup>“D. Qu'entendez-vous par différence de la définition? R. J'entends par différence de la définition, ce qui fait différer essentiellement la chose définie de toute autre chose **existante**. D. Rendez sensible par un exemple tout ce que vous venez de dire. R. Le voici par la définition de l'homme ; quand je dis, l'homme est un Etre raisonnable, je donne sa définition par son genre & sa différence. Le mot Etre est le genre, parce qu'il confond l'existence de l'homme avec l'**existence** de tout autre être sans en spécifier aucun.” (Formey 1760: 3)

<sup>37</sup>“Cependant, dès qu'une idée est négative, ou destructive de l'autre, le terme est nul, il porte le caractère évident de sa réprobation dans la **contradiction** des deux idées simples auxquelles il répond. On peut inventer des exemples dans ce genre où la chose fautive aux yeux. Une figure rectiligne formée par deux lignes, *bilineum rectilineum*, est **impossible**, & le terme qui la désigne ne se rapporte à rien.” (Formey 1764: 295-296)

*Dictionnaire instructif* (1767).<sup>38</sup> These examples of Formey's definition of *definition* faithfully reflect Wolff's own definitions of *definition*. Wolff presented his definitions most concisely in the many editions of his *Mathematisches Lexicon* and of his *Philosophisches Lexicon*,<sup>39</sup> with admirable consistency, Wolff's definitions of *definition* in these *Lexicons* bind Wolff's notions of real or genitive definition to his essential ontological ideas of *contradiction*, *impossibility*, *possibility*, and *existence*.<sup>40</sup>

Without dissecting Formey's passages on *definition*, it is obvious that he deployed terms closely associated with Wolffian philosophy: *possibility*, *impossibility*, *contradiction*, *existence*, and, particularly with respect to *definition*, the noun *mark*.<sup>41</sup> The presence or absence of these terms in the *Encyclopédie* article on *définition en logique* is therefore significant, particularly given that the first para-

---

<sup>38</sup>“Énumération des principales marques, qui caractérisent un objet, et en donnent une idée distincte. C'est en cela proprement que consiste la définition nominale. Mais la définition réelle comprend l'idée de la **possibilité** de ce qu'on définit et l'explication intelligible de ce qui constitue cette **possibilité**.” (Formey 1767: s.v. définition)

<sup>39</sup>See, for example, his *Mathematisches Lexicon* (Wolff 1716); his *Philosophisches Lexicon*, (Wolff 1737); and his *Vollständiges Mathematisches Lexicon* (Wolff 1747). In his *Vollständiges Mathematisches Lexicon*, Wolff defines *definition* so: “heisset ein deutlicher Begriff dessen, wodurch eine Sache von allen andern unterschieden wird. Man hat aber einen deutlichen Begriff, wenn die Vorstellung der Sache so beschaffen, dass man sagen kan, aus was vor **Merckmahlen** man sie erkennt, oder wodurch man sie von andern unterscheidet. Wird nun die Sache von andern ihres gleichens durch ihre besondere Eigenschafften unterschieden, so heisset es eine Wort-Erklärung, *Definitio Nominalis*, dergleichen ist, wenn ich sage: Ein Uhrmerck sey eine Maschine, welche die Stunden anzeigt, so erkläre ich das Wort Uhrmerck. Wird aber die Art and Weise angegeben, wie eine Sache **entstehen** kan, und ich zeige z. E. aus was vor Rädern und andern nöthigen Zugehör in einer gewissen Absicht das Uhrmerck zusammen gesetzt, so ist dieses eine Sach-Erklärung, *Definitio Realis*. Diese letzte hat vor der erstern diesen Vorzug, dass man gleich daraus erkennen kan, ob die Sachen **möglich** sind oder nicht, daher sie ohne fernern Beweis zum Grund einer richtigen Demonstration gebraucht werden kan; da hingegen bey der erstern zu erweisen nöthig ist, ob sie **möglich** sey oder nicht. Uiber dieses lassen sich gar öftters auch aus den Sach-Erklärungen viele Dinge ohne all Umwege demonstriren, die man sonst mit Mühe erweisen muss. . . . Im übrigen ist sehr wohl zu behalten, dass zu einer Sach-Erklärung zweyerley erfordert werde, einmal, dass man wisse, was vor Dinge darzu gehören, wenn sie entstehen soll, und sodann auch, wie viel ein jedes von ihnen zur Sache beyträget.” (Wolff 1747: s.v. Erklärung)

<sup>40</sup>In turn, these concepts borrow heavily from Gottfried Wilhelm Leibniz, but that is well outside the scope of this chapter.

<sup>41</sup>The English “mark” or “characteristic” corresponds to Wolff's use of “Merckmahl” and “Kennmerck” when he wrote in German and “nota” when he wrote in Latin.

graphs of the article treat Wolff's notions of definition favorably.<sup>42</sup> However, these paragraphs were taken in translation from Chambers's article on *definition*, who was impressed by Wolff's notions of *definition* as Wolff expressed them in his *mathematical* texts rather than his *philosophical* works.<sup>43</sup> In particular, Chambers applauded Wolff's notion of genitive definition, which Wolff terms *real definition*. Chambers's applause for Wolff's notion of genitive definition is echoed in the *Encyclopédie* article on *définition en logique*. There the Wolffian term *mark* was taken from Chambers. However, non-mathematical Wolffian terms – *possibility*, *impossibility*, *contradiction*, *existence* – do not appear in Chambers's article and they appear in the corresponding *Encyclopédie* article only in Condillac's criticism of Wolff's philosophy.<sup>44</sup>

Thus, when Formey *did* write on *definition* in other places, he consistently used Wolffian terms to invoke certain Wolffian concepts. These terms are absent from the *Encyclopédie* article on *définition en logique*.

---

<sup>42a</sup>La *définition* de nom est celle qui explique le sens ou la signification propre d'un mot ; ou, comme le dit plus exactement M. *Wolf*, c'est l'énumération qu'on fait d'un certain nombre de marques ou de caracteres suffisans pour faire distinguer la chose qu'on définit, d'avec toute autre ; de sorte qu'il ne reste point de doute sur ce que c'est que la chose qu'on a voulu faire entendre & désigner par le nom. . . . M. *Wolf* dit que la *définition* de chose est une notion distincte qui explique la génération de cette chose, c'est-à-dire la maniere dont elle est faite ou dont elle se fait. . . . La notion que nous avons donnée de la *définition* de chose, d'après plusieurs philosophes, suffit pour faire connoître en quoi elle differe de la *définition* de nom. Mais quoique cette notion ait de son côté l'avantage de l'analogie, de la clarté & de la convenance, cependant comme elle n'est elle-même qu'une *définition* de nom, c'est-à-dire une *définition* du mot, c'est sous ce point de vûe principalement que nous devons la considérer, en la regardant comme une idée attachée arbitrairement à ce mot, & que l'auteur doit toujours y conserver attachée dans toute la suite de son ouvrage. Mais cette notion ne renferme point en effet le sens ou la signification ordinaire qu'on a coûtume de donner à ce mot, & qui est beaucoup moins juste & moins distincte ; & c'est à cette signification ordinaire que nous devons principalement avoir égard." (Diderot & d'Alembert 1751: s.v. *définition en logique*)

<sup>43a</sup>*Definition of the Name, or Nominal Definition*, is that which explains the sense, or signification appropriated to a word: or, as Wolfius more accurately considers it, an enumeration of certain **marks**, or characters, sufficient to distinguish the thing *defined* from any other thing; so to leave it out of doubt, what the subject is that is intended or denoted by the name. . . . Wolfius *defines* a *real Definition* to be a distinct notion explaining the genesis of a thing, that is, the manner wherein the thing is made, or done: . . . This notion of a *Real Definition* is very strict and just; and affords a sufficient distinction between a *real* and a *nominal* one. But, though it has the Advantages of analogy, distinctness, and conveniency on its side; yet being only itself a *Nominal Definition*, i.e., a *Definition* of the term *Real Definition*, we must consider it in that light, that is, as an idea fixed arbitrarily to that word, and which the author always denotes by that word in the course of his book. But, in effect, it is not the usual sense, or acceptation of the term; which is much less convenient and distinct. And it is to that usual acceptation we are here chiefly to have regard." (Chambers 1741: s.v. *definition*)

<sup>44</sup>For Chambers's use of Wolff's work in his *Cyclopædia* article on *definition*, see Bocast (2016).

### 3.3.2 Condillac's critique of Wolff's philosophy

And now I can lay out a central argument for *disattribution* in this case. Recall those Wolffian terms, *possibility*, *impossibility*, *contradiction*, and *existence*. Recall as well Condillac's words in paragraphs 16 through 18, which quite plainly *attack* the Wolffian notions signified by those terms. Suddenly, in the flow of the article, seemingly without provocation,<sup>45</sup> Condillac lashes out:

After having defined *impossible* as that which implies contradiction, *possible* as that which does not imply contradiction, and *being* as that which can exist, these philosophers have not known how to give any other definition of *existence* than the complement of possibility. But I ask whether this definition presents any actual idea, and whether we should not show it the ridicule that we have given to some of those of Aristotle.

If the *possible* is that which does not imply contradiction, *possibility* is the non-implication of contradiction. *Existence* is then the complement of the non-implication of contradiction. What a way to talk!<sup>46</sup>

<sup>45</sup>And (1) that Formey's article contained those Wolffian terms; (2) that these Wolffian terms offended Condillac's sensibilities, prompting him to include just these three paragraphs from his *Essai*; and (3) that Condillac simply tossed Formey's manuscript into his fireplace, for winters are cold in Paris.

<sup>46</sup>§. 14. Il est donc bien important de ne pas réaliser nos abstractions. Pour éviter cet inconvénient, je ne connois qu'un moyen, c'est de savoir développer l'origine & la génération de toutes nos notions abstraites. Mais ce moyen a été inconnu aux Philosophes, & c'est en vain qu'ils ont tâché d'y suppléer par des définitions. La cause de leur ignorance à cet égard, c'est le préjugé où ils ont toujours été qu'il falloit commencer par les idées générales; car lorsqu'on s'est défendu de commencer par les particulières, il n'est pas possible d'expliquer les plus abstraites qui en tirent leur origine. En voici un exemple.

Après avoir défini l'impossible par *ce qui implique contradiction*; le possible, par *ce qui ne l'implique pas*; & l'être, par ce qui peut exister: on n'a pas sù donner d'autre définition de l'existence, sinon, qu'elle est *le complément de la possibilité*.<sup>47</sup> Mais je demande si cette définition présente quelque idée, & si l'on ne seroit pas en droit de jeter sur elle le ridicule qu'on a donné à quelques-unes de celles d'Aristote.

Si le possible est *ce qui n'implique pas contradiction*, la possibilité est la *non-implication de contradiction*. L'existence est donc *le complément de la non-implication de contradiction*. Quel langage! En observant mieux l'ordre naturel des idées, on auroit vû que la notion de la possibilité ne se forme que d'après celle de l'existence.

Je pense qu'on n'adopte ces sortes de définitions, que parce que connoissant d'ailleurs la chose définie, on n'y regarde pas de si près. L'esprit qui est frappé de quelque clarté, la leur attribue, & ne s'aperçoit point qu'elles sont inintelligibles." (Condillac 1746: 234-236).

<sup>47</sup>Wolff gave his *definition* of existence in German in his *Deutsche Metaphysik* as "der Erfüllung der Möglichkeit" and in Latin in his *Philosophia Prima, sive Ontologia* as "per complimentum possibilitatis". See Wolff (1722: 8) and Wolff (1736: 143).

Can it be imagined that Formey – known across Europe as an ardent promulgator of Wolff’s philosophy – would attach his own name – as author, as compiler, as translator, as editor, or in any other way – to an article that, at its core, rudely dismisses ideas fundamental to Wolffian philosophy and to the six volumes of his own *La belle Wolffiene*? I cannot believe that he would do so.

### 3.4 Formey considered as editor

Considering Formey as editor is trivial: Formey in Berlin had no more to do with the manuscripts of his erstwhile “Philosophical Dictionary” once they had been put into the hands of the editors of the *Encyclopédie* in Paris (Moureau 1987: 145). I suggest that Diderot handed Formey’s notes on *définition* to Condillac soon after they arrived in Paris, perhaps in 1747 (Moureau 1987: 128-9). Later, d’Alembert took up Condillac’s draft of *définition en logique* when he wrote his own article on *définition en mathématiques*. Thus, given this spatiotemporal scenario, Formey could *not* have been the editor of the printed article, even could he have countenanced Condillac’s attack on Wolffian philosophy.

### 3.5 Conclusions

On these grounds, the attribution of the article on *définition en logique* to Jean-Henri-Samuel Formey appears to have no merit. The evidence shows that he did not *author* the article nor did he *translate* it from Chambers’s English. The evidence gives no grounds to believe that Formey *compiled* the contents of the article, and he *certainly* did not *edit* the article that was printed.

I suspect that Formey *did* provide an article on *définition* to the editors of the *Encyclopédie* as, without it, what could explain Condillac’s eruption but that this article disappeared in Paris long before it was lost in Berlin. Nonetheless, the article on *définition en logique* as *actually printed* can no longer sustain an attribution to Formey.

Oddly, the editors of the *Encyclopédie* *did* nevertheless place that attribution to Formey at the end of the printed article. This may have been part of a *quid-pro-quo* leading to Diderot’s invitation to the Berlin Academy in 1751 (Moureau 1987: 136). But perhaps there is another rationale for this attribution. Recall that neither Diderot nor d’Alembert had a high regard for Formey or for the Wolffian school of thought he professed (Kafker 1989: 141); perhaps together they found

---

Condillac translated Wolff’s Latin into French as “le complément de la possibilité”. However, other possible translations of Wolff’s definition of *existence* include “the fulfillment of possibility” and “the completion of possibility”; these alternate readings might be sensible.

this attribution a great joke – a great *contradiction* – *precisely* because the printed article so bluntly mocks Formey’s intellectual hero.

In the end, however, those who study the *Encyclopédie* should *disattribute* this article from Formey.

## References

- Arnauld, Antoine & Pierre Nicole. 1712. *La logique ou l’art de penser*. 6th. Lyon: Mathieu Libéral.
- Bocast, Alexander K. 2016. *Chambers on definition*. McLean: Berkeley Bridge Press.
- Buffier, Claude. 1724. *Traité des premières veritez, et de la source de nos jugemens, où L’on examine le sentiment des Philo[so]phes de ce temps, sur les premières notions des choses*. 1st edn. Paris: la veuve Maugé.
- Chambers, Ephraim. 1741. *Cyclopaedia: Or, an universal dictionary of arts and sciences*. 4th edn. London: D. Midwinter; J. Senex; R. Gosling; W. Innys; C. Rivington; A. Ward; J. & P. Knapton; E. Symon; S. Birt; D. Brown; T. Longman; R. Hett; C. Hitch; J. Shuckburgh; J. Pemberton; A. Millar; & the Executors of J. Darby.
- Condillac, Étienne Bonnot de. 1746. *Essai sur l’origine des connoissances humaines*. 1st edn. Amsterdam: Pierre Mortier.
- Condillac, Étienne Bonnot de. 1756. *An essay on the origin of human knowledge: Being a supplement to Mr. Locke’s essay on the human understanding*. Trans. by Thomas Nugent. London: J. Nourse.
- Coste, Alain. 2009. Air, the making of. *Recherches sur Diderot et sur l’Encyclopédie* 44. 9–14. DOI: 10.4000/rde.4551. <https://journals.openedition.org/rde/4551> (26 March, 2023).
- d’Alembert, Jean Le Rond. 1995. *Preliminary discourse to the encyclopedia of diderot*. Trans. by Richard N. Schwab. Chicago: University of Chicago Press.
- d’Alembert, Jean Le Rond. 2016. *The Encyclopedia of Diderot & d’Alembert Collaborative Translation Project*. Trans. by Alexander K Bocast. <http://hdl.handle.net/2027/spo.did2222.0003.206>.
- de Colonia, Dominique. 1710. *De arte rhetorica*. 2nd edn. Lugduni (Lyon): Antoni Molin.
- Diderot, Denis & Jean Le Rond d’Alembert (eds.). 1751. *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers*. Paris: Antoine-Claude Briaçon; Michel-Antoine David l’aîné; André-François le Breton; & Laurent Durand. <http://artflsrv02.uchicago.edu/cgi-bin/philologic/getobject.pl?c.14:2565.encycopedie0513.7832259> (15 July, 2014).

- Formey, Jean-Henri-Samuel. 1741. *La belle Wolfienne*. La Haye (Den Haag): Charles le Vier (1741); Jean Neaulme (1743-1753).
- Formey, Jean-Henri-Samuel. 1746. *Conseils pour former une bibliothèque peu nombreuse mais choisie*. 1st edn. Berlin: Haude & Spener.
- Formey, Jean-Henri-Samuel. 1750. *Conseils pour former une bibliothèque peu nombreuse mais choisie*. 2nd edn. Berlin: Haude & Spener.
- Formey, Jean-Henri-Samuel. 1755. *Conseils pour former une bibliothèque peu nombreuse, mais choisie*. 3rd edn. Berlin: Haude & Spener.
- Formey, Jean-Henri-Samuel. 1760. *Abrégé de toutes les sciences à l'usage des enfants*. nouvelle edition. Bruxelles (Brussels): Antoine du Pont.
- Formey, Jean-Henri-Samuel. 1764. *Abrégé de toutes les sciences à l'usage des adolescents, et de tous ceux qui veulent s'instruire*. 1st edn. Berlin: Joachim Pauli.
- Formey, Jean-Henri-Samuel. 1767. *Dictionnaire instructif*. Halle: Jean Justin Gebauer.
- Formey, Jean-Henri-Samuel. 2016. *Definition in logic*. Trans. by Alexander K Bocast. Ann Arbor. <http://hdl.handle.net/2027/spo.did2222.0003.206>.
- Gergonne, Joseph Diaz. 1818. Essai sur la théorie des définitions. *Annales de Mathématiques Pures et Appliquées* IX(1). 1–35.
- Grosse, Annelie. 2018. *The (self-) fashioning of an eighteenth-century Christian philosopher: Religion, science and morality in the writings and life of Jean Henri Samuel Formey (1711–1797)*. Florence: European University Institute. ().
- Hubac, Jean. 2007. Review of *Lettres de l'Angleterre à Jean Henri Samuel Formey à Berlin, de Jean Des Champs, David Durand, Matthieu Maty et d'autres correspondants (1737-1788)*. *Bulletin de la Société de l'Histoire du Protestantisme Français (1903-2015)* 153. 137–139.
- Kafker, Frank A. 1989. Notices sur les auteurs des dix-sept volumes de « discours » de l'Encyclopédie. *Recherches sur Diderot et sur l'Encyclopédie* 7(1). 125–150.
- Kafker, Frank A. & Serena L. Kafker. 1988. *The Encyclopedists as Individuals: A Biographical Dictionary of the Authors of the Encyclopédie* (SVEC 257). Oxford: Voltaire Foundation.
- Krauss, Werner. 1963. Ein Akademiesekretär vor 200 Jahren: Samuel Formey. In Werner Krauss (ed.), *Studien zur Deutschen und Französischen Aufklärung*, 53–62. Berlin: Rutten & Loening.
- Locke, John. 1691. *Du Gouvernement Civil, où l'on traite de l'Origine, des Fondemens, de la Nature, du Pouvoir, & des Fins des Sociétez Politiques*. Trans. by David Mazel. 1st edn. Amsterdam: Abraham Wolfgang.
- Locke, John. 1695. *De l'éducation des enfans*. Trans. by Pierre Coste. 1st edn. Amsterdam: Antoine Schelte.

- Locke, John. 1700a. *An essay concerning human understanding*. 4th edn. London: Awnsham & John Churchill, Samuel Manship.
- Locke, John. 1700b. *Essai philosophique concernant l'entendement humain*. Trans. by Pierre Coste. 1st edn. Amsterdam: Henri Schelte.
- Locke, John. 1729. *Essai philosophique concernant l'entendement humain*. Trans. by Pierre Coste. 2nd edn. Amsterdam: Pierre Mortier.
- Luzac, Elie. 1752. *Bibliothèque impartiale*. Leide (Leiden): Elie Luzac fils.
- Mallet, Edme-François & Jean François Marmontel. 2016. Definition in rhetoric. In *The Encyclopedia of Diderot & d'Alembert Collaborative Translation Project*. Michigan Publishing, University of Michigan Library.
- Marcu, Eva Dorothea. 1953. Un Encyclopédiste oublié: Formey. *Revue d'Histoire littéraire de la France* (3). 296–305.
- Marmontel, Jean François. 1782. Définition. In *Encyclopédie méthodique: Grammaire et littérature*, 572–575. Paris: Charles Joseph Panckoucke.
- Marmontel, Jean François. 1787. *Oeuvres complètes de M. Marmontel*. 2nd edn. Paris: Née de la Rochelle.
- Moureau, François. 1987. L'Encyclopédie d'après les correspondants de Formey. *Recherches sur Diderot et sur l'Encyclopédie*. 125–145. DOI: 10.3406/rde.1987.928.
- Savile, George. 1688. *The Lady's New-years Gift: Or, Advice to a Daughter*. 1st edn. London: Randal Taylor.
- Savile, George. 1692. *Etrennes, ou conseils d'un homme de qualité a sa fille*. Trans. by unknown. London: Jaques Partridge; & Matieu Gilliflower.
- Savile, George. 1698. *Conseils d'un homme de qualité à sa fille*. La Haye (Den Haag): Louis & Henri van Dole.
- Savile, George. 1699. *Conseils d'un homme de qualité a sa fille*. La Haye (Den Haag): Louis & Henri van Dole.
- Savile, George. 1700. *Miscellanies by the Right Noble Lord, the Late Lord Marquess of Halifax*. 1st edn. London: Matt. Gillyflower.
- Savile, George. 1704. *Miscellanies by the Late Lord Marquis of Halifax*. 2nd edn. London: W. Rogers; Benj. Tooke; & D. Midwinter & T. Leigh. <http://archive.org/details/miscbyla00hali> (23 July, 2022).
- Savile, George. 1717. *Miscellanies by the most Noble George Lord Saville, late Marquis and Earl of Halifax*. 3rd edn. London: B. Tooke; D. Midwinter; & J. Hooke.
- Savile, George. 1741. *The Lady's New-year's Gift: Or, Advice to a Daughter*. 12th edn. London: D. Midwinter; & A. Ward.
- Savile, George. 1748. *Conseils d'un homme de qualité à sa fille*. nouvelle. Lausanne: J. Christ. Sievert.

- Savile, George. 1752. *The Lady's New-Year's 'Gift: or, Advice to a Daughter // conseils d'un homme de qualité à sa fille*. Jean-Henri-Samuel Formey (ed.). Berlin: Haude & Spener.
- Savile, George. 1756. *The Lady's New-Year's Gift: Or, Advice to a Daughter*. 14th edn. London: R. & J. Dodsley; & A. Ward.
- Soulard, Delphine. 2011. L'œuvre des premiers traducteurs français de John Locke : Jean Le Clerc, Pierre Coste et David Mazel. *Dix-septieme siecle* 253(4). 739–762. DOI: 10.3917/dss.114.0739.
- Thomson, Ann. 1980. *Formey, Jean-Henri-Samuel*. <https://dictionnaire-journalistes.gazettes18e.fr/journaliste/310-jean-henri-formey> (4 March, 2022).
- Wolff, Christian. 1716. *Mathematisches Lexicon*. Leipzig: Johann Friedrich Gleditsch seel. Sohn.
- Wolff, Christian. 1722. *Vernünfftige Gedanken von Gott, der Welt und der Seele des Menschen, auch allen Dingen überhaupt, den Liebhabern der Wahrheit*. 2nd edn. Halle im Magdeburg: Rengerischen Buchhandlung.
- Wolff, Christian. 1736. *Philosophia Prima, sive Ontologia*. Frankfurt: Officina Libraria Rengeriana.
- Wolff, Christian. 1737. *Philosophisches Lexicon: Darinnen die Erklärungen aus Christian Wolffens teutschen Schrifften seines philosophischen Systematis*. Heinrich Adam Meißner (ed.). Bayreuth: Johann Gottlieb Vierling.
- Wolff, Christian. 1747. *Vollständiges Mathematisches Lexicon*. 2nd edn. Leipzig: Johann Friedrich Gleditsch seel. Sohn.

## Chapter 7

# Project Cleveland: Documenting the lexicographic output of 20th-century Slovenian immigrants in the US

Alenka Vrbinc

University of Ljubljana

Donna Farina

New Jersey City University

Marjeta Vrbinc

University of Ljubljana

At the beginning of the 20<sup>th</sup> century, an enormous immigration resulted in the establishment of multiple European ethnic groups in the United States. The Slovenian immigration to the US presents a compelling story. From 1905 until World War I, the greatest mass of Slovenians arrived and settled mostly in the midwestern and northeastern US. A smaller Slovenian wave post-World War II had refugees from communism and fascism.

The present endeavor analyzes the macrostructure, with some glimpses at microstructure, of six Slovenian–English, English–Slovenian dictionaries and reference works targeted at the Slovenian immigrant group. This is part of a project investigating both microstructure and macrostructure of a greater number of books, as well as the lives of their authors. A research trip to Cleveland, Ohio as well as Joliet and Lemont, Illinois allowed for the discovery of some dictionaries and an examination of newspaper archives, genealogical and other relevant materials. Certainly, the works found have played a role in the acculturation of Slovenians in the US. Slovenian immigrants are said to have “acquired English with impressive speed and facility” (Gobetz n.d.); the ongoing project seeks in part to understand how dictionaries and reference books may have supported this effort.



## **1 Introduction**

### **1.1 Project**

This paper does not present a finished product, but rather is part of an ongoing project that seeks to analyze the lexicographic output targeted toward Slovenian immigrants in the United States; it also seeks to understand the circumstances and context of this output. During Phase 1, the researchers identified, obtained, and began the analysis of numerous Slovenian–English, English–Slovenian dictionaries and reference works (phrasebooks, grammar books, school textbooks, etc.). These were published from the turn of the 19<sup>th</sup> century to the 1940s. This time frame corresponds to the most active periods of Slovenian immigration to the US.

Phase 2 of the project was a research trip in Summer 2021 to Cleveland, Ohio; Joliet, Illinois; and Lemont, Illinois. These are just a few of the sites where Slovenian immigration has had a lasting effect. Newspaper archives, genealogical, and other available materials were examined; these provided information about dictionaries and reference books, as well as about their authors, many of whom were prominent Slovenian-Americans. During this trip, meetings with descendants of Slovenian immigrants were invaluable. We were unable to meet with Dr. Edward Gobetz, an immigrant himself and a lifelong researcher of Slovenian immigration; he died in his 90s from covid, six months before our trip. Gobetz (2014) maintained that Slovenian immigrants learned English with “impressive speed and facility”; in part, this project seeks to understand how these dictionaries and reference books supported this effort.

Future work will examine the biographies of several of the Slovenian/American authors, among them one who was a Franciscan priest (Kasimir Zakrajšek), one who was a medical doctor (Fran Javh Kern), one who started a commune in California (Peter Jos Jeram), and one who had one of his dictionaries printed in the building that today is City Hall, New York (Viktor J. Kubelka).

In this paper we first present as background an overview of the context of Slovenian immigration to the US; while it is a given that dictionaries are cultural artifacts, products of their times, this background brings home the reality of the lexicographic work. Second, here we provide a general outline of the macrostructure of six of the reference works. While we incorporate some pages from the dictionaries and reference books, a presentation of the full microstructure is beyond the objectives of this paper.

## 1.2 Slovenia and Slovenian

The Republic of Slovenia today (Figure 1) is a country of more than two million people, located in central Europe. It was part of the former Yugoslavia and was one of the six republics within the Socialist Federal Republic of Yugoslavia. Slovenia declared independence in 1991. Modern Slovenian territory was part of the Austro-Hungarian Empire from 1867 to 1918 and it was also under the Habsburg Monarchy, various parts of modern Slovenia having been annexed from 1282 to 1335.

Slovenian, the most widely spoken language in the country, is classified genetically as a South Slavic language along with other languages spoken both within the former Yugoslavia and beyond it. Although Slovenian has a relatively small number of speakers, it nevertheless has a significant lexicographic tradition; this history, like that of many other traditions (cf. Béjoint 2016, Farina & Durman 2009, Fontenelle 2016) began with needs arising from contact with other languages and cultures. For Slovenian, the main cultural and linguistic contact was with the German language within Central Europe; this contact deeply impacts Slovenian culture and lexicography to this day.



Figure 1: Map of present-day Slovenia, created by the UN Office for the Coordination of Humanitarian Affairs (OCHA)

## **2 Part 1: History of Slovenian immigration to the US**

There is a rich history of Slovenian immigration around the world that has been well documented. One part of this is the immigration to North America, particularly to the US. Here, we present a chronological overview with an emphasis on the Slovenian presence in the US at the end of the 19<sup>th</sup> to the mid-20<sup>th</sup> century.

### **2.1 The beginnings**

Up to the 1850s, when Slovenia was still a part of the Hapsburg monarchy, mobility of people was limited, and thus outmigration was constrained. Nevertheless, there were clusters of Slovenians that set out for other lands before the mid-19<sup>th</sup> century. For example, following the Counter-Reformation in Europe, some Protestants from Slovenia settled in the US state of Georgia by way of Germany (Žugel 1987).

Beginning in 1680, a few Slovenian settlements were founded in the US, primarily associated with the activities of Catholic missionaries (Gobetž 2014). Sporadic emigration continued into the 18<sup>th</sup> century; there are Slovenian last names among the soldiers in the 1775 continental army of George Washington (Prisland 1950). Until 1782, Austria restricted the movement of its citizens, and those who managed to leave were mostly exiles and people avoiding conscription or deserting from the army.

Starting in 1830, the missionary work of the Slovenian bishop Frederic Baraga (1797–1868) among the Native Americans spurred Slovenian migration activity. His books were widely read in Central Europe and awakened a curiosity about America (Mazi-Leskovar 2003). In the preface to his Ojibwe dictionary, Baraga mentions Samuel Johnson (see Figures 2 and 3). Baraga is currently moving toward canonization in the Catholic Church.

### **2.2 Mass migration and mass immigration**

Slovenian emigration began to pick up as part of the mass global movement of peoples. The Industrial Revolution of the late 18<sup>th</sup> and early 19<sup>th</sup> centuries unleashed internal migration, a flow of rural migrants to European towns and cities. The internal migration of Slovenians was exacerbated by the abolition of serfdom, resulting in the disintegration of farms and a dire economic situation for many (Žugel 1987). It is not surprising that the less developed Austrian south (see Figure 4, area labeled as Carniola), including territories that are part of modern Slovenia, saw substantial out-migration (Drnovšek 2001).

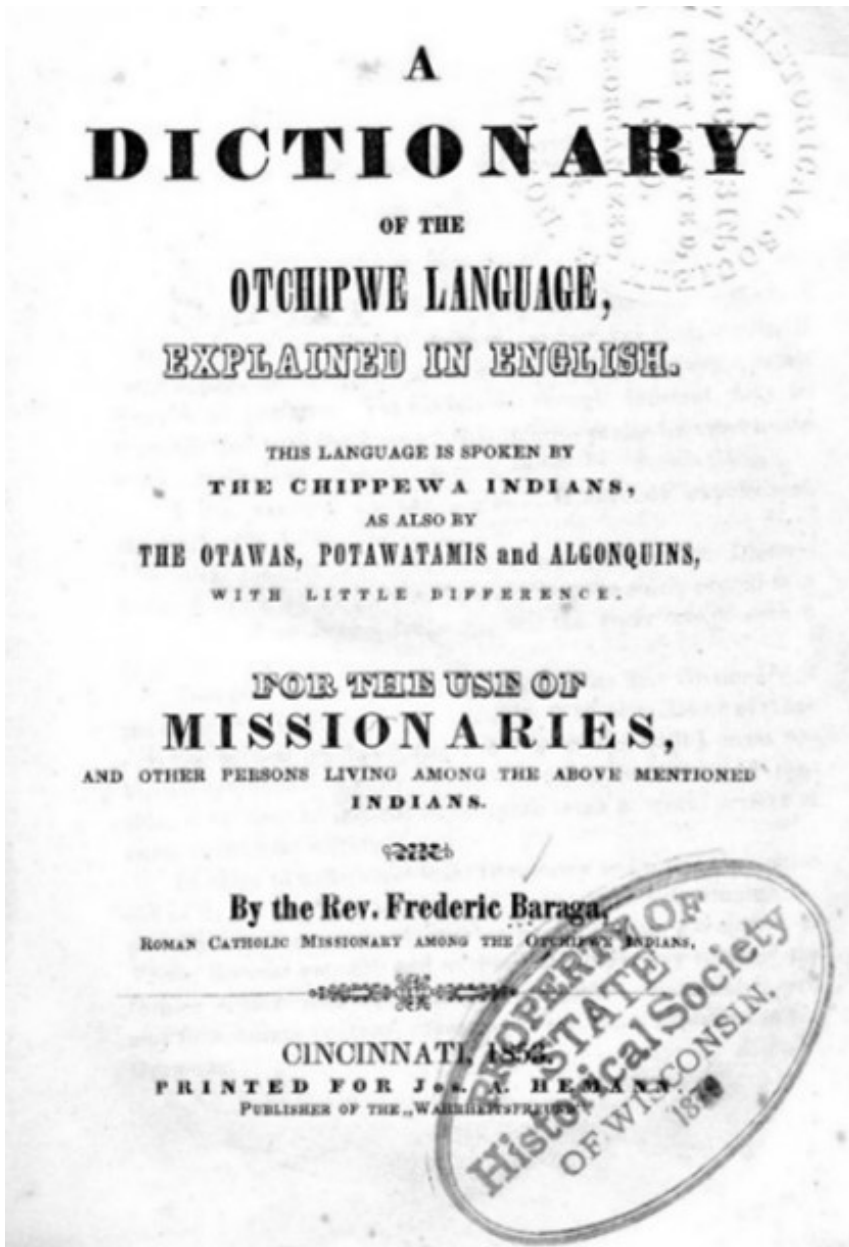


Figure 2: Title page of Baraga's (1853) Ojibwe dictionary

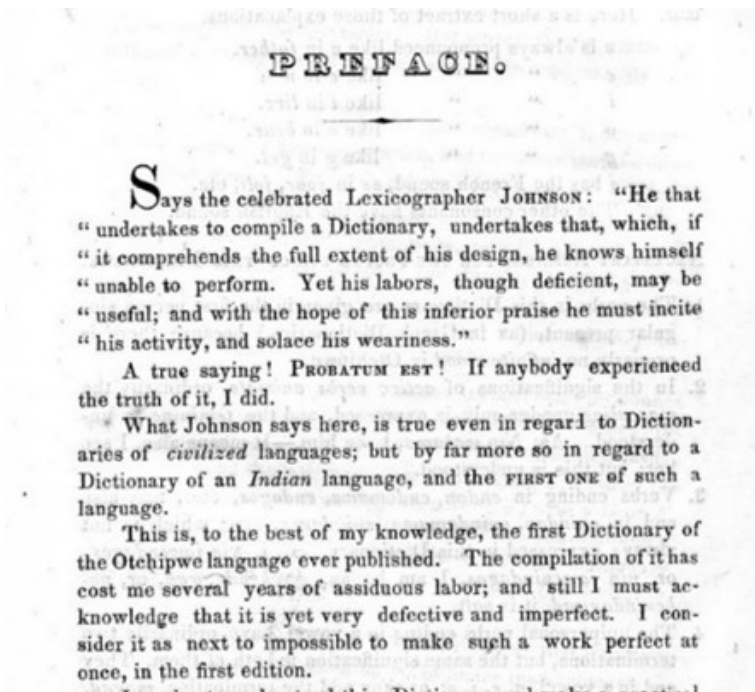


Figure 3: Preface of Baraga's (1853) Ojibwe dictionary

The increasing commercial viability of steam power after 1764 led to the development of rail and sea transport, making the whole world more accessible. The population of Europe increased to 300 million in 1900, from 128 million in 1750. This increase is especially significant in light of the huge emigration; from 1820 to World War I, 55 million Europeans left, about 39 million of them arriving in North America (Ferrie & Halton 2013).

In the 20<sup>th</sup> century, the face of European emigration to the US began to change. As noted by Gerber (2011):

Between 1871 and 1900, 11.7 million immigrants arrived [in the US]; between 1901 and 1920 alone 14.5 million did. The points of origin were changing dramatically; while in the nineteenth century, western and northern Europeans predominated, now southern, central, and eastern Europeans did. (p35)

The years 1905 to 1913 saw the largest wave of Slovenian immigrants arrive in North America (Gobetž 2014). According to US census data, over 183,000 people said that their native language was Slovenian in 1910; ten years later almost



Figure 4: The ethnic groups of Austria-Hungary in 1910 according to *Distribution of races in Austria-Hungary* by William R. Shepherd, 1911. This work is in the public domain.

209,000 people made this claim (Drnovšek 1994). It is estimated that prior to World War I, about one-third of the entire Slovenian population left their home region and went abroad. The Slovenian arrivals in North America were by and large peasants, but very few were able to continue farming. Most began working in coal mines or quarries; on railroad tracks, as masons, longshoremen, and lumbermen or in iron- and steel-working (Žugel 1987, Priland 1950). Among the non-economic reasons for immigration were the desire to evade military service or the avoidance of criminal prosecution (Drnovšek 2001).

A list of the early Slovenian settlements in the US (see Table 1) indicates that the ethnic group's presence was significant and widespread.

The Midwestern US (Pennsylvania, Ohio, Illinois, Minnesota, Wisconsin, and Michigan) was home to the most Slovenians; there were also settlements in Colorado and California (Gobetz 2014, Klemenčič 1986, Žugel 1987.)

Table 1: Early Slovenian Settlements in the US (Gobetz 2014, Klemenčič 1986, Žugel 1987)

---

---

1865	Brockway, Minnesota
1868	Omaha, Nebraska
1873	Joliet, Illinois
1878	New York City
1881	Cleveland, Ohio
By 1914	Slovenian settlements in 34 US states and three Canadian provinces

---

---

### 2.3 Backlash

Americans did not look favorably on the immigrants of the pre-World War I period; many feared that they would not assimilate as the northern European and Protestant immigrants before them had done. Their religion, physical appearance (called their “race” at that time), clothing, and poverty fostered suspicion (Gerber 2011). Slovenians were subjected to pejorative names, including “Grainers” or “Grenish” from “Krainers,” people from the province of Krain, Kranjska, or Carniola (Gobetz 2014). American concerns about this new wave of immigrants resulted in the Immigration Act of 1891 (Hoyt 2017) as well as the restrictive quota immigration laws of the 1920s.

After World War II, a more sympathetic attitude on the part of the American public allowed exceptions to or expansions in the strict quotas of earlier immigration laws. Many Slovenians who entered the US during this postwar period were displaced persons, refugees either from fascism or from communism (the Tito regime). Approximately 100,000 Slovenians left communist Yugoslavia, with the US among their top destinations (Baker et al. 2023). These new immigrants layered onto the existing Slovenian–American communities, not without religious, political, and ideological conflicts among them. One author describes this situation as a “schism” and indicates that there were dueling organizations, separation of living areas within cities, and even avoidance of intermarriage between groups (Milharčič-Hladnik 2007); another writer refers to the situation as a “cleavage” (Susel 2022).

After the mid-20<sup>th</sup> century, there was little migration out of Slovenia, to the US or anywhere else (Baker et al. 2023). In the US today, the era of the close-knit and active Slovenian communities has passed, but the descendants of immigrants whom we met during this project have been motivated and persistent in keeping the knowledge base of their culture alive, through the maintenance of continuing

organizations and through connections to organizations in Slovenia. The enormous archive of Edward Gobetz will be transported to Slovenia for preservation.

### 3 Part 2: Macrostructure

To date we have identified 11 different Slovenian–English, English–Slovenian reference books and dictionaries targeted at the above-described Slovenian–American immigration. Most of the books were obtained as intact, full copies; in some cases, only excerpts are available. The analysis of the books is ongoing. Here we will discuss six of these books, intended for the Slovenian immigrants who arrived up to World War I. The discussion is focused mostly on the macrostructure, with some glimpses of microstructure. The patterns in macrostructure emerging from the six books examined lead to the expectation that subsequent analyses of the remaining books will not result in many (or any) new macrostructure discoveries.

#### 3.1 Jeram’s *Slovensko-angleška slovnica* [Slovenian–English Grammar], 1895.

Peter Josip Jeram (see Figure 5) was born in 1857<sup>1</sup> in the village of Smoleva; he died in Covelo, California in 1897 (Cesar-Nedzbala 1990, Sedmak 2021a). Jeram attended a seminary (St. Francis de Sales, today in the Milwaukee Archdiocese, Wisconsin); he became a priest in 1880. He first practiced as a priest in Minnesota and South Dakota, writing the *Slovensko-angleška slovnica* [Slovenian–English grammar] while in South Dakota, from about 1889 to 1892 (Cesar-Nedzbala 1990, Sedmak 2021a,b,c).

Jeram’s (1895) *Slovensko-angleška slovnica* of 1895 is a reference book rather than a dictionary; it contains some segments with similarities to dictionaries. It is the oldest of the books found for the Slovenian–American immigration. Its place of publication, Tower, Minnesota, is part of a larger area where early Slovenian immigrants settled. Brockway, Minnesota, the first Slovenian settlement (1865) is about a four-hour drive south from Tower (see Table 1).

The title page of Jeram’s grammar is in Slovenian only with no English (Figure 6), indicative of this early stage in the Slovenian immigration when the native language was still primary for most users. The grammar has all explanations in Slovenian. There is not a full foreword but just a short quote from a well-known Slovenian writer, Janez Trdina.

---

<sup>1</sup>Another source, Grafenauer (2013), gives a birthdate of around 1865, but this is unlikely.



Figure 5: Peter Josip Jeram (Archives of the Archdiocese of St. Paul and Minneapolis).

Pages three to 18 (16 pp.) contain a very detailed description of English pronunciation. Figure 7 shows the treatment of pronunciation irregularities. For example, A) §43.1 discusses the pronunciation of what is called the English vowel *a*, provides examples of English words containing *a* (for example, *have*, *scarce*, *scarcely*, etc. with their translations into Slovenian), and compares the English pronunciation with words in Slovenian containing *e* or *a*. Overall, the explanations here and elsewhere in the pronunciation section were confusing to three trained linguists and most likely would have been incomprehensible to uneducated Slovenian–American users.

Part One of the grammar, pages 17–81 (65 pp.), consists of translation exercises; Exercise II is shown in Figure 8. This exercise begins with some general information, presumably that would enable a reader to undertake a translation. The information provided is a list of Slovenian words and their English equivalents related to *Družina* The Family (as the title of the exercise indicates); the verb *imeti* ‘(to) have’ is also presented. The actual exercise that follows consists of English sentences related to the theme of family that use the English words previously presented. The sentences are related to each other by theme but could not be called a paragraph. The learner is apparently expected to translate the sentences into Slovenian, although no instructions are given.

The backdrop for this exercise format is that translation exercises for language-learning purposes date at least to the bilingual manuals of Ancient Greek and Latin used in third-century Rome; in 18<sup>th</sup>- and 19<sup>th</sup>-century Europe, the use of grammar–translation was widespread for language teaching (Hernández

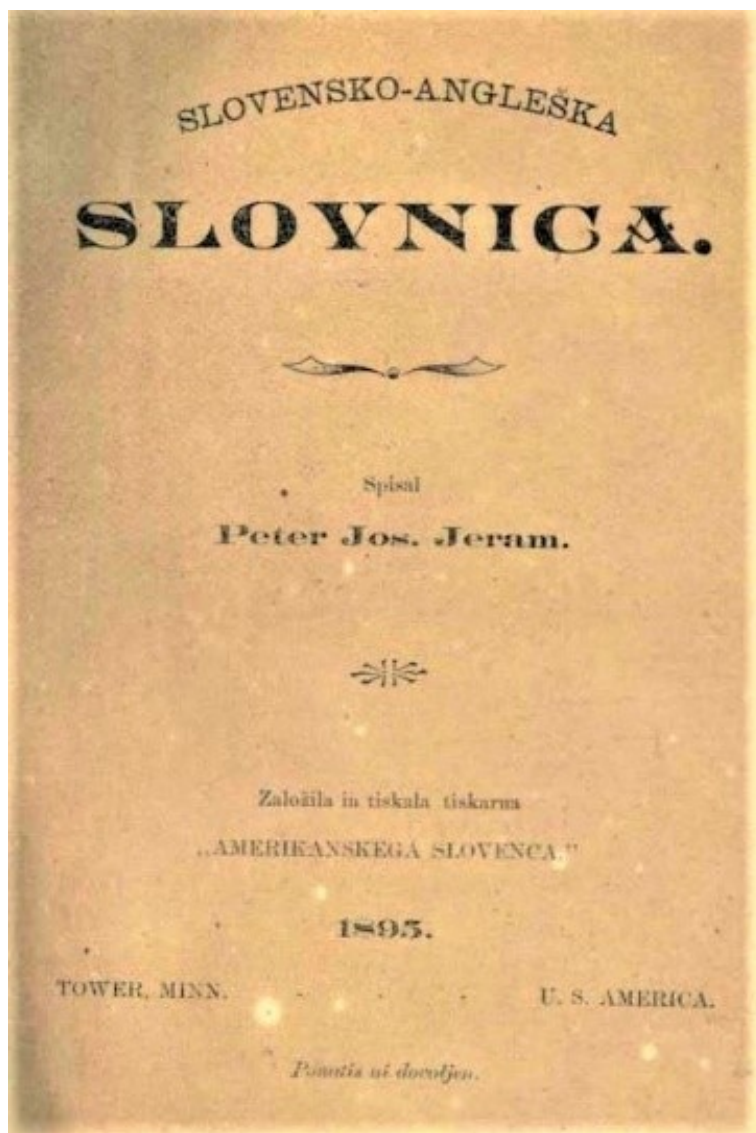


Figure 6: Title page from Jeram 1895.

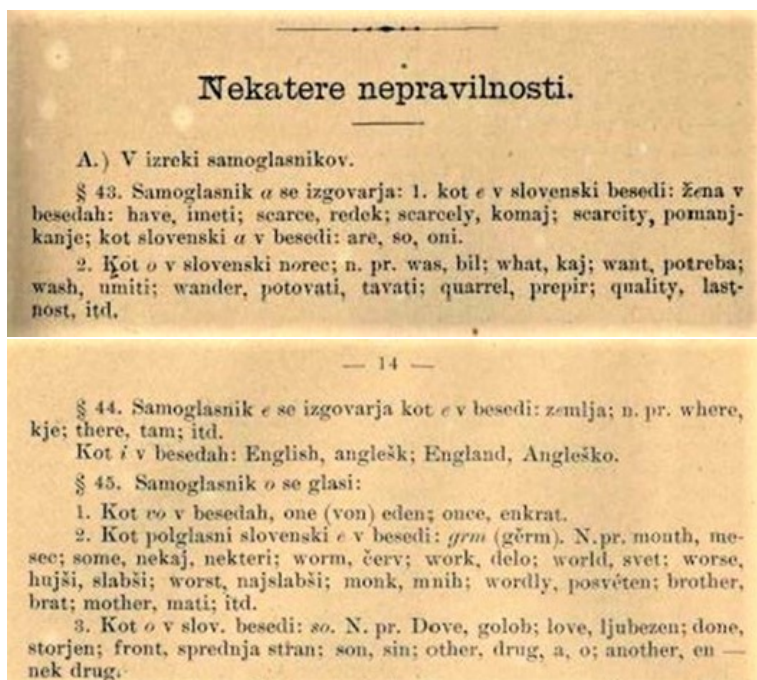


Figure 7: Irregularities of English pronunciation (Jeram 1895: 13–14).

Jaramillo 2019; see also Zhou & Niu 2015. It would not be surprising if Jeram based his book on the grammar–translation methods of his time. However, one reviewer of the *Slovensko-angleška slovnica* (Anonymous 1895) linked Jeram’s approach to that of Heinrich Gottfried Ollendorff (also known as Henri Godefroy Ollendorf), whose own method was more conversationally based.

Underneath the English sentences to be translated (19), Jeram presents a rule (*Pravilo*) explaining that English does not decline nouns in the nominative, accusative, singular, or plural. This is followed by a note (in Slovenian): “Look at every word carefully, how it is spelled, if you want to spell English words correctly and if you want to speak nicely and correctly.” The next part of Exercise II contains Slovenian sentences related to the theme of family; apparently the reader is supposed to translate them – though again, no instructions are given.

The other exercises of Part One follow essentially the same pattern as Exercise II; a few of the themes covered by these exercises are listed in Table 2. Within each theme, grammar explanations are also provided; for example, modal verbs, tenses, and plural formation, to name a few. Sample correspondence (letters) related to the themes is also given.

**II. Naloga.**

**Družina. The Family.**

Oče, father,	otrok (dete) child,	naka, ne, no; noben, a, o, no.
mati, mother,	stric, uncle,	da (ja), yes,
brat, brother,	teta, aunt,	ne, not,
sestra, sister	bratranec, sestrična, cousin,	tako, so,
sin, son,	velik, a, o, tall,	in, and,
hči, daughter,	majhen, a, o, small,	pa, ali, but,
deček, boy,	slaboten, a, o, weak,	kakor, kot, as,
deklica, girl,	močan, a, o, strong,	je, is; on je, he is (is — iz).

**Imeti, (to) have.**

I have, imam,	Have I, ali imam (jaz)?
thou hast ) imaš,	hast thou, ) ali imaš (ti)?
you have {	have you, {
he has, (on) ima,	has he, ali ima (on)?
she has, (ona) ima,	has she, ali ima (ona)?
it has, (ono) ima,	has it, ali ima (ono)? n. pr. dete;
we have, imamo,	have we, ali imamo?
you have, imate,	have you, ali imate?
they have, imajo.	have they, ali imajo?

*V.A.J.A.*

I have a father and a mother, an uncle and an aunt. You have a brother, but no sister. My father has a brother and a sister. His brother is my uncle, and his sister is my aunt. Our uncle has a son and a daughter.

— 19 —

This boy is tall, but that girl is small. Is that boy your brother? No, he is my cousin. This child is weak. My brother is not so strong as my cousin.

*Pravilo.* Pri imenih (rečovnikih) je tožilnik vedno enak imenovalniku v vseh treh spolih, v edinem in množnem številu.

*Opomba.* Dobro poglej vsako besedo, kako se piše, če želiš pravilno angleški pisati in lično pa pravilno govoriti znati.

*V.A.J.A.*

Ali imaš očeta in mater? Da, jaz imam očeta, matere pa nimam. Moj brat ni tako velik, kot moj oče. Imamo strijca in teto. Moj stric ima sina in hčer. Njegov sin je moj bratranec. Moj brat ni tako močan in tako velik, kot moja sestra in moja sestrična. Moja mati so (is) slabotni. Ta otrok je slaboten in majhen. Ali je ova dekla tvoja sestra? Ne, ona je moja sestrična.

**III. Naloga.**

Figure 8: Part One, Exercise II, thematically-based translation exercise (Jeram 1895: 18–19).

Table 2: Part One exercises, a sampling of themes covered (Jeram 1895).

Meals	The Garden	Minerals
Division of Time	The Flower-Garden	Geographical Names
The School	Vegetable-Garden	The Universe
The Church	The Orchard	

Part Two of the book, pages 82–173 (92 pp.) is called Elementary Grammar. Unlike Part One, this section is not organized by conceptual themes but by points of grammar. It is interesting that some titles are in English, others are in Slovenian, and some are in both languages.

As an example, we can examine Exercise XLVII (p47), shown in Figure 9. The grammar point presented is nouns with irregular plurals; the section begins with a list of such English nouns, usually separated from their Slovenian equivalent by a comma. English *Man* is an exception in that two equivalents are given, separated from each other by a semi-colon due to semantic differences. The first equivalent has the meaning ‘male’ whereas the second one means ‘human being, person’. Similar to what we saw in Part One, the text for translation is not a paragraph, but just a string of sentences, usually with no apparent connection between them. For example, while it is possible to loosely connect: “Wine has drowned more men than [sic] water” with “The more women look into their glasses, the less they look to their homes,” these two sentences are not conceptually linked to: “Geese have very valuable feathers” or “The Dutch are renowned for their cleanliness.” After the sentences provided for translation, we see a short glossary designed to help the user carry out the translation – a list of English words and their Slovenian equivalents. Following the list there is a description of the grammatical topic, irregular plurals.

While each section in Jeram’s Part Two usually begins immediately with the sentences to be translated, sometimes the ordering of elements can be altered. A section can sometimes begin with vocabulary or else with the grammatical point being described, with the translation exercises appearing afterwards. Some of the section headings in Jeram (1895) reflect terms for grammatical categories that are not in line with modern usage. To take one example, while today the English term *indefinite pronoun* is used, Jeram has “Undefined Pronouns”; the Slovenian term he used, *nedoločni zaimek*, is the same as in modern terminology.

Following Part Two, there are two pages on weights and measures and two pages of errata. The existence of the errata pages might indicate that there was a

**XLVII. Naloga.**

**Nepravilna množna števila. Irregular Plurals.**

1. Man, mož; človek,	men, možje, ljudje.	goose, gos,	geese, gosi.
woman, žena,	women, žene.	mouse, miš,	mice, miši.
child, otrok,	children, otroci.	louse, uš,	lice, uši.
foot, noga,	feet, noge.	ox, vol,	oxen, voli.
tooth, zob,	teeth, zobje.	sheep, ovca,	sheep, ovce.

2. Englishman, Anglež. Englishmen, Angležje, statesman, državnik.  
Dutchman, Holanec, Dutchmen, Holanci, statesmen, državniki.

— 88 —

Irishman, Irec,	Irishmen, Irci,	workman, delavec.
Scotchman, Skot,	Scotchmen, Skotje,	workmen, delavci.

**EXERCISE.**

Selfish men meet with no assistance in time of need. Wine has drowned more men than water. No man should be both accuser and judge. Frailty! thy name is woman. The more women look in to their glasses, the less they look to their homes. Geese have very valuable feathers. Small feet are considered pretty. Teeth are often spoiled by taking food to hot or too cold. The cat is the greatest enemy of mice. In some countries people use oxen to draw the plough. Sheep are amongst the most useful domestic animals. The French and the English have often been at war with each other. We met Frenchmen and Irishmen on our journey. The Dutch are renowned for their cleanliness. Our workmen receive their wages every Saturday.

To drown, vtoniti, vtopiti.	hollow, votel.
accuser, tožnik.	to ache, boleti.
judge, sodnik.	cold, prehlajenje.
frailty, slabotnost, omahljivost,	to catch cold, prehladiti se,
glass, ogledalo.	to draw out, izvleči.
factory, tovarna.	the French, Francozi,
to employ, rabiti koga,	the English, Angleži,
prone, nagnjen.	the Irish, Irci.
to avenge, maščevati.	the Dutch, Holanci.
injury, krivica.	the Scotch, Škoti.
stupid, neroden, trd.	to be at war, vojskovati se.
to indicate, naznanjati.	renowned, sloveč, znamenit.
silly, aboten, neumen.	cleanliness, svažnost,
watery, voden, vodnat.	passenger, potovalec.
to injure, škodovati, po-	wages, plača, zaslužek.

**Pravila.**

1. Samostalniki, zloženi iz dveh besed, katerih zadnja je: *man*, spreminjajo to besedo v množ. štev. v: *men*. N. pr. English-man, Englishmen, itd.

2. Če pa samostalniki niso zloženi, temveč se le tako z *man* končujejo, je njihovo množ. štev. pravilno. N. pr. German, Nemec; Germans, Nemci; Roman, Rimljan; Romans, Rimci, itd.

Opomba: Besede: French, Francozi, the English, Angleži, the Irish, Irci, itd. zaznamujejo celi narod; Frenchman, itd. zaznamuje le eno osebo in Frenchmen, več posameznih oseb izmed naroda.

Dialect, narečje.	understood, razumel — en.
to help, pomagati,	wharf, pristanišče.
to understand, razumeti,	native tongue, materni jezik,

Figure 9: Part Two, Exercise XLVII, translation exercise linked to grammar (Jeram 1895: 87–88).

printing before the book we have seen; however, we have found no evidence of a previous printing. It is also possible that during the process of preparation of the final book, Jeram or someone else noticed errors which there was no time to correct. At the end of the errata section, Jeram added an interesting final sentence (Jeram 1895: 177): *Druge manjše popravke si učenec lahko sam popravi* ('The learner himself can correct other minor mistakes'). The final two pages are the table of contents.

Jeram died in 1897 at the age of 40, two years after his grammar was published. While he apparently envisaged a dictionary as a companion to the grammar dictionary itself (Jeram 1895: 16), none ever materialized.

### **3.2 The dictionaries of Viktor Kubelka**

Viktor J. Kubelka was a New York City-based immigrant. While he was born in Ljubljana in 1879, his last name as well as genealogical sources indicate that he was partially of ethnic Czech origin. He died in Sullivan, New York in 1961.

#### **3.2.1 Kubelka's Slovenian–English pocket dictionary to facilitate the study of both languages, Slovensko-angleški žepni rečnik v olajšavo naučenja obeh jezikov, 1904.**

This apparently is the first reference book Kubelka published. The first page is a bilingual Slovenian and English title page, listing Kubelka as editor. Next comes a full-page advertisement (in Slovenian) for Frank Zotti & Co., headquartered in New York City, which offered banking services (including the possibility for immigrants to send money to the homeland), sold steamer tickets, and provided other services likely to be needed by immigrants. The next page states only (in English): "POCKET DICTIONARY"; the following page states (in English): "Copyright by V. J. KUBELKA. 1904 ALL RIGHTS RESERVED." Finally, we have two additional pages that correspond to the normal formula for title pages. The first is in English: *Slovenian–English Pocket Dictionary to Facilitate the Study of Both Languages*, and the second is in Slovenian: *Slovensko-angleški žepni rečnik v olajšavo naučenja obeh jezikov*. Unlike Jeram (1895), where the sole title page is in Slovenian, nine years later in Kubelka's publication, the English title page has first position, reflecting possibly a transition toward English already beginning as the Slovenian immigration to the US continues.

A preface of one page in English followed by a one-page preface in Slovenian come after the two title pages. The next component is one page with the English alphabet and the pronunciation of each letter, followed on the same page by

the Slovenian alphabet. The front matter also contains an explanation of the pronunciation of English vowels, diphthongs, and consonants.

The central part of this reference book is the dictionary itself (p24–122) (99 pp), in which Slovenian lemmas in column 1 are followed by English equivalents in column 2 and the pronunciation of the English equivalent in column 3. The dictionary’s outer access structure consists of two running heads per page. As seen in Figure 10, the top left corner of the left-hand page contains the first lemma of the left-hand page and the top right corner of the left-hand page contains the last lemma of the left-hand page.

The dictionary itself is followed by a section called “Easy dialogus,” written in English (122–131; see Figure 11). It is arranged according to the following topics: Salutations; A Meeting; To make Inquiry; The Necessaries [sic] of Life; and Looking for work.

The next two chapters provide additional information about the US. For example, the names of US states are given along with their abbreviation and pronunciation – with the title in Slovenian rather than in English (see Figure 12).

Other information includes money in the US, with the name of a coin or bill, the pronunciation, how to write it, and the value in crowns. Under the topic “Names,” Slovenian first names are followed by their equivalents in English; the pronunciation of the English names is provided (see Figure 12). On the next page of the book is the table of contents (see Figure 13).

After the table of contents there are two pages with Slovenian-language advertisements. One page has an advertisement that contains the name and address of the official representative of different shipping companies as well as information for those wishing to send money back to their homeland. On the next page is an ad for the same Frank Zotti & Co. advertised in the front matter, who provides tickets for “the best, the quickest, the safest and the cheapest steamers in the world.” This advertisement lists ports in Europe, the US, South Africa, Asia, and Australia; also included is the purchase of train tickets.

### 3.2.2 Kubelka’s *Slovensko-angleški razgovori*, Slovenian–English Interpreter, 1912a.

Kubelka (1912b) is a small phrasebook rather than a dictionary proper. The entire book has 136 pages, as compared to Kubelka (1912a) (below) which has 431 pages including the ads. The phrasebook starts with a title page in Slovenian and English, a copyright page, and a table of contents in Slovenian. The copyright page bears the notation: “Press of Stettiner Bros., New York.” These Austrian printers and brothers worked out of the building that is now New York City Hall.

REČNIK. DICTIONARY. (dikšjoneri).		A.	
Slovensko	Angleško	Izgovorjanje	
A	and	end	
a	but	bet	
abeceda	alphabet	afabhet	
abota	foolishness	fušines	
"	foolish action	fušik akšja	
abotnica	foolish women	fušis vurne	
abotnik	fool	fuš	
ajda	black-wheat	buk vit	
ako	if	if	
ako ravno	though	dzo	
"	although	aldzo	
ali	but	bet	
"	yet	jet	
"	if	if	
"	whether	vedker	
angelj	angel	endžel	

Čujem	— 36 —	Dar
Slovensko	Angleško	Izgovorjanje
čujem	I am awake	aj em evčk
čutiti	feel	fil
čuvaj	watchman	vačmen
čuvati se	beware of	biver of
čvrst (močan)	strong	strong

D.		
Da	yes	jes
da (dasiravno)	wether	veder
dajatelj	giver	giver
dajati	give	gif
daleč	far	far
daljava	distance	distens
dajje	farther	fardzer
daljen	wide	vajd
dati	give	giv
dan	day	dej
danes	to day	tu dej
današnji	to days	tu dejs
danica	morning star	morning star
dar	present,	present,
"	gift	gift

Figure 10: Two pages from the dictionary section (Kubelka 1904: 24,36).

— 127 —

*Življenju*     *The Necesseries* (*dze neseries*  
*potrebne stvari.*     *of Life.*     *of laif.*)

Ali (si) ste lačni? — Are you hungry? — *ar ju hongri.*

Da, jaz sem jako lačen — Yes I am very hungry — *jes aj em veri hongri.*

Ne jaz nisem lačen, ampak žejen sem — No, I am not hungry but, I am thirsty — *no aj em not hongri, bot aj em dersti.*

Kaj želite piti? — What will you drink? — *vat vil ju drink.*

Čašo (kozarec) pive — A glass of beer — *e gles of bir.*

Ali si (ste) truden? — Are you tired? — *ar ju tajrd.*

Da, jaz sem otrujen — Yes, I am a little fatigued — *jes aj em e litle fatižd.*

Si se (ste) prehladil? — Are you cold? — *ar ju kold.*

Ne, zaspan sem — No I, am sleepy — *no aj em slipi.*

Pojdi v posteljo — Go to bed — *go tu bed.*

Da, takoj hočem iti v posteljo — Yes I shall go to bed immediately — *jes aj šel go tu bed ime djeteli.*

Figure 11: “Easy dialogus” (Kubelka 1904: 127).

Alabama	— 132 —	Missouri
<b>IMENA ZJEDINJENIH DRŽAV.</b>		
<i>Ime</i>	<i>Skrajšano</i>	<i>Izgovarjanje</i>
Alabama	Ala.	<i>alabama</i>
Arkansas	Ark.	<i>arkenscs</i>
Arizona	Ariz.	<i>arizona</i>
California	Cal.	<i>kalifornia</i>
Colorado	Colo.	<i>kolorado</i>
Connecticut	Conn.	<i>konektikut</i>
Delaware	Del.	<i>delever</i>
District of Columbia	D. C.	<i>distrikt of kolumbia</i>
Florida	Fla.	<i>florida</i>
Georgia	Ga.	<i>džordžia</i>
Idaho	Io.	<i>ajdaho</i>
Illinois	Ill.	<i>ilinojs</i>
Indiana	Ind.	<i>indiana</i>
Iowa	Ia.	<i>ajova</i>
Kansas	Kas.	<i>kansas</i>
Kentucky	Ky.	<i>kentuki</i>
Lousiana	La.	<i>lujziana</i>
Maine	Me.	<i>mejn</i>
Maryland	Md.	<i>merilend</i>
Massachusetts	Mass.	<i>masačuzec</i>
Michigan	Mich.	<i>mišigen</i>
Minnesota	Minn.	<i>minesota</i>
Missouri	Mo.	<i>mizuri</i>

Figure 12: Names of US states (Kubelka 1904: 132).

— 135 —

---

IMENA.	NAMES.	(nems.)
~~~~~		
<i>Slovensko</i>	<i>Angleško</i>	<i>Izgovarjanje</i>
Adolf	Adolphus	<i>edolf</i>
Andrej	Andrew	<i>endru</i>
Anna	Annie	<i>eni</i>
Anton	Anthony	<i>entoni</i>
France	Frank	<i>frenk</i>
Johana	Johnny	<i>džoni</i>
Jakob	Jacob	<i>džekob</i>
„	Jake	<i>džek</i>
„	James	<i>džems</i>
„	Jimmy	<i>džimi</i>
Joan	John	<i>džon</i>
Josip	Joseph	<i>džosep</i>
„	Joe	<i>džo (žo)</i>
Jurij	George	<i>džorč</i>
Lovro	Lauernce	<i>lorens</i>
Matija	Matthew	<i>metju</i>
„	Mat	<i>met</i>
Miha	Mike	<i>majk</i>
Nikolaj	Nick	<i>nik</i>
Peter	Petre	<i>pit</i>
Stefan	Steve	<i>stiv</i>
Tomaž	Tommy	<i>tomi</i>

Figure 13: Slovenian first names with English equivalents and pronunciation (Kubelka 1904: 135).

DEC 30 1904

— 136 —

KAZALO.

~~~~~

|                                                   |       |        |
|---------------------------------------------------|-------|--------|
| Angleška abeceda . . . . .                        | Stran | 11     |
| Slovenska abeceda . . . . .                       | "     | 11     |
| Izgovarjanje samoglasnikov . . . . .              | "     | 12     |
| Izgovarjanje sestavljenih samoglasnikov . . . . . | "     | 13     |
| Izgovarjanje soglasnikov . . . . .                | "     | 14     |
| Izgovarjanje sestavljenih soglasnikov . . . . .   | "     | 14     |
| Letni časi . . . . .                              | "     | 15     |
| Meseci . . . . .                                  | "     | 15     |
| Dnevi tedna . . . . .                             | "     | 16     |
| Posebni dnevi. . . . .                            | "     | 16     |
| Človeško truplo . . . . .                         | "     | 17     |
| Številke . . . . .                                | "     | 20     |
| <b>Rečnik</b> . . . . .                           | "     | 24—121 |
| Lahki pogovori:                                   |       |        |
| Pozdravi . . . . .                                | "     | 122    |
| V družbi . . . . .                                | "     | 123    |
| Povprašanja . . . . .                             | "     | 125    |
| Življenju potrebne stvari. . . . .                | "     | 127    |
| Za delom . . . . .                                | "     | 128    |
| Imena Ziednjenih Držav . . . . .                  | "     | 132    |
| Denar Zjedinenih Držav . . . . .                  | "     | 134    |
| Imena . . . . .                                   | "     | 135    |

LB 1117-12

Figure 14: Table of contents, (Kubelka 1904: 136).

Most immigrants of this era were ignorant of the languages of the countries through which they traveled, and few spoke English. Instructions written in their native language were accordingly more than welcome. While this work is just a phrasebook, its importance for a new immigrant should not be underestimated. As was emphasized by Drnovšek (2006): “Even on Ellis Island (New York), many emigrants were given bilingual handbooks such as ... Kubelka’s ... (1912). These contained simple dialogues, which immigrants needed to know in order to communicate in their new environment” (p10).

There are 16 different topics identified in the table of contents; some examples are: *The Voyage; Forwarding of Baggage; The Rail Road: At the Depot; On Ellis Island; Looking for Work; Taking Lodgings*. After the list of topics, the central part of the phrasebook follows, in which each topic is treated. As one would expect from a phrasebook of this type, each topic contains typical conversational phrases, short dialogs, or other texts considered useful for Slovenian immigrants. In Column 1, Slovenian phrases are listed; Column 2 provides translations of the Slovenian phrases into English, and Column 3 contains the pronunciation of the English phrases – this is the same organization we have already seen in Kubelka (1904). An example of a dialog that might have been used at Ellis Island is shown in Figure 14.

—9—

#13 (6

**NA ELLIS ISLANDU**  
(Naselniški otok.)  
**ON ELLIS ISLAND.**  
(*on elis ajlend.*)

| SLOVENSKO              | ANGLĚŠKO                     | IZGOVARJANJE             |
|------------------------|------------------------------|--------------------------|
| Kako je vaše ime?      | <i>What is your name?</i>    | vot is jur nem?          |
| Koliko ste stari?      | <i>How old are you?</i>      | hau old ar ju?           |
| Ali ste oženjeni?      | <i>Are you married?</i>      | ar ju merid?             |
| Ne, gospod, samec sem. | <i>No, Sir, I am single.</i> | no, ser, aj em<br>singl. |

Figure 15: On Ellis Island (Kubelka 1912b: 9).

The pronunciation presented is neither IPA nor a respelling, but a spelling geared toward a native speaker of Slovenian with no knowledge of the English phonological system. It is very simplified; the American English phonemes not present in Slovenian (e.g., /θ/, /ð/, /æ/, /ŋ/) are completely ignored.

At the end of the phrasebook, there is a full-page advertisement for the Austro-American Line (with an office address in Trieste). On the back cover is a picture of a steamer ship and a list of steamers that sail between Trieste and New York, along with the deadweight tonnage (DWT) of each (see Figure 15).

**3.2.3 Kubelka's *Slovensko-angleška slovnica, tolmač, spisovnik in navodilo za naturalizacijo. Angleško-slovenski in slovensko-angleški slovar*, Slovenian–English Grammar Interpreter, Letterwriter and Information on Naturalization. English–Slovenian and Slovenian–English Dictionary, 1912b.**

Kubelka (1912a) is the most comprehensive of Kubelka's three reference books both in terms of its structure and content. As was previously mentioned, the book has 431 pages including the ads. After the two title pages, first in Slovenian and then in English, there is a preface in Slovenian written in August 1912. In it, the author notes that because the demand for his various Slovenian–English books published in the previous eight years has been so great, he decided to write a new, improved, and more extensive book. The new book, Kubelka says, is the most complete book Slovenians have for learning English. Kubelka maintains that “The book is organized so that it meets every single demand and *everyone* using it will most certainly be satisfied” (p9).

The book is divided into the following six parts: English Grammar, Every-Day Conversations, Slovenian–English Letterwriter, Information on Naturalization, English–Slovenian Dictionary, Slovenian–English Dictionary. The first part (English Grammar, 10–86) provides essentially the same type of information that we saw in Jeram (1895). While the format does not correspond exactly to that of Jeram, it is strikingly similar. Like Jeram, all of the explanations in Kubelka's English Grammar section appear in Slovenian. This section treats: alphabet (containing English and Slovenian alphabets), the pronunciation of vowels, diphthongs and consonants, articles, nouns, verbs, adjectives, numerals, as well as many more topics. The internal organization of each grammar topic follows a similar pattern. For example, the topic dealing with articles (27) contains a brief discussion (in Slovenian) of both the definite and indefinite articles, followed by one translation exercise. The translation exercise begins with a short glossary with words that the learner can make use of in the sentences to be translated.



**Parobrodi Avstro-Amerikanske**  
**Proge:**

Novi ekspresni parobrod (na dva vijaka):  
 "Kaiser Franz Joseph I" - 18,000 ton

Drugi brzi parobrodi (na dva vijaka):

|                     |   |        |   |
|---------------------|---|--------|---|
| "Martha Washington" | - | 14,500 | " |
| "Laura"             | - | 10,500 | " |
| "Alice"             | - | 10,500 | " |
| "Argentina"         | - | 10,000 | " |
| "Oceania"           | - | 10,000 | " |
| "Atlanta"           | - | 10,300 | " |
| "Columbia"          | - | 10,000 | " |

in  
 "Sofia Hohenberg," "Francesca," "Eugenia,"  
 "Georgia"

Figure 16: Back cover (Kubelka 1912b).

The sentences to be translated follow a theoretical discussion and are divided into two paragraphs. The first paragraph has Slovenian sentences to be translated into English, and the second paragraph has English sentences to be translated into Slovenian. Like in Jeram, the sentences are semantically unrelated and form neither a true paragraph nor a narrative.

While, as mentioned, the internal organization of each grammar topic follows a similar pattern, not all of the topics have an identical structure and the same level of complexity. Unlike the Article section with no subsections, the Verb is a very long section (50–76) that contains four subsections: Auxiliary Verbs, Regular Verbs, Conjugation of Regular Verbs, and Irregular Verbs.

The second part of the book (p88–146) is entitled *Everyday Conversations*. The structure here resembles that of Kubelka (1904) in terms of topics; however, Kubelka (1912a) has a wider variety of topics. The content of this part of Kubelka (1912a) is not identical to the earlier book, indicating a reworking of the 1904 material. In Kubelka (1912a), some phrases are identical to Kubelka (1904), others differ slightly, and some are new. The third part of the book, with the title *Slovenian–English Letterwriter* (147–191), contains guidance on how to write a letter in English, including beginnings, endings, salutations, what to write on an envelope, numerous sample letters, etc. Then follows *Information on Naturalization* (192–208), where the naturalization procedure is described, including questions that could be posed by a judge to a person applying for American citizenship.

As in Kubelka (1904), the central part (Parts 5 and 6) of this more comprehensive book is a bidirectional dictionary: *English–Slovenian Dictionary*, *Slovenian–English Dictionary*. Unlike Kubelka (1904), there is no outer access structure at all, no running heads to represent the first or the last lemmas on a page. While an outer access structure is not used, the new dictionary is more sophisticated than Kubelka (1904), which (as we have seen) is organized into three columns per page: lemma + equivalent(s) + pronunciation. Kubelka (1912a) looks more like a familiar 20<sup>th</sup>-century dictionary: Each page has two columns, each of which contains a list of entries.

The English–Slovenian portion of the dictionary contains 85 pages (210–294). The structure, as shown in Figure 16, is as follows: The English word (which does not appear in boldface) is followed by a comma, followed by Slovenian equivalents written in italics and separated by commas.

It is noteworthy that no semi-colons at all are used to separate the equivalents. The ordering of the equivalents appears to be by part of speech; however, it also appears that this ordering is not entirely consistent. Nouns and verbs are listed first and second, but there is fluctuation in the ordering of adjectives and adverbs.

—254—

|                                                |                                                                         |
|------------------------------------------------|-------------------------------------------------------------------------|
| lofty, visok, visien.                          | madam, gospa.                                                           |
| long, dolg, krepneti.                          | madman, presmudnec.                                                     |
| longitude, zemljepisna dol-<br>žina.           | magic, čarovnija, čaroben.                                              |
| look, pogled, pogledati, gle-<br>dati, iskati. | magistrate, mirovni sodnik,<br>uradniš gospe.                           |
| looking-glass, zrcalo, ogled-<br>dalo.         | magnet, magnet.                                                         |
| loom, tkalni stol, stavce.                     | magnify, povečati, povekša-<br>ti.                                      |
| loose, rahel, prost, nemaren,<br>raspušteni.   | mahogany, mahogonovina.                                                 |
| Lord, Bog, gospod.                             | mail, devica, deklica.                                                  |
| lose, izgubiti, napraviti.                     | mail, pošta, po pošti poslati.                                          |
| loss, izguba, škoda.                           | main, veliki, ogromen, gla-<br>ven.                                     |
| lot, zгода, delež, temeljite.                  | maintain, obdržati, držati.                                             |
| loud, glasen, glasno.                          | majesty, veličanstvo.                                                   |
| love, ljubezen, ljubiti.                       | majestic, veličanstven.                                                 |
| lover, ljubimec, ljubček.                      | major, major, polkolonel.                                               |
| low, nizek, nizek, globok.                     | majority, večina.                                                       |
| lower, spuščati v globino.                     | make, delo, oblika, storiti,<br>napraviti.                              |
| luck, sreča, naključje.                        | maker, izdelovatelj.                                                    |
| lucky, srečen.                                 | malady, bolezen.                                                        |
| lumber, les (starbi les).                      | male, molki.                                                            |
| lump, gruda, kupa.                             | malice, hudobija, jeza.                                                 |
| lunatic, blazen.                               | malicious, hudoben, jezen.                                              |
| lunatic asylum, norišnica.                     | mallet, leseno bičivo.                                                  |
| lunch, prigrizek, drugi voj-<br>tek, lunč.     | malt, slad.                                                             |
| lungs, pljuča.                                 | man, mož, človek.                                                       |
| luxury, obilnost, razkošje.                    | manage, oskrbovati, uprav-<br>ljati, voditi, izzviti, nad-<br>zorovati. |
| lye, lug.                                      | management, oskrbovanje,<br>uprava, vodstvo, nadzor-<br>stvo.           |
| lyric, ljudska sodba, lin-<br>čanje.           | manager, nadzornik, oskrb-<br>nik, upravitelj, voditelj.                |

**M**

Figure 17: Part 5, English–Slovenian Dictionary (Kubelka 1912a: 254).

For example, *yell* has three equivalents; the first two are nouns (*krik*, *kričanje*) followed by one verb (*kričati*). The ordering of adverb and adjective differs in the entries for *round* and *wrong*. *Round* has four equivalents separated by commas in the following order: noun (*krog*), verb (*krožiti*), adverb and preposition (*okolo*),<sup>2</sup> and finally adjective (*okrogel*). The word *wrong* has four equivalents, again separated by commas, in the order: noun (*krivica*), two adjectives (*neprav*, *krivičen*) and adverb (*narobe*).

The Slovenian–English part of the dictionary consists of 128 pages (p296–423). The Slovenian word appears first in the entry in boldface – unlike the English words in the English–Slovenian section which are not in boldface. After the Slovenian word, a comma follows and then the English equivalents (also not in boldface). Immediately following each English equivalent is the pronunciation of the English word in parentheses. Each pronunciation is followed by a comma, and if applicable, another English equivalent. The pronunciations follow the same system used in Kubelka (1904): neither IPA nor a respelling, but a spelling that Slovenian speakers are likely to understand. Just like in the English–Slovenian section of Kubelka (1912a), no semi-colons are used to separate equivalents.

The ordering of equivalents in the Slovenian–English section follows a different system as compared to the English–Slovenian section. This is due to the difference in the two languages. While a single English word can have Slovenian equivalents of several parts of speech, Slovenian words are clearly marked in their morphology for part of speech, so that English equivalents are all a single part of speech. Because of the lack of ambiguity in Slovenian words regarding part of speech, there are usually fewer English equivalents for each Slovenian word, as compared to the number of Slovenian equivalents for each English word. When, infrequently, there is more than one English equivalent listed for a Slovenian lemma, the ordering of the equivalents does not appear to follow any discernible rules – semantic, alphabetical, or otherwise. For example, the Slovenian verb *zmenjati* has two English equivalents: *to exchange for*, and *to change*. Note that they are not listed in alphabetical order; it is not possible to understand whether there is a semantic motivation for the ordering.

After the Slovenian–English dictionary, there is a table of contents in Slovenian and six advertisements. The first advertisement is for the Austro-Americana Line, although this is not the same advertisement that appeared in Kubelka

---

<sup>2</sup>This is an obsolete Slovenian word used as an adverb and preposition. It is not included in any contemporary Slovenian reference books but appears in a 16<sup>th</sup>-century Slovenian dictionary and in a late 19<sup>th</sup>-century bilingual Slovenian–German dictionary.

(1912b) (and a New York address is now included). Another advertisement is for the main representative of the Austro-Americana Line in the west, with an address in Chicago. The third advertisement (see Figure 17) is for The National Herald [*Narodni vestnik*] weekly newspaper and for its publisher, the Slovenian Print & Publishing Company [Slovenska tiskovna družba] in Duluth, Minnesota. There is also an advertisement for another newspaper, The People's Voice [*Glas naroda*], which calls itself “the **only** Slovenian newspaper in America” (despite the ad in the same publication for The National Herald). There is also an ad for Kubelka himself, publisher of Slovenian–English books. On the back cover is, again, as in Kubelka (1912b), a picture of a steamer and a list of steamers sailing between Trieste and New York.

### 3.2.4 Košutnik's Ročni slovensko-angleški in angleško slovenski slovar: Zlasti namenjen izseljencem v Ameriko [A Pocket Slovenian–English and English–Slovenian Dictionary: Intended for Immigrants to America], 1912.

Silvester Košutnik's dictionary is the sole book being examined that was published outside of the US (Ljubljana: Anton Turk), although clearly it was intended for Slovenians planning to emigrate. The title page and the copyright page (Figure 18) are followed by a preface in Slovenian. According to the preface: “... in this dictionary, the most common words are collected, i.e., those that are used most frequently in everyday life” (p1). The preface contains an explanation of the pronunciation of English.

The dictionary that follows is a bidirectional dictionary, first 59 pages of Slovenian–English (5–63), and then 85 pages of English–Slovenian (64–148). Page 149 contains a list of books published by Anton Turk, the publisher of this dictionary.

The Slovenian–English dictionary (Figure 19) has a very simple microstructure: lemma in Slovenian, dash, equivalent in English, comma if there is another equivalent, period at the end. The English–Slovenian dictionary (Figure 20) is about 25 pages longer than the Slovenian–English section. It has the following structure: English lemma, comma followed by a simplified pronunciation (with primary word stress provided as an apostrophe immediately after the stressed syllable), then a dash and the equivalent(s) separated by a comma, even if they are not synonyms. In both the Slovenian–English and the English–Slovenian parts, a semi-colon is used to separate different parts of speech. For example, in Figure 20, English *holy* is followed by the Slovenian adjective *svet*, which is separated by a comma from the adjective *prazničen*, which means more ‘holiday’ than ‘holy’.

**Naročajte se**  
na največji slovenski tednik na Severozapadu

**“NARODNI VESTNIK”**

Kateroga izdajo:

**SLOVENSKA TISKOVNA DRUŽBA**  
31 East Mich. St., Detroit, Minn.

Naročnina za Ameriko:

|                      |        |
|----------------------|--------|
| za celo leto .....   | \$2.00 |
| za pol leta .....    | 1.00   |
| za četiri leta ..... | .75    |

Za Evropo in Kanado:

|                    |        |
|--------------------|--------|
| za celo leto ..... | \$2.50 |
| za pol leta .....  | 1.25   |

Izdaja vsak četrtak na obeh straneh.  
Najcenoje:  
“Priljubljen in napredni narod,  
V uspehu in blagi dežji.”

**TISKARNA**  
SLOVENSKE TISKOVNE DRUŽBE  
31 East Mich. St., Detroit, Minn.

izvedaje vse v tiskarsko strojno spajajoča dela kakor:  
dravitvena pravila, platneni papir, nalepke, vabila,  
brnjake, vzgojnice, račune, vianke, cirkularje itd. ter  
se pripravijo s. dravitvam, trgovcem in občanom.

Cene gornje. Dela Minn. Povračila točno.

Dajte nam za ceno in vsega raznovrstnih tiskovin  
prema naročila draga.

Figure 18: Advertisement from the final pages (Kubelka 1912a).



Figure 19: Title page (Košutnik 1912).

These two adjectives are separated by a semi-colon from the noun *praznik* meaning ‘holiday’.

On the other hand, note the entry for *home* (Figure 20), which diverges from the dictionary’s usual procedure with semi-colons. Here, *dom* and *domovje*, both nouns meaning ‘home’, are separated from the two following nouns – the same part of speech – meaning ‘apartment’ and ‘dwelling’. Uncharacteristically, the semi-colon is being used to indicate a semantic difference. The final three words in the entry are an adjective and two adverbs. *Domač*, an adjective, means ‘domestic’; a semi-colon is not used to separate it from the adverb *domu* ‘going toward home’, or from the adverb *doma* ‘at home’.

**V.**

|                                                      |                                                       |
|------------------------------------------------------|-------------------------------------------------------|
| <b>v</b> - in, into, at, to, within.                 | <b>včas</b> i - frequently, often, some-<br>times.    |
| <b>vabiti</b> - to lade in, to load in to<br>invite. | <b>vč</b> eraj - yesterday.                           |
| <b>vaditi</b> - to exercise, to practise.            | <b>vd</b> ova - widow, dowager.                       |
| <b>vaja</b> - exercise, practice.                    | <b>vd</b> ovec - widower.                             |
| <b>vajen</b> - exercised, versed, expert.            | <b>ve</b> č - more.                                   |
| <b>valjati</b> - to roll, to turn about.             | <b>ve</b> čen - eternal, everlasting, per-<br>petual. |
| <b>varati</b> - to cheat, to deceive.                |                                                       |

Figure 20: Slovenian–English dictionary, (Košutnik 1912: 56).

### 3.2.5 Kern’s A Complete Pronouncing Dictionary of the English and Slovene Languages for General Use, *Popoln angleško-slovenski besednjak z angleško izgovarjavo*, 1919*Popoln angleško-slovensk.*

Fran/Frank Javh Kern (Figure 21) was born in 1887 in Breznica and died in Cleveland, Ohio in 1977. He arrived in the US at the age of 16 (1903) to attend the Saint Paul Seminary in Saint Paul, Minnesota (established 1894). He brought with him a good gymnasium education from his home country. He later became a physician in Cleveland as well as a prominent community member.

Kern’s A Complete Pronouncing Dictionary of the English and Slovene Languages for General Use, *Popoln angleško-slovenski besednjak z angleško izgovarjavo* (Kern 1919), was published in 1919, at a time when immigration to the US (and worldwide) was generally slowing due to WWI. At that time, the number of people of Slovenian origin in the US was quite high, especially in Cleveland, where Kern lived.

Kern lists the English title first on the title page (Figure 22).

hold, hold - držati, obdržati,  
ceniti, meniti, dobiti, kratko-  
časiti, držati se, vztrajati.  
hole, hol - luknja, votlina, iz-  
govor.  
hollow, hol'lo - izvotljen, ne-  
prav; votlina, luknja, hodnik,  
rov.  
holy, ho'li - svet, prazničen;  
praznik.  
home, hom - dom, domovje;  
stanovanje, bivališče; domač,  
domu, doma.

Figure 21: English–Slovenian dictionary, (Košutnik 1912: 103).

A preface in English (III) and another in Slovenian (IV–V) immediately follow the title page and the copyright page. Kern’s treatment of pronunciation is explained in the English preface:

... I have followed the latest authorities, particularly the New Standard Dictionary,<sup>3</sup> which employs the phonetic method of indicating the pronunciation, and which is similar to the phonetic system of spelling and writing employed in the Slovene (Slovenian) language. (III)

In the Slovenian preface, Kern explained how he selected Slovenian equivalents; he considered

... the general use and most contemporary books. In technical and scientific terms Slovenian terminology is not determined precisely enough for me to be able to get an appropriate term in each and every case. In more important words, I added several synonyms (similar expressions) for explanation. (1919, IV).

---

<sup>3</sup>The first edition of the Funk and Wagnalls dictionary with “new” in the title appeared in 1913.

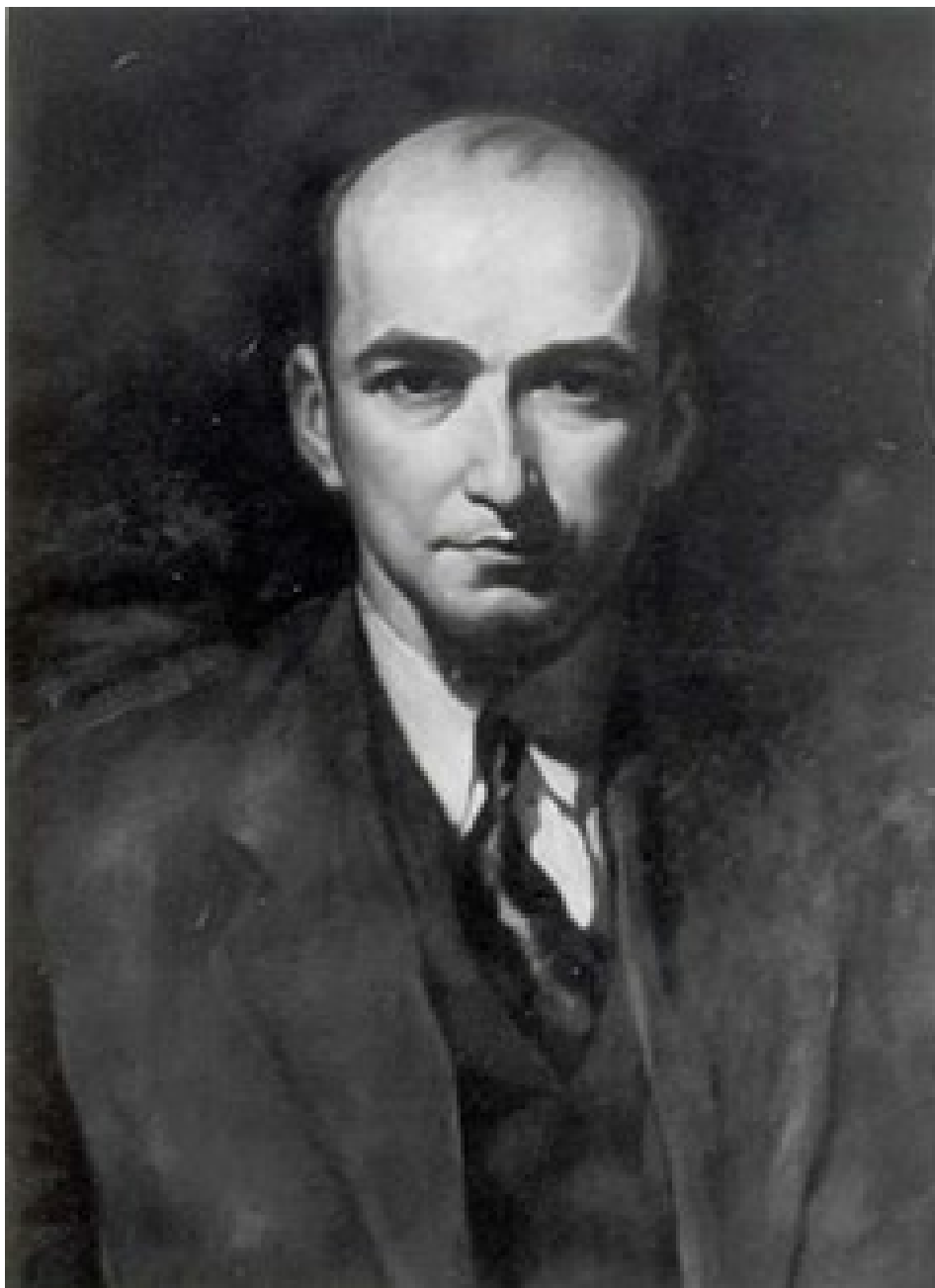


Figure 22: Frank/Fran Javh Kern.

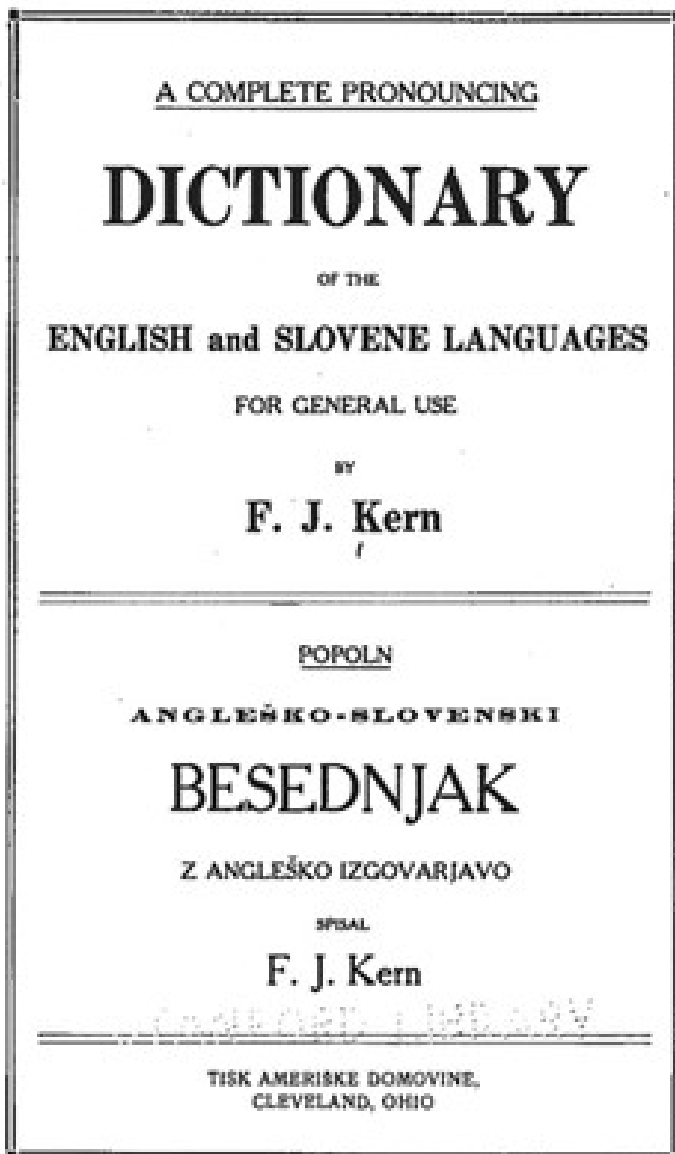


Figure 23: Title page (Kern 1919).

Also noted in the Slovenian preface:

Spelling and pronunciation of English words follows the American tradition and is somehow different from continental English.

Most likely Kern is the first among the authors addressed here to notice that British and American spelling and pronunciation differ.

Following the two prefaces, on pages VI–VII, we find a user's guide (*Pojasnila*) which mostly treats pronunciation; there is a very brief clarification of how to use the dictionary (an explanation of running heads and alphabetical arrangement of entries) and a brief note on compounds.

The Kern dictionary has two-column pages with running heads (i.e., an outer access structure). Each page has two running heads consisting of three letters (Figure 23). The left running head contains the first three letters of the first lemma on the page; as would be expected, the first three letters of the page's last lemma appear in the running head on the right. Among all the works examined here, only Kubelka (1904) and Kern have running heads. Kern's English lemma is printed in boldface and is followed by a comma; next, pronunciation appears in parentheses followed by a comma; finally, we have the Slovenian equivalent(s).

The use of punctuation in Kern is more systematic than, for example, in Košutnik (1912). There are two uses for semi-colons. First, they separate equivalents of the same part of speech when there are perceived semantic differences. For example (Figure 23), for English utter, four adjectives appear as equivalents: *skrajn* 'extreme', *prevelik* 'too big', *popoln* 'complete', and *cel* 'complete, whole'. The first two, *skrajn* and *prevelik*, are separated by semi-colons due to their perceived semantic differences. Kern clearly considered the last two adjectives, *popoln* and *cel*, as closer semantically; so, they are separated by a comma. We could quibble over Kern's choice of adjective equivalents or his judgments about semantic closeness, but it is clear that he is employing a systematic approach.

While part of speech is not listed in Kern's dictionary, the second use for the semi-colon is to separate equivalents of different parts of speech. For example, in Kern's *utter* (Figure 23), two equivalents – verbs – follow the adjective *cel* 'complete': *izreči* 'say' and *govoriti* 'talk, speak'. These are separated by a semi-colon from the preceding adjective *cel*. It is interesting that a semi-colon is not used to separate the two verbs from each other; instead, Kern uses a comma – despite the fact that the two verbs are not synonyms.

At the end of the dictionary, we can find a table of the irregular verbs spanning the final three pages (p271–273).

| 258                                                                     | USU | VAP                                                                     |
|-------------------------------------------------------------------------|-----|-------------------------------------------------------------------------|
| <b>usurper</b> , (juzr'pr), vsiljivec, prilastovalec.                   |     | daljen; jako, zelo.                                                     |
| <b>usury</b> , (jū'žuri), skopušstvo.                                   |     | <b>Utopia</b> , (jutō'piā), utopija (kraj vseh populnosti).             |
| <b>utensil</b> , (juten'sil), priprava, orodje, potrebščina.            |     | <b>utopian</b> , (jutō'pien), sanjarski, nemo-goč; sanjar.              |
| <b>uterine</b> , (jū'terajn), maternični.                               |     | <b>utter</b> , (ātr), skrajen; prevelik; popoln, cel; izreči, govoriti. |
| <b>uterus</b> , (jū'teres), maternica.                                  |     | <b>utterance</b> , (ā'terens), izrek; govoriga.                         |
| <b>utility</b> , (juti'liti), rabljivost, korist; javna naprava.        |     | <b>utterly</b> , (ā'trli), skrajno, popolnoma.                          |
| <b>utilization</b> , (jūtilizēj'sn), uporabnost.                        |     | <b>uttermost</b> , (ā'trmōst), glej <b>utmōst</b> .                     |
| <b>utilize</b> , (jū'tilajz), uporabiti, izkoristiti.                   |     | <b>uvula</b> , (jū'vjulā), mali jeziček.                                |
| <b>utmōst</b> , (āt'mōst), najvišji, najbolj od-                        |     | <b>uxorious</b> , (juksō'ries), pretirano podložen ženi.                |
| <b>V</b>                                                                |     |                                                                         |
| <b>vacancy</b> , (vēj'kensi), praznosta; prazno mesto; neoddana služba. |     | <b>valet</b> , (vē'let), strežaj, lakaj.                                |
| <b>vacant</b> , (vēj'kent), prazen; neoddan; nemisleč.                  |     | <b>valetudinarian</b> , (vē'litjū'dinē'rien), kroničen bolnik.          |
| <b>vacate</b> , (vēj'kējt), izprazniti; zapustiti.                      |     | <b>valiant</b> , (vē'ljent), hraber, neustrašen; —ly, hrabro, junaško.  |
| <b>vacation</b> , (vekēj'sn), počitnice; počitek; izpraznjenje.         |     | <b>valid</b> , (vē'lid), veljaven; pravomočen; dober.                   |
| <b>vaccinate</b> , (vēk'sinējt), cepiti koze.                           |     | <b>validity</b> , (veli'diti), moč, pravomočnost; veljava, vrednost.    |
| <b>vaccination</b> , (vēksinēj'sn), cepljenje koz ali drugih vakcin.    |     | <b>valise</b> , (velis'), ročni popotni kovčeg.                         |
| <b>vaccine</b> , (vēk'sin), cepilo, vakcina.                            |     | <b>valley</b> , (vē'li), dolina.                                        |
| <b>vacillate</b> , (vē'silējt), omahovati.                              |     | <b>valor</b> , (vēlr), hrabrost.                                        |
| <b>vacillation</b> , (vēsilēj'sn), omahovanje, negotovost.              |     | <b>valorous</b> , (vē'lores), hraber.                                   |
| <b>vacuity</b> , (vekjū'iti), praznina; pomanjkljaj; ničnost.           |     | <b>valuable</b> , (vē'ljuebl), dragocen, velike vrednosti.              |
| <b>vacuous</b> , (vē'kjues), prazen.                                    |     | <b>valuation</b> , (vē'ljuēj'sn), cenitev.                              |
| <b>vacuum</b> , (vē'kjuem), brezračje, va-                              |     | <b>value</b> , (vē'lju), vrednost, cena; ceniti.                        |
|                                                                         |     | <b>valueless</b> , (vē'ljules), brez vrednosti.                         |
|                                                                         |     | <b>valve</b> , (vēlv), zaklopka, zaklopnica.                            |

Figure 24: Kern 1919: 258.

## 4 Conclusion

The dictionaries and other reference works examined here were designed to help the large population of Slovenian immigrants, “liberated from the Hapsburg yoke” (Kern, 1919, III) in the late 19<sup>th</sup> and early 20<sup>th</sup> centuries, to learn English and acculturate to their strange new American environment. The present effort has focused on six books; there are others to be analyzed and likely still others that have not yet been identified. While we have mainly examined macrostructure, there is much more to be said about the microstructure, beyond the parsimonious remarks above. The investigation so far leads us to the following reflections and preliminary conclusions.

Some of the books are relatively standard 20<sup>th</sup>-century print dictionaries (Kern 1919 and Košutnik 1912) that look familiar to a modern eye. On the other hand,

other books are more hybrid works (Kubelka 1904), (Kubelka 1912b), and (Kubelka 1912a) that aimed to meet the needs of a specific time and place and did not concern themselves with replicating traditional scholarly models (if they were even known). In a class by itself, Jeram (1895) grammar, while apparently following prevailing language-learning models popular at the time, may have been rather mysterious to the mostly uneducated Slovenian immigrants who used it; it certainly does not contain exercises of the type that are given to language learners today.

However we may regard the quality of these reference works today, all indications are that their authors were motivated by a sense of duty to a population in need. Some of the books contained language in the form of dialogs on topics that would have been quite relevant in everyday conversations with American English native speakers. What is more, these books gave Slovenian immigrants instructions on how to become a naturalized citizen – not an easy thing to accomplish in any era of American history. As we have seen, these books were constructed quite simply for the most part (Jeram 1895 is perhaps the exception). But if you are uneducated or poorly educated and not in the habit of using any dictionary or reference work at all, isn't simplicity better?

Finally, many of these books contained advertisements. Likely the ads themselves had the purpose of offsetting the books' publishing costs, but it is also evident that the types of ads included would have been useful to the target audience. The ads for US–Slovenian-language newspapers pointed their public toward consequential sources of social cohesion for people thrust into a new milieu. The ads for steamer ships helped orient people planning to emigrate or hoping to bring relatives to the United States. Ads also helped people maintain contact with those back in Slovenia, because they provided help in sending money home.

The books examined here and others of this type targeted toward Slovenian immigrants were published in very small print runs. Today, they are dispersed in libraries and archives of varied size, sometimes in paper and sometimes in digital format. Not all (it seems) have been digitized and some lie in poor condition, in little-used, dusty collections. Their continued study and preservation is imperative, and this is why our voyage through Slovenian–American lexicography must continue.

## **Acknowledgements**

It is our pleasure to acknowledge the Slovenian Research Agency for providing funding for the grants *Slovarji in izkušnje Slovencev–emigrantov v Združenih*

*državah Amerike (konec 19. in začetek 20. stoletja)/Dictionaries and the Slovenian Immigrant Experience in the United States (Late 19th, early 20th Century)* [BI-US/22-24-042] and *Stare besede, nove besede, novi svet: Življenje slovenskih leksikografov - priseljencev v ZDA/Old Words, New Words, New World: The Lives of Slovenian Immigrant Lexicographers in the U.S.* [BI-US/22-24-030].

## References

- Anonymous. 1895. *Rev. of slovensko-angleška slovnica by peter jos.* Tower, Minnesota: Ljubljanski zvon 15.9: 575.
- Baker, Thomas M, John B. Allcock, Karl Lavrencic & Anton Gosar. 2023. *Slovenia.* <https://www.britannica.com/place/Slovenia> (25 March, 2023).
- Baraga, Frederic. 1853. *A dictionary of the Otchipwe language, explained in English. This language is spoken by the Chippewa Indians, as also by the Ottawas, Potawatamis and Algonquins, with little difference. For the use of missionaries, and other persons living among the above mentioned Indians.* Cincinnati: Jos. A. Hemann.
- Béjoint, Henri. 2016. Dictionaries for general users: History and development; current issues. In Philip Durkin (ed.), *The Oxford handbook of lexicography*, 7–74. Oxford: Oxford University Press.
- Cesar-Nedzbala, Polonca. 1990. Rajska domovina. *Dve domovini/Two homelands* 1. 83–106.
- Drnovšek, Marjan. 1994. Mass emigration and Slovenes. *Dve domovini/Two Homelands* 5. 19–36.
- Drnovšek, Marjan. 2006. *Ljubljana: Križišče na poti v svet. Množično izseljevanje slovencev v ameriko. Ljubljana: The crossroads to the world.* Ljubljana: ZRC SAZU & Numizmatično društvo Slovenije. <http://ljubljana-kps.zrc-sazu.si/Brosura/Ljubljana%20-%20Krizisce%20na%20poti%20v%20svet.pdf> (18 July, 2022).
- Drnovšek, Marjan. 2001. *The Causes for Emigration of Slovenes in the Last Two Centuries.* <http://www.theslovenian.com/articles/drnovsek.htm>.
- Farina, Donna M. T. Cr. & George Durman. 2009. Bilingual dictionaries of English and Russian in the eighteenth to the twentieth centuries. In Anthony P. Cowie (ed.), *The Oxford history of English lexicography*, vol. 1. : General-purpose Dictionaries, 105–126. Oxford: Oxford University Press.
- Ferrie, Joseph P. & Timothy J. Halton. 2013. *Two centuries of international migration.* Discussion Paper 7866. Bonn: Institute for the Study of Labor. <https://www.iza.org/publications/dp/7866/two-centuries-of-international-migration> (15 July, 2022).

- Fontenelle, Thierry. 2016. Bilingual dictionaries: History and development; current issues. In Philip Durkin (ed.), *The Oxford handbook of lexicography*, 44–61. Oxford: Oxford University Press.
- Gerber, David A. 2011. *American immigration: A very short introduction*. Oxford: Oxford University Press.
- Gobet, Edward. 2014. Slovenian Americans. In Thomas Riggs (ed.), *Gale encyclopedia of multicultural America*, vol. 4, 223–239. Farmington Hills, Michigan, Gale. <https://go.gale.com/ps/i.do?p=GPS&u=wikipedia&id=GALE%7CCX3273300166&v=2.1&it=r&sid=GPS&asid=e849a75f> (9 July, 2022).
- Gobet, Edward. n.d. *Slovenian Americans*. <https://www.everyculture.com/multi/Pa-Sp/Slovenian-Americans.html>.
- Grafenauer, Ivan. 2013. Jeram, Peter Josip (okoli 1865–okoli 1897) - slovenska biografija. In *Slovenska biografija*. Ljubljana: Slovenska akademija znanosti in umetnosti, Znanstvenoraziskovalni center SAZU. <https://www.slovenska-biografija.si/oseba/sbi255628/#slovenski-biografski-leksikon>.
- Hernández Jaramillo, Alejandro. 2019. *A historical overview of the “grammar-translation” method for teaching foreign languages*. <https://www.noslangues-ourlanguages.gc.ca/en/blogue-blog/methode-grammaire-traduction-grammar-translation-method-eng?wbdisable=true#auteur-author> (22 July, 2022).
- Hoyt, Joanna Micham. 2017. *A wary welcome: The history of U. S. attitudes toward immigration*. Middletown, DE: Skinny Bottle Publishing.
- Jeram, Peter Jos. 1895. *Slovensko-angleška slovnica*. Tower, Minnesota: Amerikanski Slovenec.
- Kern, Frank J. 1919. *A complete pronouncing dictionary of the English and Slovene languages for general use. Popoln angleško-slovenski besednjak z angleško izgovarjavo*. Cleveland, Ohio: Tisk ameriške domovine.
- Klemenčič, Matjaž. 1986. Research on Slovene immigration to the United States: Past achievements and future directions. *Slovene Studies: Journal of the Society of Slovene Studies* 8(2). 9–14. DOI: 10.7152/ssj.v8i2.3630. <https://journals.lib.washington.edu/index.php/ssj/article/view/3630> (17 April, 2023).
- Košutnik, Silvester. 1912. *Ročni slovensko-angleški in angleško slovenski slovar: Zlasti namenjen izseljencem v ameriko*. Ljubljana: Anton Turk.
- Kubelka, Victor J. 1904. *Slovensko-angleški žepni rečnik v olajšavo naučenja obeh jezikov. Slovenian-English pocket dictionary to facilitate the study of both languages*. New York. <https://babel.hathitrust.org/cgi/pt?id=loc.ark:/13960/t4qj8p026&view=1up&seq=1>.

- Kubelka, Victor J. 1912a. *Slovensko-angleška slovnica, tolmač, spisovnik in navodilo za naturalizacijo. Angleško-slovenski in slovensko-angleški slovar. Slovenian-English grammar interpreter, letterwriter and information on naturalization. English-Slovenian and Slovenian-English dictionary*. 1st edn. New York: Viktor J. Kubelka.
- Kubelka, Victor J. 1912b. *Slovensko-angleški razgovori. Slovenian-English interpreter*. New York: Press of Stettiner Bros. <https://hdl.handle.net/2027/nyp.33433021125889>.
- Mazi-Leskovar, Darja. 2003. Domestication and foreignization in translating American prose for Slovenian children. *Meta : journal des traducteurs / Meta: Translators' Journal* 48(1). 250–265. DOI: 10.7202/006972ar. <https://www.erudit.org/fr/revues/meta/2003-v48-n1-2-meta550/006972ar/> (10 July, 2022).
- Milharčič-Hladnik, Mirjam. 2007. Marie Prisland – her role in preserving Slovenian culture and tradition among Slovenian migrants in the United States. *Dve domovini/Two Homelands* 25. 229–247.
- Prisland, Marie. 1950. The Slovenians, most recent American immigrants. *The Wisconsin Magazine of* 33(3). 265–280.
- Sedmak, Jay. 2021a. Strangers to paradise [part 1]. *Prosveta: SNPJ News* 114(1). 2.
- Sedmak, Jay. 2021b. Strangers to paradise [part 2]. *Prosveta: SNPJ News* 114(2). 7.
- Sedmak, Jay. 2021c. Strangers to paradise [part 3]. *Prosveta: SNPJ News* 114(3). 2.
- Susel, Rudolph M. 2022. *Encyclopedia of Cleveland History*. <https://case.edu/ech/articles/s/slovenes> (17 July, 2022).
- Zhou, Gang & Xiaochun Niu. 2015. Approaches to language teaching and learning. *Journal of Language Teaching and Research* 6(4). 798. DOI: 10.17507/jltr.0604.11. <http://www.academypublication.com/issues2/jltr/vol06/04/11.pdf> (11 April, 2023).
- Žugel, Franc. 1987. Ameriški Slovenci. In Marjan Javornik, Dušan Voglar & Alenka Dermastia (eds.), *Enciklopedija Slovenije*, vol. 1 (A–Ca), 58–62. Ljubljana: Mladinska knjiga.



## Chapter 8

# The incorporation of proper nouns of non-Slavic origin into the 16th-century Slovenian literary language

Alenka Jelovšek

Research Centre of the Slovenian Academy of Sciences and Arts, Ljubljana, Slovenia

After the establishment of the Slovenian literary language as the result of the Reformation movement in the 16th century, a significant number of proper nouns of Semitic, Romanesque and Germanic origin were introduced into it, which posed a great challenge in terms of their orthography, phonology and inflection. As the Protestant translated mainly from Luther's works and some of them also used Latin translations of the Bible, the use of proper names was highly influenced by those languages. The analysis of proper names used in 16th-century Slovenian literary language, based on the complete excerpt of the Slovenian printed works published between 1550 and 1595, shows that the authors various strategies for transferring proper names of different origins: in some cases they were transferred in their original form, but generally they were at least morphologically adapted, as Slovenian is an inflection-rich language; they were also inconsistently adapted orthographically and phonologically, creating various exonyms, which resulted in high degree of variation in the literary language of the period.

### 1 Introduction

Proper names have received much scholarly attention in various disciplines of linguistics in recent decades, including contact linguistics and translation studies, but not many studies deal with the question of integration of foreign proper names into a (literary) language in its older periods, especially those originating



in languages which are not genetically closely related or are even unrelated.<sup>1</sup> This study aims to present the methods of incorporation of proper names of non-Slavic origin into the (Slavic) Slovenian literary language at the time of its establishment in the 16th century, taking into the account the specific linguistic situation in Slovenian lands in the analysed period and the distinctive features of proper names in a language system.

Proper names, in contrast to common nouns/semantic words, do not have lexical meaning, i.e. descriptive content with which they would denote concepts. Their function is to refer and identify; accordingly, their relationship to their referent is direct and not mediated via the lexical content (Schlücker & Ackermann 2017: 312). Nicolaisen (1996: 549) describes them as semantically opaque: “Proper names have content, not meaning, in order to function as identifying and individuating devices”.

Hermans (2015) points out an additional formal property of proper names: their minimal integration into the language system. “Phonologically as well as morphologically they exploit the possibilities available to a given language to a larger extent than is the case in the rest of the lexicon” (Hermans 2015: 12). Similarly, Schlücker & Ackermann (2017) argue that proper names differ from common nouns morphosyntactically.

Because of their specific features, the transfer of proper nouns from one language to another in language contact situation is easier in comparison with semantic words. As Raukko (2017: 94) states, “there is no requirement for target language users to understand the semantics of the foreign items, as long as the identifying function is evident”. Similarly, Nicolaisen (1996: 549) claims that “[a]s long as at least minimal content is perceived in their usage it can be transferred in a linguistic shell and thus keep the name alive”.

The aim of this chapter is to illustrate that such proper names transfer between genetically unrelated or distantly related languages based on the historical material from the period of the establishment of the Slovenian literary language in the 16th century. As a result of numerous translations created in a short period of time, there was a large influx of proper names of non-Slavic origin into Slovenian, which posed a great challenge in terms of their orthography, phonology and inflection. Lacking an established tradition, the translators applied various strategies of transfer which resulted in a high level of variation.

---

<sup>1</sup>For historical integration of names from closely related languages see e.g. Casanova 2012; for integration of names from classical languages, see e.g. Gordon 2012; for adaptation of Hebrew names in various translations of the Bible, including Slovenian, see Krašovec 2006 and Krašovec 2010.

The structure of the paper is as follows: Section 2 brings a general discussion on the strategies of transferring proper names between languages and outlines the general terminology used in the remainder of the chapter. Section 3 gives a brief historical overview of the Slovenian literary language in the analysed period, and Section 4 outlines the general properties of proper names of foreign origin in Slovenian before and at the time of the establishment of the literary languages, as well as giving the description of the material for the analysis. Following this general discussion, Section 5 presents and discusses the findings from the analysis, divided into three subsections, described in more detail at the end of the Section 4. Finally, the concluding remarks section considers the main findings of the analysis and outlines their meaning for Slovenian linguistics.

## 2 The strategies of inter-lingual proper name transfer

Linguists have enumerated various strategies of transferring proper names from one language into another. Hermans (2015) speaks of *copying*, i.e. reproducing proper names in the target text exactly as they were in the source text; *transcription*, under which he subsumes transliteration and adaptation on the level of spelling, phonology, etc.; *substitution*; and *translation* in cases when a proper name acquires "meaning" (quotation marks in original text) in the source language. Somewhat differently, Vermes (2003) speaks of *transference* (= Hermans's copying); *substitution*, which for Vermes encompasses replacing a proper name from the source language with a conventional correspondent in the target language (= Hermans's substitution), as well as traditional transliteration (defined as the replacement of graphological units of the source language by graphological units of the target language, based on conventionally established correspondences, where the target language form makes explicit the phonological value of the original expression (Vermes 2003: 93)); *translation* (rendering the meaning into another language); and *modification* (replacing the original name with one which gives rise to similar contextual implications). Raukko (2017) assumes Vermes's division but considers substitution more closely. He considers any adapted form of a proper name in the target language that differs from the form in the source language an exonym, extending the term from geographical names to any type of proper names (Raukko 2017: 97). According to the types and extent of adaptation of proper names, he distinguished between a) exonyms formed as a result of historical and cultural importance of its referent; b) exonyms arising from shared onomastic heritage (such as names of monarchs); c) transliterated forms, in which graphemes from the source language alphabet are replaced with the

graphemes of the target language alphabet; d) exographs, which change orthography but, ideally, keep the original pronunciation (or change the orthography in order to bring the pronunciation closer to the original; e) grammatically adapted forms that are a result of a grammatical necessity, such as adding a required suffix to express gender, f) exophones, whose written form is not changed, but the phonetic or phonological form is: “In the target language form the individual sounds and the phonotactic form are adjusted to the target language phonological system” (Raukko 2017: 99). “Phonological translation” is typical not only for exophones but generally also for full exonyms. The author also mentions “semantic translation” which occurs only in cases where the original name can be seen not only as a name but as a carrier of semantic content (Raukko 2017: 108). He also draws attention to the influence of “transmitting” or “intermitting” languages that spread their own exonyms to other languages (Raukko 2017: 106).

In the context of 16th-century literary Slovenian, Raukko’s detailed division of forms of adaptation is the most appropriate as a basis for the study of the incorporation of proper names into Slovenian, and will serve as a framework for arranging the analysed data. His use of the term *exonym* for adapted forms of various classes of proper names will also be adopted in the analysis, and the verb *transfer* will be from now on used for describing cases which would fall into the Hermans’s category of *copying* and Vermes’s *transference*.

### 3 The establishment and development of the Slovenian literary language in the 16th century

Literary Slovenian was established in the 16th century as a direct result of a Reformation movement in so-called Inner Austrian lands that were ruled by the Hapsburgs and inhabited by mainly speakers of Slovenian dialects (duchies Carniola with a Slovenian-speaking majority and Carinthia and Styria with significant Slovenian-speaking minorities alongside German-speaking majority *Historical atlas of East Central Europe* (Magocsi & Matthews 1993: 99)). The first two printed Slovenian books, a catechism and a primer, were published in 1550 by an exiled Catholic priest turned Protestant, Primož Trubar, who thus established the first model for literary Slovenian, which he upgraded and partially modified in his subsequent translations of the New Testament (cf. Cooper 1985). In 45 years, 49 Slovenian books, mainly with religious content, were printed by the Slovenian Protestant authors, with the financial help of the Inner Austrian estates and some German Protestant princes (cf. Ahačič 2014: 64–73). In addition, Slovenian was also included in multi-language dictionaries compiled by the German

scholar Hieronymus Megiser (1592, 1603). The main achievement of the period was the translation of the entire Bible by Jurij Dalmatin in 1584. Slovenian was thus the 14th world language with a complete translation of the Bible. In the same year, Adam Bohorič published his Slovenian grammar *Arcticae horulae* (written in Latin). With the publication of the Bible and the grammar, Slovenian literary language achieved the level of standardization and the Bible became the basis of the Slovenian literary standard for the following two centuries (Pogorelec 2011: 35).

#### 4 Proper names in the 16th-century literary Slovenian

As the majority of the published texts were translations of either the Bible texts or German works with religious content, such as postils, hymnals or prayer books, a significant number of proper nouns related to the Christian faith and the Church (both Catholic or Protestant) of Semitic, Romanesque and Germanic origin were introduced into Slovenian, which posed a great challenge in terms of their orthography, phonology and inflection, especially because, before the establishment of the literary language, Slovenian was a seldom written vernacular with no existing orthographic tradition (although some oral supradialectal tradition is presumed to have existed in clerical circles; cf. Pogorelec (2011). Although some records of proper names exist, e. g. in medieval land registers (Jakopin 1986: 73–74) and in the Černjeja manuscript (cf. Mikhailov 1998: 109–133), these were ad-hoc recordings and show high degree of variation.

Not all names of non-Slavic origin appearing in the works of the Slovenian Protestants were equally foreign to speakers of Slovenian in the analysed period. Some of them were familiar to them from church services and sermons, others from their contacts with non-Slavic surroundings, especially in the context of the Holy Roman Empire as the wider political frame into which Slovenian lands were incorporated and to which they were culturally connected; some personal names were also already incorporated into the Slovenian language as Christian names which gradually replaced original Slavic first names by the 13th century (Jakopin 1986: 70). But many names, especially less frequent ones from the Old Testament, were introduced into Slovenian for the first time in the translations of the Bible. As the Protestants translated mainly from Luther's works and some of them also used Latin translations of the Bible, the use of proper names was highly influenced by the two source languages. Their influence extended not only to newly introduced proper nouns but also to the already established Christian names, most often at the orthographical level. Similarly, many Aramaic and He-

brew biblical names were adopted in their Latinized form and therefore exhibit similar adaptations as originally Latin names.

The following Section 5 describes the strategies of incorporating non-Slavic proper names into the 16th century Slovenian. The analysis is limited to proper nouns as the prototypical proper names (Vermees 2003: 92–93), more precisely, to personal names (anthroponyms) and place names (toponyms). As their number in the analysed material is still high, as is the degree of variation on the levels of orthography, phonology and morphology, an exhaustive description of the topic is not possible in the limits of this chapter. Therefore, the analysis will focus on selected proper nouns with higher numbers of occurrences in the 16th-century Slovenian texts.

The first subsection discusses the incorporation of names of Latin origin and those adopted in their Latinized form; the second subsection deals with the names of Germanic origin (mostly from the southern German dialects and New High German; and the third subsection presents names of Semitic (mostly Aramaic and Hebrew) origin, with a high emphasis on the “intermitting” languages – Latin and German.

The material for the analysis of proper names used in 16th-century Slovenian literary language is the complete excerpt of the Slovenian printed works published between 1550 and 1595. The excerpt is in the form of a paper concordance which was compiled as part of the collection of material for the *Dictionary of the 16th-Century Slovenian Literary Language*.<sup>2</sup> The paper concordance is kept at the Section for the History of the Slovenian Language, Fran Ramovš Institute of the Slovenian Language, Research Centre of the Slovenian Academy of Sciences and Arts.

## **5 The incorporation of non-Slavic proper names into 16th-century Slovenian**

### **5.1 Latin or Latinized names**

In the case of names taken from or via Latin, we must distinguish between names from the Christian tradition which were, in the time of the establishment of the literary language, already relatively well incorporated into Slovenian, and other names which only entered the language via Bible translation.

---

<sup>2</sup>The first volume of the *Dictionary of the 16th-Century Slovenian Literary Language*, containing entries from A to D, was published in 2021 and is available online at the Slovenian dictionary portal *Fran.si* since 2022. The dictionary does not contain proprial lemmas, i.e. proper names as dictionary entries (cf. Langendonck 2007).

The names from the first group were already phonologically and morphologically adapted to Slovenian: the Latin nominative endings *-us* and *-as* were commonly replaced with *-už/-ož/-o* and *-ež*, respectively (Lat. *Primus* > Sln. *Primush/Primosh* [Primož], Lat. *Lucas* > Sln. *Lukesh*, Lat. *Marcus* > Sln. *Marko*), or even omitted (Lat. *Augustinus* > Sln. *Augustin*, Lat. *Paulus* > Sln. *Paul/Paval*, Lat. *Petrus* > Sln. *Peter* with an added *schwa* between the final two consonants); the palatalization of sibilants was common not only in the endings but also in the middle of a word (Lat. *Augustinus* > Sln. *Augu**sh**tin*); some of the names underwent further phonological adaptations, such as vowel and consonant reductions etc. (eg. Lat. *Iohannes* > Sln. *Jansh* [Janž] > *Ansh*, Lat. *Georgius* > Sln. *Iuri*). Such names can be considered full exonyms, according to Raukko (see Section 2), differing from the original names orthographically as well as phonologically and morphologically. Some of them are attested as Christian names of the Protestant writers: *Primosh* (*Truber*), *Iuri* (*Dalmatin*), *Iansh* (*Shvvager*), but the authors also sometimes signed their names in the Latinized forms: *Primus Truber*, *Georgius Jurishitz*. The Latin (or German) influence can also be seen in the orthography of the first names: the author of Slovenian Protestant hymns, Marko Kumprecht, spelled his name as *Marco* (< *c* > and not with the < *k* > that was occasionally used in Slovenian works when referring to Mark the Evangelist (*Marko, Markus*).<sup>3</sup>

Although such names had been well established as exonyms in Slovenian, under the influence of the German and Latin translations of the Bible as their source texts, Slovenian translators often transmitted source forms of proper names into their translations. Therefore, in the Bible translations the Latin(ized)<sup>4</sup> forms of names were predominantly used, while in other texts exonyms were more frequent, although sporadically they can also be found in the Bible translations; e.g. in *Dalmatin's* Bible there are 8 occurrences of the exonym *Peter* and 104 occurrences of the Latin form *Petrus* in the Bible verses, while the exonym *Iansh* does not appear in the Bible verses; the only name where the exonym is almost as frequent as the Latin form is Paul (34 occurrences of *Paul* vs. 49 occurrences of *Paulus*). In comparison, *Dalmatin's* predecessor *Trubar* used *Paulus* much more frequently; in the same Bible verses, he used *Paulus* only in four instances. But exonyms are often used in other Slovenian texts, as well as in titles, marginal notes and other additions to the Bible verses in the Bible translations.

<sup>3</sup>Such variants of personal names also reflect the multilingualism of the authors: as no Slovenian schools existed in the 16th century, intellectuals were by necessity multilingual (Ahačič 2014: 23, 43–44) and, as a result of their education in German and Latin schools, often more fluent in those languages than in their native tongue, especially in written communication. As a result, German and Latin orthographical norms greatly influenced their writing in Slovenian.

<sup>4</sup>See further discussion on orthography in possible phonological variation of the original forms.

- (1) INu Ioannefu fo njegovi Iogri letu vfe osnanili <S. Iansh fvoja dva Iogra h'Christufu pohle> (Dalmatin, *Biblia*, 1584)

'And the disciples of John shewed him of all these things <Saint John sends two of his disciples to Christ>'.<sup>5</sup>

Also, some of the authors, such as Sebastijan Krelj and Jurij Juričič, showed preference to source language forms of proper names over exonyms also outside Bible translation (Jakopin 1986: 71).

The second group consists of names of Latin origin that had not already been incorporated into the Slovenian at the time of the establishment of the literary language. Their nominal forms were typically transferred into Slovenian texts, although for Latin names ending in *-ius* and *-as*, apparent exonyms with nominative ending *-a* or *-i* can be found in the Bible translations. But the comparison with the Latin and German translations shows that such forms were usually not true exonyms but transfers of Latin or Luther's German vocative forms into Slovenian, which at the analysed period no longer had distinctive case endings for vocative, which had been replaced with nominative endings.

- (2) **Trubar 1557:** Ta ifti ie uidil utim uidenu ozhitu okuli te deuete ure tiga dne, eniga Boshyga Angela noter kfebi gredozh, inu ie dial knemu, **Corneli**.

**Dalmatin 1584:** Ta je vidil v'eni prikasni ozhitu, ob deveti uri, po dnevi, eniga Boshjiga Angela k'febi noter gredozh, ta je djal k'njemu: **Corneli**.

**Vulgate:** vidit in visu manifeste quasi hora nona diei angelum Dei introeuntem ad se et dicentem sibi **Corneli**

**Luther:** Der sahe in einem Gesichte offenbarlich / vmb die neunde stunde am tage / einen Engel Gottes zu jm eingehen / der sprach zu jm / **Corneli**.

'He saw in a vision evidently about the ninth hour of the day an angel of God coming in to him, and saying unto him, Cornelius.'

---

<sup>5</sup>The sources for the Slovenian quotations in this chapter were either the *Corpus of texts by the 16th-Century Slovenian Protestant writers* (Sl. *Korpus besedil slovenskih protestantskih piscev 16. stoletja*), available at <https://fran.si/korpus16> or the *Slovenian Bible Portal* (*Biblija. Net* n.d.), which includes several Slovenian historical and modern translations, as well as some foreign language translations of the Bible, including Luther's *Bible* (1545) and the Vulgate, from which German and Latin equivalents of Slovenian quotations are copied. English translations of the Bible texts are copied from the King James Version, also available at the *Slovenian Bible Portal*; marginal notes (marked by <>) and other non-Biblical texts were translated by the author of the chapter.

- (3) **Trubar 1577:** INu **Agrippas**<sub>NOM</sub> rezhe Hpaulufu. Tebi ie perpuřzhenu famimu fa fe gouriti. Tedai Paulus iftegne to roko, inu ie fa fe uti maffi gouuril. Lubi Kral **Agrippa**<sub>VOC</sub>, Ieft fe shtimam fa frezhniga, de fe ieft bō danas pred tebo odgouariel, od ufeh tih rizhi, fa katerih uolo fem od tih Iudou fatoshen

**Dalmatin 1584:** **AGrippas** pak je rekàl h'Paulu: Tebi je pèrpuřzhenu, de fam sa fe govorih. Natu fe je Paul sagovarjal, inu je roko istegnil: Meni je filnu lubu, lubi Krajl **Agrippa**, de fe jeft imam danas pred tabo odgovoriti, vřiga tiga, kar mene Iudje dolshé:

**Vulgate:** Agrippa vero ad Paulum ait permittitur tibi loqui pro temet ipso tunc Paulus extenta manu coepit rationem reddere de omnibus quibus accusor a Iudaeis rex **Agrippa** aestimo me beatum apud te cum sim defensurus me hodie

**Luther:** **AGrippas** aber sprach zu Paulo / Es ist dir erleubet fur dich zu reden. Da verantwortet sich Paulus / vnd recket die hand aus / Es ist mir seer lieb / lieber könig **Agrippa** / das ich mich heute fur dir verantworten sol / alles des ich von den Jüden beschuldiget werde

'Then Agrippa said unto Paul, Thou art permitted to speak for thyself. Then Paul stretched forth the hand, and answered for himself: I think myself happy, king Agrippa, because I shall answer for myself this day before thee touching all the things whereof I am accused of the Jews.'

Besides transferred nominal forms and full exonyms, names from both groups forms were frequently adapted orthographically, forming **exographs**: the most common changes were replacing Latin diphthongs and digraphs representing aspirated plosives, which are not a part of Slovenian phonological system, as well as omitting Latin (post-Classical) silent <h> and replacing Latin <c> with Slovenian <k>: e.g.,; for Lat. *Iohannes*, besides the transferred form exographs *Ioannes* and *Ioanes* are attested; similarly, for Lat. *Philippus*, we find exographs *Philipus*, *Phillipus*, *Filippus* and *Filipus* (besides equally orthographically variant exonyms *Philip*, *Fillip*, *Filip*); for the Latinized name *Caiphas* exographs *Kaifas*, *Kajfas*, *Kajphas*, *Kaiffas* are attested as well as exonyms *Kajfash*, *Kaifef*, *Kaifesh*, *Caifesh* and *Kajphesh*; for Lat. *Mathaeus* we find only exographs *Matheus*, *Mateus*, *Matteus* (besides exonyms *Matheush*, *Mateush*, *Matthevs*, *Matteush*), original Latin spelling is not attested. The exonyms, as well as exographs, show high variation in spelling and demonstrate the high influence of Latin on the orthography of proper names in the 16th-century Slovenian literary language. It should be noted

that in the entire Bible translation (Dalmatin, 1584), some of the earlier substitutions (such as *ph* > *f*) were omitted in favour of spelling closer to that in Luther's Bible, while the others, such as the omission of the silent *h*, were retained.

- (4) **Trubar 1557:** Sakai Paulus ie bil naprei ufel, mumu **Efefa** fe pelati  
**Dalmatin 1584:** Sakaj Paulus fi je bil naprej vsel mimu **Ephesa** fe zhes pelati  
**Luther:** Denn Paulus hatte beschlossen / fur **Epheso** vber zu schiffen  
'For Paul had determined to sail by Ephesus.'
- (5) **Dalmatin 1578:**<sup>6</sup> Od Keelate fo shli inu fo legli vtih Gorrah **Saffer**. Od Gur **Saffer** fo shli inu fo legli vAradi  
**Dalmatin 1584:** Od Keelate fo fhli, inu fo legli v'téh gorrah **Sapher**. Od gur **Sapher** fo fhli inu fo legli v'Aaradi (Dalmatin, *Biblia*, 1584)  
**Luther:** Von Kehelatha zogen sie aus / vnd lagerten sich im gebirge **Sapher**. Vom gebirge **Sapher** zogen sie aus / vnd lagerten sich in Harada.  
'And they went from Kehelathah, and pitched in mount Shapher. And they removed from mount Shapher, and encamped in Haradah.'

While the written source material does not enable us to definitively prove the existence of **exophones** in the 16th-century, it is probable that all forms of proper names that correspond to their Latin equivalents were not pronounced as in Latin. Jakopin (1986: 72–73) believes that masculine names orthographically ending in *-us* (*-os*), *-as*, *-es* were pronounced with a final post-alveolar fricative [ʃ], same as their exonyms where such pronunciation is orthographically confirmed, or that variant high/low pronunciation existed. Some variants of inflectional forms and word-forming morphemes seem to corroborate his thesis. For example, possessive adjectives derived from the names *Jesus* and *Kristus* 'Christ', which are predominantly spelled according to Latin and German traditions (*Christus*), are often variantly formed with the suffix *-ev*, which is usually used after palatal consonants, instead of the expected *-ov* (*Iesusev* : *Iesusov*, *Christusev* : *Christusov*), indicating the pronunciation [Jezuf] (or after the trans-syllabic consonant assimilation [Jezuf] and [Kristuf/Kriftuf]). Such pronunciation seems to be confirmed by rare variants of the name *Iesush*, *Jeshush* and the possessive adjective *Kristushev*, as well as the forms *Krishtus*, *Chrishtusev* which retain the prevalent final <s>

---

<sup>6</sup>While other Bible quotations were copied from [www.biblija.net](http://www.biblija.net), the quotation from Dalmatin's 1578 translation of the Pentateuch was taken from the *Corpus of texts by the 16th-Century Slovenian Protestant writers*.

but the change  $s > \text{ʃ}$  in the middle of the word indicates trans-syllabic consonant assimilation and therefore post-alveolar pronunciation of the final consonant. However, rare examples of possessive adjectives with the suffix *-ev* used after non-palatal consonants (e.g., *Aristobulev* from *Aristobulus*) make it impossible to confirm Jakopin's thesis.

So far, the discussion focused on nominative forms which could be transferred from Latin to Slovenian without the change of the ending. But because Slovenian is an inflection-rich language, it requires distinctive endings in the oblique cases and therefore morphological adaptation of proper names is almost obligatory to enable them to fully function in texts. Some transfers of Latin names in oblique cases are attested in the analysed material, but they are extremely limited: in some texts, Latin oblique forms were used when referring to various Bible books (6) or, in Dalmatin's Bible, oblique forms of toponyms were transferred from Latin (7, 8); in earlier Trubar's translation such toponyms were morphologically adapted.

- (6) Sato tudi Chriftnu gouori, **Marci** na 9. cap. Vše rizhi fo mogozhe, temu ker Veruie. (Juričič, *POSTILLA*, 1578)

'So says Christ in Mark's Gospel, chapter 9, all things *are* possible to him that believeth'

- (7) **Trubar 1557**: Inu od Mileta poshle on uto **Efeto**

**Dalmatin 1584**: OD Mileta pak je on pošlal v'Efhefum

**Luther**: ABer von Mileto sandte er gen **Ephesum**

**Vulgate**: a Mileto autem mittens **Ephesum**

'And from Miletus he sent to Ephesus.'

- (8) **Trubar 1557**: Ty Sholnery pag, koker ie nim bilu naloshenu, ufamo Paulufa, inu ga pelaio po nozhi uto **Antipatrído**

**Dalmatin 1584**: Ty Sholnerji fo vseli Paula, kakòr je nym bilu porozhenu, inu fo ga po nozhi pelali v'**Antipatriden**.

**Luther**: DJe Kriegsknechte / wie jnen befolhen war / namen Paulum / vnd füreten jn hin bey der nacht gen **Antipatriden**

**Vulgate**: milites ergo secundum praeceptum sibi adsumentes Paulum duxerunt per noctem in **Antipatridem**

There are also some cases of the transfer of oblique forms of rarely used names of Latin origin in non-Biblical texts, e.g. for names of pagan deities.

- (9) TY Babilonery fo imeili nih Boga **Belum**. Egiptary **Ifidem** inu **Ofyryn**. Affri **Neptunum**, Rodifary Sonce. Ta Meifta Samos **Iunonem**, Pafos **Venerem** Delfos **Apollinem**, Rym **Quirinum**, Atene **Mineruam**, Ephesus **Dianam** (Trubar, *CATEHISMVS SDVEIMA ISLAGAMA*, 1575)

‘The Babilonians had their God Baal. The Egyptians had Isis and Osiris. The Africans(?) had Neptun, the Rhodians had the Sun. The city of Samos had Iuno, Paphos had Venus, Delphi had Appolo, Rome had Quirinus, the Athens had Minerva, the Ephesus had Diana.’

In rare cases, the degree of morphological adaptation of a name depended on its referent: thus for Thomas the Apostle an exonym *Thomash* (and rarely an exograph *Tomas*) was used and inflected according to the rules of Slovenian morphology, while for (rarely mentioned) Saint Thomas of Aquinas, the Latin vocative form *Thoma* was used in nominative and locative cases.

In general, masculine proper names of Latin origin formed oblique forms from the same base as in Latin but some retained Latin nominative suffixes as a part of a Slovenian base word. For some names, such forms were more common than for others: e.g., *Ioannes* (< Lat. *Iohannes*) always retains its final -es in oblique cases (Gen *Ioannesa*, Dat *Ioannesu*); *Paulus* forms the genitive form with Latin nominative suffix much more frequently than *Petrus*: while the genitive form *Petrusa* occurs only in two works, genitive form *Paulusa* as a variant to the form *Paula* can be found in various texts, including (Bible) translations, sermons, catechisms, hymns, etc. Also, some authors used the longer forms more frequently than others.

- (10) **Trubar 1557**: Ampag ty Iudi fo gori fdrashili ene andohtie inu poshtene shene, inu te Vifshe tiga Meifta, inu fo obudili enu preganene zhes **Paulufa**<sub>ACC</sub> inu **Barnabafa**<sub>ACC</sub>, inu fo nyu uunkai is shnih krayeu ifpahnili

**Dalmatin 1584**: Iudje pak fo nadraftili te andohtlive inu poshtene Shene, inu Vifhe tiga Méfta, inu fo enu preganenje obudili zhes **Paula** inu **Barnaba**, inu fo nyu vun is fvoih pokrain pahnili.

**Luther**: Aber die Jüden bewegeten die andechtigen vnd erbarn Weiber / vnd der stad Obersten / vnd erweckten eine verfolgunge vber **Paulum** vnd **Barnaban** / vnd stiessen sie zu jren Grentzen hinaus.

**Vulgate**: Iudaei autem concitaverunt religiosas mulieres et honestas et primos civitatis et excitaverunt persecutionem in **Paulum** et **Barnaban** et eiecerunt eos de finibus suis

In some cases, both types of declension can be found in the same passage:

- (11) **Petru**<sub>NOM</sub> pak ta zhař ſtoij ſûnai pred vratmi, Nato gre ta drugi Ioger, kija ie Velikimufariu ſnan bil, inu govori ſ'vratarico, inu pelia **Petrufa**<sub>ACC</sub> notâr. [Vratarica] Pravi k'**Petru**<sub>DAT</sub>: Néřili tudi tiga Zhloveka mlaiřhieh edan? (Krelj, *Postilla Slovenska*, 1567)

'But Peter stood at the door without. Then went out that other disciple, which was known unto the high priest, and spake unto her that kept the door, and brought in Peter. Then saith the damsel that kept the door unto Peter, Art not thou also *one* of this man's disciples?'

On the other hand, genitive *Marcusa/Markusa* < *Marcus/Markus* is not attested, oblique forms of the name are only formed from Latin base form (Gen *Marca/Marka*, Dat *Marcu/Marku*).

As the base form for transferred Latin nominative forms and their Slovenian exonyms are often the same (*Petrus* > Gen *Petra*; *Peter* > Gen *Petra*), it is often impossible to determine to which nominative form oblique cases should be attributed if more than one is attested. An example of this is the name *Andreas* or *Andrej* 'Andrew'. In nominative the Latin nominal form *Andreas* is the most common (see Table 1; numbers in brackets indicate number of occurrences in the 16th-century Slovenian). The adapted nominal form is *Andrej*, which also occurs in phonologically reduced form *Andri*. In one case, a contaminated form *Andreias* appears. In oblique cases the root without suffix -as, typical for Latin and for Slovenian, is the norm, with only one occurrence of the extended base *Andreasa* in the accusative. Therefore, it is difficult to determine whether those oblique cases should be counted as occurrences of *Andreas* or *Andrej*.

One criterion for making the distinction could be the orthographic closing of the hiatus with <j> or its spelling variant <i>, but such closing of hiatus is a matter of orthographical convention and not phonological variation (it is always orthographically closed for *eu*, as can be seen in the dative and the locative, but not necessarily for *ea* and *eo*, while the pronunciation of all vocal clusters is identical – the hiatus is closed).

Also, same oblique forms appear in the same context as different nominative forms.

- (12) Filipus pride inu pouei **Andreiu**<sub>DAT</sub>, **Andreas**<sub>NOM</sub> řpet inu Filipus poueiřta Iefufu (Trubar, *TA PERVI DEIL TIGA NOVIGA TESTAMENTA*, 1557)

Filipus pride inu pouei **Andreiu**<sub>DAT</sub>, **Andreias**<sub>NOM</sub> řpet inu Filipus poueiřta Iefufu (Trubar, *TA CELI NOVI TESTAMENT*, 1581–82)

'Philip cometh and telleth Andrew: and again Andrew and Philip tell Jesus.'

Table 1: The declension of ‘Andrew’ in Latin and Slovenian.

| Latin declension |          | 16th-century Slovenian forms |                                                                  | Modern Slovenian declension |
|------------------|----------|------------------------------|------------------------------------------------------------------|-----------------------------|
| Case             | Singular | Case                         | Singular                                                         |                             |
| Nominative       | Andrĕās  | Nominative                   | Andreas (33)<br>/ Andrei<br>(10)/ Andri<br>(3) /<br>Andreias (1) | Andrej                      |
| Genitive         | Andrĕae  | Genitive                     | Andrea (12) /<br>Andreja (1)                                     | Andreja                     |
| Dative           | Andrĕae  | Dative                       | Andreiu (3) /<br>Andreju (2)                                     | Andreju                     |
| Accusative       | Andrĕam  | Accusative                   | Andrea (19)                                                      | Andreja                     |
| Ablative         | Andrĕā   | Locative                     | Andreiu                                                          | Andreju                     |
| Vocative         | Andrĕā   | Instrumental                 | Andreom (2)<br>/ Andrejom<br>(1)                                 | Andrejem                    |

- (13) Kadar je pak Iesus raven Galilejfkiga Morja hodil, je sagledal dva Brata, Simona, kateri je imenovan Petrus, inu **Andrea**<sub>ACC</sub>, njegoviga Brata, ta fta fvoje Mrĕshe v’Murje metala: Sakaj ona fta bila Ribizha <Peter inu S. **Andri**<sub>NOM</sub> k’Apostolom poklizana> (Dalmatin, *BIBLIA*, 1584)

‘And Jesus, walking by the sea of Galilee, saw two brethren, Simon called Peter, and Andrew his brother, casting a net into the sea: for they were fishers <Peter and St. Andrew called to be the Apostles>.

- (14) Kerftnik perpravi **Andrea**<sub>ACC</sub> hChriftufu, **Andreas**<sub>NOM</sub> Petra (Trubar, *HISHNA POSTILLA*, 1595)

‘The Baptist sends Andrew to Christ, Andrew brings Peter.’

Another morphological variation occurs in toponyms. As, unlike the majority of anthroponyms, they do not have a natural gender, a grammatical gender had to be assigned to them in Slovenian for them to be included in the appropriate declension and for the purposes of syntactic agreement. For several toponyms, variant gender assignment is attested, either according to the general tendencies of gender assignment for Latin words in Slovenian based on their

form (Ahačič 2011) or according to their Latin gender. In nominative, they could either retain the Latin ending or an exonym without the Latin ending or – for neuter and feminine gender – an exonym with the appropriate Slovenian ending was used: e.g., for Lat. *Corinthus*<sub>f</sub> exograph *Corintus*<sub>m/n?</sub> as well as exonyms *Corint/Korint/Korinth*<sub>m</sub> were used, as the Latin words ending in *-us* were generally assigned masculine gender when borrowed into Slovenian (Ahačič 2011: 116), but the feminine accusative form *Corinto*<sub>f</sub> is also attested, as well as locative form *Corintij* which could be interpreted as a feminine form with an added suffix *-ija* (*Corintija*<sub>NOM</sub>, *Corintiji*<sub>LOC</sub>), but as the spelling for /i/ and /j/ and their combinations was highly irregular in the 16th-century Slovenian literary language (the combination <ij> could stand for /i/, /j/, /ij/, /ji/ and /iji/ (Jelovšek 2020: 293), it is impossible to reliably determine the base word. In Trubar's works, we also find sporadic cases in the neuter gender, probably under the influence of the gender of the noun *mesto* 'town, city': *Corintu*<sub>n</sub>.

Similarly, *Jerusalem* (which was adopted from its German and not Latin form) could function as a masculine noun, as it ends in a consonant (Ahačič 2011: 116), but also as a feminine or neuter noun (both genders were attested in Latin); unlike *Korinth*, when feminine or neuter, *Jerusalem* did not acquire an appropriate case ending in nominative; the gender is determined by the verb form or the definite article used.

- (15) v'ta nebefki Ierusalem (Krelj, *POSTILLA SLOVENSKA*, 1567)

to the *A<sub>CC.SG.MASC</sub>*· heavenly *A<sub>CC.SG.MASC</sub>*· Ierusalem<sub>*A<sub>CC.SG</sub>*</sub>·  
„to the heavenly Ierusalem”

- (16) *Ierusalem pufta leshy* (Dalmatin, *BIBLIA*, 1584)

Ierusalem<sub>*NOM.SG*</sub>· bare<sub>*NOMSG.MASC*</sub>· lies  
„Ierusalem lies bare”

- (17) *terpi tu Ierufalem inu tui Folk leto framoto* (Trubar, *ENA DVHOVSKA PEISEN SVBPER TVRKE*, 1567)

suffers<sub>*3RD.SG*</sub>· the<sub>*NOM.SG.NEUT*</sub>· Ierusalem<sub>*NOM.SG*</sub>· and your people this  
shame  
„Ierusalem and your people suffers this shame”

## 5.2 Names of Germanic origin

For originally Germanic names, determining the exact type of transfer strategy is often much more difficult, as the majority of them were adopted into Slovenian from various dialects and chronolects of German, in which case the source and the target language were not only in direct geographical contact but German played the role of the prestige language in the Slovenian lands for centuries (Ahačič 2014: 17–20, Javor Briški 2012: 590–591); furthermore, it was itself undergoing intense dialectal development in the centuries before the establishment of the Slovenian literary language (cf. Šekli 2020). In the 16th century, Slovenian authors were not only acquainted with the newly established Luther's literary norm but they were also in direct contact with a variety of German dialects in the southern parts of the Holy Roman Empire. Therefore, it is difficult to determine whether a nominative form of a proper name is a true exonym, phonologically as well as orthographically distinct from the source form, or merely an exograph, orthographically modified to better reflect original (regional) German pronunciation in the period when the form entered the Slovenian language. For example, the locative form *Vshalspurgi* 'in Salzburg' could, at least theoretically, reflect Bavarian post-alveolar pronunciation of the Middle High German *s* as [š/ž], typical for the period 750–1500, and the development of Bavarian Old High German *b-* > *p-* (appearing around 770) (Šekli 2020: 10), while the form *Salzburg* could be borrowed from the later literary form of the name; as we see from the first variant, which appears in the oblique case, a morphological adaptation (addition of Slovenian case endings) was also required in majority of cases.

On the other hand, rare adaptations from non-German Germanic names, such as *Viklef* for English Wycliffe, could be confidently labelled as exographs, replacing the foreign orthography in a way to reflect original pronunciation.

The orthographical adaptations were the dropping of the original silent <h> and the replacement of the <ch> with <h>; in cases where it was not phonologically adapted to /b/, German <w> could be either retained or orthographically replaced by <vv>, <vu> or <v>: *Vuilelm* (< *Wilhelm*), *Vvolfgang/Vuolfgang* (< *Wolfgang*) alongside *Bolfgang*

The most common phonological adaptations attested in names adapted from German were the palatalization of alveolar sibilants (similar as in names of Latin origin), the de-voicing of voiced stops (or, less frequently, voicing of the aspirated voiceless stops) and the sporadic replacement of German /w/ with /b/, as well as the simplification of consonant clusters. In some cases, vowel rounding also occurred, while German rounded vowels /œ/ and /ɤ/ were commonly replaced with /e/ and /i/. For illustration, the exonyms and exographs of German

toponyms *Wittemberg* and *Augspurg* are *Wittemberg*, *Witeberg*, *Biteberg*, *Bitemberg*; *Augspurg*, *Augshpurg*, *Auspurg*, *Aushpurg*; some of less frequent, but interesting examples are *Ludvih/Ludbih* (< *Ludwig*), *Hiltepront* (< *Hildebrand*) and *Tonaberd* (< *Donauwörth*).

For some anthroponyms, forms influenced by the Latinized forms of the name can also be found in the analysed material: *Ludovik/Ludovic* (< Lat. *Ludovicus*, Ger. *Ludwig*) 'Louis', *Bernardus* (< Lat. *Bernardus*, Ger. *Bernhard*) alongside *Bernhard* and a contaminated form *Bernhardus*. For German *Karl* 'Charles' only the form *Karol* from Latinized *Carolus* is attested.

In cases where German and Slovenian exonyms existed for an anthroponym from the shared onomastic heritage (cf. Raukko 2017: 98), Slovenian exonym was predominantly used: *Vyd Djetrih* < Ger. *Veit Dietrich*, *Iansh Brencius* < Ger. *Johannes Brentz*, Latinized *Brentius*.

In rare cases, a partial or a full translation of German proper names is also attested. The name *Felix od Barteperg* < Ger. *Felix von Wartenberg* 'lit. Felix from Wartenberg' could be understood as an example of a partial translation, in which only the nobility particle *von* was translated with the Slovenian preposition *od*, while the family name was adapted phonologically, but not morphologically (the full translation would require the family name in the genitive case after the preposition *od*).

For rare frequently used and historically important toponyms, unrelated Slovenian exonyms existed (and still exist), such as *Dunaj* for *Wien* 'Vienna', but an exonym *Kostnica* is also attested for the German town of Konstanz which has no direct historical connection to Slovenian lands and was probably a result of morphological re-analysis in Slovenian (cf. Nicolaisen 1996: 553), in which German name was reinterpreted as a similarly sounding but semantically unrelated meaningful Slovenian word *kostnica* 'ossuary'.

### 5.3 Names of Semitic origin

Names of Semitic origin, especially from the Old Testament, which did not become a part of the Christian names inventory and also did not acquire typical Latin endings in the Latin or German translations of the Bible, nevertheless show high influence of the intermediary languages – Latin and German – in 16th-century Slovenian texts, as direct transcriptions from Luther's Bible or the Vulgate were even more frequent than in the other two groups of proper names. Therefore, the attested variation is often the result of the similar variation in the source (inconsistent adaptation of Semitic names is typical for both Luther's Bible and the Vulgate; cf. Krašovec (2006: 352ff).

**Dalmatin 1584:** Osnanite v'Egypti, povéjite tu v'**Migdali**

**Luther:** Verkündigt in Egypten / vnd sagts an zu **Migdal**

**Vulgate:** adnuntiate Aegypto et auditum facite **Magdolo**

'Declare ye in Egypt, and publish in Migdol.'

**Dalmatin 1584:** Od Etame fo fhli, inu fo oftali v'ti dolini Ahiroth, katera leshy pruti BaalZephtonu, inu fo legli pruti **Migdoli**.

**Luther:** Von Etham zogen sie aus / vnd blieben im grund Hahiroth / welchs ligt gegen Baal Zephton / vnd lagerten sich gegen **Migdol**

**Vulgate:** inde egressi venerunt contra Phiahiroth quae respicit Beelsephon et castrametati sunt ante **Magdolum**

'And they removed from Etham, and turned again unto Pihahiroth, which is before Baalzephon: and they pitched before Migdol.'

**Dalmatin 1584:** Taku pravi **Sanherib**, Affyrerfki Krajl

**Luther:** so spricht **Sanherib** der könig zu Assur

**Vulgate:** haec dicit **Sennacherib** rex Assyriorum

'Thus saith Sennacherib king of Assyria'

**Dalmatin 1584:** Koku je **Senaherib**, inu ftu inu pet inu offemdeffet taushent Mosh shnym red, v'eni nozhi bilu konzhanih.

**Luther:** Da dich die Boten des königes **Sennaherib** lesteren / schicktest du einen Engel / der schlug tod / hundert vnd fünff vnd achzig tausent man.

**Vulgate:** et sub **Sennacherim** centum octoginta quinque milia ut perierunt

'how they were delivered, when under Sennacherib an hundred fourscore and five thousand perished'

As can be seen from the examples, Dalmatin made similar orthographical adaptations for Semitic names as for Latin(ized) names (ommiting the silent *h*, replacing the aspirated stops etc.).

Besides frequent vowel variation, we can also find some typical variation in consonants, such as final *-n* / *-m*, which was not necessarily a result of a variation in source texts: e.g. *Absalom* – *Absalon* – *Absolom* – *Absolon*.

Some variation in Semitic proper names in the 16th-century Slovenian arose because the authors used different source translations of the Bible. While Dalmatin in his translation of the Old Testament closely followed Luther (Krašovec

2010: 3), other authors, when writing of persons from the Old Testament in their non-Biblical texts, often used the name forms from the Vulgate. The best example for this is the name of the king Nebuchadnezzar with thirteen variants (see Table 2). This name also illustrates other typical variations in Semitic proper names, such as the trans-syllabic vowel and consonant assimilation, as well as the variation between /z/ (spelled with <s>), when based on the Latin form, and /c/ (spelled with <z>), when based on the German form, and the variation between the final *-er* and *-ar* which is not limited to proper names but is also typical for common nouns borrowed from German into Slovenian at the period (Legan Ravnikar 2012: 141). On the other hand, two occurrences of *NebudNezar* with a missing syllable are probably a mistake.

Table 2: Forms of the name *Nebuchadnezzar* in the 16th-century Slovenian.

| Via German: <i>NebucadNezar</i> | Via Latin: <i>Nabuchodonosor</i> |
|---------------------------------|----------------------------------|
| <i>Nabukadnecar</i>             | <i>Nabugdnosor</i>               |
| <i>Nabukadnecer</i>             | <i>Nabuhodonosar</i>             |
| <i>Nabukadrekar</i>             | <i>Nabuhodonosor</i>             |
| <i>Nebudnecar</i>               | <i>Nebuhodnosor</i>              |
| <i>Nebukadnecar</i>             | <i>Nubugdonosor</i>              |
| <i>Nebukadnecer</i>             |                                  |
| <i>Nebukadrekar</i>             |                                  |
| <i>Nebukadrecer</i>             |                                  |

## 6 Conclusion

The analysis of the methods of incorporation of proper names into the 16th-century Slovenian literary language has shown some differences based on the origin of the name: Latin and Latinized names were more often transferred than Germanic names, who underwent more extensive phonological adaptation that was reflected also in their orthography, or their orthography was changes to more accurately reflect their original phonological form in Slovenian; as only written sources can be studied, it is hard to determine whether Latin names also underwent phonological adaptation which was not reflected in their orthographical form, due to the prevailing influence of Latin orthography, as Latin was prized

above all other languages in intellectual circles (Ahačič 2014: 23). For Semitic names, high influence of source translations is attested.

The analysis has also shown the high level of prestige the Bible text had among the translators, who predominantly kept the original forms of proper names in their Bible translations despite using existing exonyms in their other texts. The changes made in the entire translation of the Bible in 1584, which decreased the level of (orthographical and morphological) adaptation of some proper names compared to previous translations, and brought them again closer to German and Latin forms, further corroborate the prestigious status of the source translations, especially Luther's Bible 1545.

In some aspects, the findings of the analysis agree with the findings of the previous research on linguistic borrowing of semantic words (e.g. Striedter-Temps 1963 and Legan Ravnikar 2012 for German; Ahačič 2011 for Latin); proper names underwent some similar orthographical and phonological adaptations as semantic words, but due to their specific features, they were much more frequently transferred (even if in a form of an exograph) from the source language than semantic words; they also exhibit some special adaptations in terms of morphology and a higher degree of variation. The findings of the study therefore provide further insight into issues of language borrowing and adaptation of loanwords, enriching the findings of Slovenian contact linguistics by encompassing heretofore rarely considered material.

## Acknowledgements

The article was written in the context of the research programme IMAGE – WORD – KNOWLEDGE. The Transmission and Transformation of Ideas on the Territory Between the Eastern Alps and the North Adriatic 1400–1800 (P6-0437). The author acknowledges the project was financially supported by the Slovenian Research Agency from the state budget.

## References

Ahačič, Kozma. 2011. Nekaj vidikov umeščanja slovenščine v evropski jezikovni prostor 16. stoletja: Od uporabnosti slovenskega jezika do latinskih tujk. In Kozma Ahačič & Petra Testen (eds.), *Jeziki, identitete, pripadnosti med središči in obrobji: V počastitev 500. obletnice rojstva Primoža Trubarja*, 109–122. Ljubljana: Založba ZRC, Inštitut za kulturno zgodovino ZRC SAZU.

- Ahačič, Kozma. 2014. *The history of linguistic thought and language use in 16th century Slovenia*. Frankfurt am Main: Peter Lang.
- Biblija. Net*. N.d. <https://www.biblija.net/biblija.cgi?Bible=Bible&l=sl>.
- Casanova, Emili. 2012. On the adaptation of names of Romance origin (Spanish, French, Occitan, Italian) into Valencian Catalan. In Oliviu Felecan (ed.), *Name and naming: Synchronic and diachronic perspectives*. 65–73. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Cooper, H. R. 1985. Primoz Trubar and Slovene literature of the sixteenth century. *Slovene Studies* 7(1). 35–50.
- Gordon, Octavian. 2012. Denominational differences in the Romanian rendering of Greek and Latin-origin proper names in ecclesiastical literature: A translation theory perspective. In Oliviu Felecan (ed.), *Name and naming: Synchronic and diachronic perspectives*. 2–9. Newcastle upon Tyne: Cambridge Scholars Pub.
- Hermans, Theo. 2015. On translating proper names, with reference to De Witte and Max Havelaar. In Michael J. Wintle & Paul Vincent (eds.), *Modern Dutch studies: Essays in honour of Professor Peter King on the occasion of his retirement* (Bloomsbury academic collections. History: European history.), 11–24. London, UK, New York, NY: Bloomsbury Academic.
- Jakopin, Franc. 1986. Jezikovna in pisna adaptacija imen v slovenskih besedilih 16. stoletja. In Marjan Dolgan & Darko Dolinar (eds.), *Družbena in kulturna podoba slovenske reformacije*, 69–75. Ljubljana: Slovenska akademija znanosti in umetnosti.
- Javor Briški, Marija. 2012. Slowenien. In Albrecht Greule, Jörg Meier, Arne Ziegler, Melanie Glantschnig, Jakob Reichsöllner & Elisabeth Scherr (eds.), *Kanzleisprachenforschung: Ein internationales Handbuch*, 589–598. Berlin: De Gruyter.
- Jelovšek, Alenka. 2020. Prikaz pisne in glasovne variantnosti iztočnic v SSKJ16. In Marko Jesenšek (ed.), *Slovensko jezikoslovje, književnost in poučevanje slovenščine: Slovar slovenskega knjižnega jezika 16. stoletja. Veliki madžarsko-slovenski spletni slovar*, 287–308. Maribor: University of Maribor Press. DOI: 10.18690/978-961-286-365-4.
- Krašovec, Jože. 2006. Proper names in translations of the Bible and interdependence of the use of their forms. In Jože Krašovec, Majda Merše & Hans Rothe (eds.), *Matthäus-Evangelium (1555), Paulus, Römerbrief (1560), Paulus-Briefe (1561, 1567), Psalter (1566), Neues Testament (1581-1582), Jesus Sirach (1575), Pentateuch (1578), Proverbia (1580). Kommentare* (Biblia slavica Serie 4, Südslavische Bibeln 3,2), 327–368. Paderborn: Schöningh.

- Krašovec, Jože. 2010. *The transformation of biblical proper names* (T & T Clark library of biblical studies). New York, London: T & T Clark.
- Langendonck, Willy van. 2007. *Theory and typology of proper names* (Trends in linguistics 168). Berlin, New York: Mouton de Gruyter.
- Legan Ravnikar, Andreja. 2012. Značilnosti sprejemanja besed iz nemščine v knjižno slovenščino 16. stoletja. *Jezikoslovni zapiski* 18(2). 131–148.
- Magocsi, Paul R. & Geoffrey J. Matthews. 1993. *Historical atlas of East Central Europe*. Toronto: University of Toronto Press.
- Mikhailov, Nikolai. 1998. *Frühslowenische Sprachdenkmäler: Die handschriftliche periode der slowenischen Sprache (XIV. Jh bis 1550)*. Amsterdam, Atlanta (GA): Rodopi.
- Nicolaisen, Wilhelm F. H. 1996. Language contact and onomastics. In Hans Goebel, Peter H. Nelde, Zdeněk Starý & Wolfgang Wölck (eds.), *Kontaktlinguistik: Ein internationales Handbuch zeitgenössischer Forschung*, vol. 1, 549–554. Berlin, New York: Walter de Gruyter.
- Pogorelec, Breda. 2011. *Zgodovina slovenskega knjižnega jezika: Jezikoslovni spisi i*. Kozma Ahačič (ed.). Ljubljana: Založba ZRC ; ZRC SAZU.
- Raukko, Jarno. 2017. Names in contact: Linguistic and social factors affecting exonyms and translated names. In Terhi Ainiala & Jan-Ola Östman (eds.), *Socio-onomastics: The pragmatics of names*, 94–125. Amsterdam: John Benjamins Publishing Company. DOI: 10.1075/pbns.275.05rau.
- Schlücker, Barbara & Tanja Ackermann. 2017. The morphosyntax of proper names: An overview. *Folia Linguistica* 51(2). 309–339. DOI: 10.1515/flin-2017-0011. (23 April, 2024).
- Šekli, Matej. 2020. Relativna in absolutna kronologija (bavarsko)staro- in -srednjevisokonemških izposojenk v slovenščini. *Jezikoslovni zapiski* 26(1). 7–25. DOI: 10.3986/Jz.26.1.1. <https://ojs.zrc-sazu.si/jz/article/view/8557> (23 April, 2024).
- Striedter-Temps, Hildegard. 1963. *Deutsche Lehnwörter im Slovenischen*. Berlin: Osteuropa-Institut.
- Vermes, Albert Péter. 2003. Proper names in translation: An explanatory attempt. *Across Languages and Cultures* 4(1). 89–108.

## Chapitre 9

# Dictionnaires manuscrits dans l'histoire de la lexicographie croate : Des recueils de mots aux trésors linguistiques et culturels

Ivana Franić

University of Zagreb, Croatia

The paper discusses the status of manuscript dictionaries created in Dubrovnik from the end of the 16th to the beginning of the 19th century and their role in the history of the Croatian lexicography. Firstly, I provide a brief overview of these dictionaries, with a particular attention given to the *Dictionarium Latino-Illiricum* (1715–1716), a bilingual dictionary, which is a combination of an explanatory and a translational general dictionary with encyclopaedic elements (FRANIĆ 2014). With this analysis, we want to show that manuscript dictionaries, although they do not have the communication potential of published (or printed) dictionaries represent an important link in the unbroken chain of lexicographic influences. Only a thorough analysis of manuscript dictionaries can lead to a better evaluation of the overall lexicographic production of individual cultures (DUBOIS 1970), both vertically within historical lexicographic movements in a particular linguistic and cultural community and horizontally at the level of mutual influences of European cultures and languages. Only if manuscript dictionaries stand alongside the published ones will it be possible to better highlight the mutual influences of models that exist in lexicography (LANDAU 2001, REY 2008, SAMARDŽIJA 2013), so each dictionary can be considered as a sum of dictionaries that precede it and a model for dictionaries that follow.



## 1 Introduction

Ouvrages fondamentaux créés par le travail lexicographique, les dictionnaires sont traditionnellement considérés comme des livres qui « contiennent une liste de mots d'une langue avec des définitions de leur sens ou avec des équivalents dans une langue étrangère » (SAMARDŽIJA 2019 : 10). Grâce au développement du numérique et compte tenu des besoins de plus en plus croissants de la communication entre langues et cultures, différents enjeux méthodologiques se posent lors de la confection de dictionnaires. C'est pour cette raison que, de nos jours, les dictionnaires sont plutôt conçus comme

une collection d'éléments lexicaux présentés par certains médias, dans lesquels des informations sont données et qui doivent être organisées de manière à ce que l'utilisateur y ait un accès rapide et facile. (SAMARDŽIJA 2019 : 11, notre traduction)

Jean (DUBOIS 1970 : 44) Jean Dubois parle du dictionnaire comme d'un « trésor » dans lequel se trouvent des mots inconnus et avec lequel chaque individu « mesure » son trésor linguistique et le compare au trésor de sa communauté linguistique ; néanmoins, bien que le dictionnaire ne soit pas un moyen d'expression, il représente la norme ou la mesure de la richesse ou de la pauvreté d'un énoncé.

Il existe différentes typologies de dictionnaires, plusieurs auteurs abordent les critères de leur différenciation et leurs spécificités dont nous ne soulignons que ZGUSTA (1971), puis LANDAU (2001), DUBOIS (1970) et surtout REY (2008) et PRUVOST (2021). Plus récemment dans le contexte croate, il convient de mentionner l'ouvrage de SAMARDŽIJA (2019).

Un dictionnaire bilingue représente un type particulier de dictionnaire dans lequel les unités lexicales de deux différents systèmes linguistiques sont juxtaposées :

Les dictionnaires bilingues se contentent le plus souvent de donner le terme équivalent dans une seconde langue, ce qui implique que l'utilisateur ait une certaine connaissance du système lexical de celle-ci. (COLLIGNON & GLATIGNY 1978 : 32)

Selon Landau, le dictionnaire bilingue comprend deux fonctions :

for comprehension, as in reading, of the source language, by a person who knows the source language ; or as an aid in expression, as in writing, of the

target language, by a person who knows the target language. (LANDAU 2001 : 9)

Le même auteur souligne le fait que parler de l'étendue d'un dictionnaire n'est pas chose facile, le lexique étant une catégorie ouverte. (LANDAU 2001 : 7)

Les dictionnaires servent d'habitude de « pont interculturel » (REY 2007), ils sont des moyens précieux de communication et de compréhension interculturelles (WIERZBICKA 1988 : selon TALLARICO 2013, MURANO 2013, VAXELAIRE 2005, TALLARICO 2013). Les dictionnaires peuvent également contenir des données encyclopédiques hautement culturelles et relèvent de l'histoire culturelle (MEŠTROVIĆ 1994) ou de la *culture partagée* (GALISSON 1988). Les dictionnaires fournissent des images linguistiques qui témoignent de l'essence linguistique de chacune des deux langues (MEŠTROVIĆ 1994) et des représentations des deux langues et deux cultures. Le dictionnaire a alors aussi une fonction symbolique car il devient :

instrument de reconnaissance identitaire et égalitaire qui témoigne de la culture de l'Autre et amène à le connaître et à le reconnaître. (FREY 2007 : 209)

Les dictionnaires bilingues ont aussi la particularité d'aborder inévitablement le champ des spécificités culturelles : ils regroupent bien souvent des notions propres à certaines cultures et qui n'ont pas d'équivalents univoques dans la langue cible (coutumes, croyances, création artistique, artisanat, alimentation, habitat, etc.). Cela inclut une compréhension très variable et culturellement colorée de certaines parties de la réalité extralinguistique (faune, flore) (REY 1991). C'est la finalité du dictionnaire qui détermine aussi bien son contenu que sa structure, comme le prétendent (COLLIGNON & GLATIGNY 1978 : 12). Gabrić-Bagarić soutient que dans le cas du dictionnaire conçu par un auteur, qui sélectionne les entrées indépendamment du corpus, la portée du dictionnaire est conditionnée par le but et le mode de création (GABRIĆ BAGARIĆ 2010 : 19).

## 2 Début de l'activité lexicographique

Les débuts de l'activité lexicographique sont généralement liés à l'histoire ancienne, d'où proviennent des listes plus ou moins longues de mots moins connus (mots obsolètes, dialectaux ou étrangers), ou glossaires. (SAMARDŽIJA 2019 : 9). Le regain d'intérêt tant pour les langues classiques que pour les langues vivantes

grandit à la Renaissance, ce qui donne un nouvel élan à la création d'œuvres similaires. L'invention de l'imprimerie marque sans doute un tournant important quant à l'élaboration de ces « listes de mots », si bien que les dictionnaires – d'abord mis en annexe puis volumes indépendants ou autonomes – acquièrent peu à peu l'attribut de livres prestigieux de l'époque. Ainsi, depuis le 15<sup>ème</sup> siècle, de nombreux dictionnaires de langues classiques et modernes ont été publiés sous des titres différents.

A la différence du dictionnaire élaboré avant la période de l'imprimerie, comprenant pour la plupart des explications de mots inconnus ou peu connus, la tâche particulière du dictionnaire à l'époque de la Renaissance, est de relier les contextes culturels, de permettre la communication entre les individus et surtout de rapprocher les oeuvres littéraires aux lecteurs (SAMARDŽIJA 2019). Un siècle plus tard, la Réforme catholique va radicalement changer la pratique lexicographique ainsi que la finalité des dictionnaires, qui ont depuis lors une part active en tant que supports au service pastoral. (KATIČIĆ 1981 : 42) C'est ainsi que les premiers manuels, dictionnaires et grammaires sont confectionnés justement pendant cette période. En ce temps-là, c'est par rapport au latin que les langues « se mesurent », l'objectif étant de répondre, en premier lieu, aux besoins éducatifs des élèves et des missionnaires ainsi que de tous ceux qui ont besoin d'une bonne maîtrise du latin. De fait, le latin représentait à l'époque la « langue littéraire de l'érudition baroque » (KATIČIĆ 2013 : 13), toute cette époque étant, selon Katičić, la période de l'« érudition vivante » où naissent l'historiographie critique, la lexicographie et la grammaire, l'époque de l'émergence de l'ensemble des sciences humaines (KATIČIĆ 2013).<sup>1</sup> Le latin était à la fois la langue de la liturgie, mais aussi des documents officiels, de la science et de l'alphabétisation, tandis que le droit canonique, partout reconnu, influençait fortement les codes civils (KRASIĆ 2009 : 19). Les grammaires et les dictionnaires représentaient donc, à cette époque-là, des ouvrages indispensables, une sorte d'outil linguistique qui aide à la maîtrise du latin.

En ce sens, un bon dictionnaire bilingue, comportant de bons équivalents est particulièrement important : on y répertorie et explique les mots latins. Ce répertoire de mots latins doit aider les utilisateurs à comprendre les mots et constructions latins. Si l'on ajoute à tout cela l'activité très fructueuse de nombreuses académies qui recensent les trésors lexicaux de nombreuses langues depuis la

---

<sup>1</sup>Les activités de la Compagnie de Jésus (S. I.) portaient à l'époque, entre autres, sur la rédaction et l'élaboration de grammaires et manuels d'apprentissage du latin en combinaison avec les langues de différentes populations (vernaculaires), entre autres le croate, afin de répandre son enseignement parmi le peuple.

seconde moitié du 17<sup>ème</sup> siècle, on peut dire que la lexicographie acquiert alors le statut d'une véritable discipline.

A cette époque, une attention particulière est portée aux grands dictionnaires nationaux qui revêtent une grande importance pour l'histoire linguistique et culturelle en général. Comme le souligne Matasović, les résultats les plus significatifs des efforts de normalisation des langues en Europe au 17<sup>ème</sup> et au 18<sup>ème</sup> siècle sont les dictionnaires des académies nationales : *Academia della Crusca* (1611), *Dictionnaire de l'Académie française* (1694), de l'Académie espagnole (1726–1739) et de l'Académie russe (1789–1794). (MATASOVIĆ 2013 : 514)

En ce qui concerne le rapport entre deux langues en combinaison lexicographique, très souvent dans les dictionnaires ayant la composante latine, on observe un certain déséquilibre. Plus précisément, les équivalents proposés dans d'autres langues ne sont pas toujours en mesure d'expliquer de manière simple ou en un seul mot (un synonyme ou une suite de synonymes) la signification du mot latin figurant dans l'entrée.

Par conséquent, dans ce type de dictionnaire (bilingues, mais souvent trilingues), la ou les langues cibles sont vues à travers le miroir du latin (FURNO 2003 : 34).

C'est ainsi que les lexicographes de l'époque, guidés par le latin comme langue de départ, mais aussi la langue prestigieuse autour de laquelle tout était mesuré, et en même temps désireux de préserver l'esprit et l'expression de la langue vernaculaire, ajustent soigneusement la langue cible au latin afin d'expliquer les mots latins de la manière la plus fiable possible, en fonction des besoins de la description lexicographique et de la finalité du dictionnaire. (FURNO 2003 : 28–29).

Outre les grands dictionnaires imprimés bilingues et trilingues, les ouvrages lexicographiques qui sont alors créés sont très souvent manuscrits.

Dans ce qui suit, nous allons d'abord identifier quelques particularités des œuvres lexicographiques manuscrites et examiner de plus près quelques dictionnaires manuscrits créés par des auteurs qui appartiennent à ce que nous désignerons de « cercle lexicographique de Dubrovnik » (cf. § 4).

### **3 Dictionnaires manuscrits au sein de la lexicographie : Leurs propriétés et leur(s) destin(s)**

Les dictionnaires manuscrits constituent des maillons importants dans l'histoire de la lexicographie. Bien que de par leur structure, leur forme, leur méthodologie et leur objectif ils diffèrent largement des dictionnaires imprimés, leur étude est

indispensable pour une présentation complète des réalisations lexicographiques dans un contexte linguistique et historico-culturel.

Le destin d'un dictionnaire fut le plus souvent étroitement lié, au cours de l'histoire, au sort de son auteur, ou comme le soutient REY (2008 : 118), à des modèles analogues des dictionnaires antérieurs et notamment à l'idiolecte du lexicographe. En particulier, dans les dictionnaires manuscrits plus petits, les lexicographes s'appuient souvent sur leur intuition linguistique et répertorient des éléments dialectaux ainsi que des éléments de la langue parlée. Bien entendu, de telles œuvres refléteront également la créativité de l'auteur puisque l'auteur, en l'absence de bons équivalents en un mot, essaie souvent de trouver la meilleure expression (tournure ou explication) dans la langue cible.

On peut donc dire que le rôle du lexicographe dans la création d'un dictionnaire est très souvent crucial. En fait, l'auteur du dictionnaire « tend à l'anonymat » et s'identifie en même temps avec l'objet de sa création, à son dictionnaire (ainsi les dictionnaires portent souvent le nom de leurs auteurs), mais aussi à la langue décrite dans le dictionnaire (DUBOIS (1970 : 41).

Les lexicographes ne sont pas seulement des locuteurs de cette langue et donc « possesseurs » de ce « trésor commun », ils sont aussi des médiateurs de ce trésor en direction des lecteurs (utilisateurs). Et comme les lexicographes sont guidés presque régulièrement par les modèles de dictionnaires de leurs précurseurs, on peut constater que leur tâche consiste non seulement à transmettre ce trésor à leurs lecteurs, mais également de le laisser comme un héritage et une marque authentique dans le temps. C'est ainsi que le dictionnaire devient un véritable pont qui relie des générations de locuteurs, séparées par des années ou des décennies, parfois par des siècles entiers, mais aussi réunies par l'identité linguistique et par l'utilisation du même idiome.

Reprenons ici le constat de Collignon et Glatigny sur le rôle décisif de la destination du dictionnaire qui détermine son contenu ainsi que son organisation (cf. Introduction). C'est précisément une partie de la réponse à la question de savoir pourquoi, d'une part, certains dictionnaires ont vu le jour sous forme imprimée et pourquoi, d'autre part, certains dictionnaires sont restés en manuscrit et tombés dans l'oubli. Nous essaierons de fournir au moins une partie de réponses à partir de données dont nous disposons sur les dictionnaires manuscrits que nous avons étudiés.

A l'instar de tous les manuscrits, les dictionnaires manuscrits sont des œuvres écrites par la main de leurs auteurs, lexicographes. Tout dictionnaire manuscrit est donc unique : il représente un texte authentique (ou le matériel lexical), ce texte n'est pas multiplié par des procédés de reproduction mécanique. En outre, le manuscrit, y compris le dictionnaire, a son propre format, nombre de lignes

par page (souvent variable), style d'écriture, qui peut être rempli de caractères spéciaux d'auteur, d'abréviations, de corrections, d'interpolations, de notes en marge, etc. Il n'y a donc pas deux dictionnaires manuscrits identiques, tout dictionnaire manuscrit représente un recueil original : de par la structure du répertoire, des déterminants, des types de définition, des manières de proposer l'équivalent. Cependant, cet « autographe authentique » témoigne des efforts des individus pour composer eux-mêmes de telles œuvres, en s'appuyant sur les modèles existants, souhaitant laisser leur propre trace ou interprétation des mots d'une autre langue.

Un dictionnaire peut rester sous forme de manuscrit pour différentes raisons : la finalité du dictionnaire, le caractère inachevé du dictionnaire, son caractère pratique ou tout simplement le fait que le manuscrit d'un dictionnaire n'est souvent que la transcription d'un autre dictionnaire (que l'auteur a consulté). Par ailleurs, le contexte social et économique, les circonstances culturelles et historiques ont dans la plupart des cas déterminé le sort du dictionnaire. Comme le soutient Rey :

la décision de publier un dictionnaire relève plus de facteurs économiques ou politiques que d'un noble souci didactique ou scientifique REY (2008 : 14).

Il convient de mentionner ici le fait que de nombreux dictionnaires manuscrits et leurs différentes rédactions ont été perdus, puisqu'ils ont été écrits en un seul exemplaire. En tant que manuscrits, ils n'ont été utilisés que par un nombre bien restreint de lecteurs et leur utilisation était donc limitée.

Malgré leur faible potentiel de communication et leur réception assez limitée, nous pouvons supposer que ces répertoires manuscrits ont servi d'appui aux autres lexicographes, dans leurs recherches de modèle convenable, comme le souligne REY (2008 : 118) :

La lexicographie décrit habituellement un système lexical qu'elle définit a priori et qui correspond à un modèle élaboré selon des règles non formulées, subjectives et souvent variables : intuition de l'unité lexicale, de son importance fonctionnelle, de sa valeur culturelle (dans le cas du purisme normatif). Ce modèle est défini en partie par référence à l'idiolecte du lexicographe, en partie par référence aux modèles analogues (les dictionnaires précédents) et en partie par référence à une sélection du corpus indéfini des phrases, c'est à dire un ensemble de faits de discours.

En fait, aucun dictionnaire n'aurait pu être créé *ex nihilo*, sans modèles et précurseurs, comme le prétend Landau :

The earliest English lexicographers by and large copied the definitions of their predecessors. (LANDAU 2001 : 190)

C'est là que réside la plus grande valeur des dictionnaires manuscrits plus petits ou plus grands, plus pauvres ou plus riches. Sans doute est-il parfois difficile de prouver la chaîne des influences lexicographiques, notamment celles qui lient un dictionnaire plus ancien et un autre plus récent. Cependant, il existe de nombreux exemples qui nous permettent de déterminer assez facilement ces influences (dans FRANIĆ (2014) nous avons expliqué plus en détail les rapports qui existent à l'intérieur d'un corpus qui a précédé la confection de *Dictionarium Latino-Illiricum*, cf. § 4.2., ainsi que dans les dictionnaires qui ont été rédigés postérieurement à celui-ci).

En conséquence, si les œuvres manuscrites n'avaient pas un potentiel de communication égal à celui des œuvres imprimées – n'étant pas largement répandues ni connues – nous trouvons que, dans les circonstances de l'époque, elles étaient suffisamment disponibles aux lexicographes qui auraient pu les atteindre au moment où ils en avaient besoin. Juste titre que les grands dictionnaires (imprimés) doivent leur naissance, au moins partiellement, entre autres, à un nombre d'œuvres manuscrites.

C'est ainsi que dans l'histoire de la lexicographie croate, de nombreux dictionnaires manuscrits, de portée plus ou moins grande, pour la plupart bilingues, mais aussi trilingues, ont marqué certaines périodes, comme l'estime Samardžija en ajoutant que ces dictionnaires représentent une source précieuse pour l'histoire de la langue littéraire croate :

En plus d'être un riche trésor du lexique croate, les dictionnaires croates de l'époque, tant ceux publiés que ceux qui nous sont parvenus en manuscrits à cause des circonstances (non) favorables, sont des sources précieuses pour l'histoire de la langue littéraire croate [...]. (SAMARDŽIJA (2013 : 453), notre traduction)

Au cours des décennies passées, de nombreux auteurs ont réalisés des recherches dans ce domaine (GABRIĆ BAGARIĆ 1996, 1998, 2000, 2002, VAJS 2003, VONČINA 1992, 2003, SIRONIĆ BONEFAČIĆ 1992, PUTANEC 1991, 2000, MEŠTROVIĆ & VAJS 1995, HORVAT 1990). Il convient de souligner tout particulièrement le *Lexicon Latino-Illyricum* (1700–1709) de Pavao Ritter Vitezović, trois dictionnaires de Ivan Tanzlingher Zanotti : *Talijansko-ilirsko-latinski rječnik* (1699), *Vocabolario di tre nobilissimi linguaggi, italiano, illirico e latino*<sup>2</sup> (1699) et *Vocabolario italiano ed*

<sup>2</sup>Le dictionnaire a été numérisé en 2008 (avec une autre édition en 2020) et peut être consulté sur le site [https://tanzlingher.disll.unipd.it/xml/index.php?\\_lang=it](https://tanzlingher.disll.unipd.it/xml/index.php?_lang=it).

*illirico* (1699) ainsi que *Dictionarium latino-illyricum et germanicum* (1772–1779) de Adam Patačić. Etant donné que Vitezović et Tanzlingher Zanotti avaient une conception bien claire de la langue littéraire croate fondée sur trois dialectes : shtokavien, kaïkavien et ikavien (SAMARDŽIJA 2019 : 47–48), leurs dictionnaires s'avèrent être incontournables pour donner un aperçu tout d'abord d'un corpus intégral lexicographique et lexical de la langue croate et puis pour faire un bilan des contacts linguistiques du croate, à l'époque, avec les langues de communication, aussi bien générale que scientifique, les langues de l'art, avant tout le latin et l'italien.

Nous nous proposons d'examiner, dans la section suivante, un corpus lexicographique manuscrit qui d'ailleurs rejoint les grands dictionnaires imprimés de l'époque ayant pour base l'idiome de Dubrovnik – *Blago jezika slovinskoga* (1649–1651) de Jakov Mikalja et *Dizionario italiano-latino-illirico* (1728) de Ardelio Della Bella. Ce corpus manuscrit, par sa portée et son étendue s'inscrit par-là dans le patrimoine linguistique et culturel croate et européen.

En conséquence, nous ne décrivons pas ici les grands dictionnaires imprimés du cercle lexicographique de Dubrovnik car ils sont déjà bien décrits dans de nombreuses études menées dans les décennies passées. En revanche, nous nous pencherons sur les dictionnaires manuscrits moins connus ou presque inconnus qui ont sans doute leur part dans la production lexicographique intégrale de Dubrovnik.

#### 4 Corpus lexicographique manuscrit au sein du cercle lexicographique de Dubrovnik

Le statut de Dubrovnik comme point de rencontre d'événements culturels et littéraires au cours des siècles, le prestige de l'idiome de Dubrovnik, mais aussi le besoin constant d'enseignement, notamment l'enseignement du latin et l'interprétation de ses spécificités ont favorablement influencé l'intérêt pour la réalisation de divers dictionnaires manuscrits. Ces œuvres provenaient principalement des auteurs qui maîtrisaient bien le latin, des prêtres et des religieux de l'époque. Dans les couvents de Dubrovnik, cette activité était souvent accessoire, parallèle à la transcription et à l'interprétation d'œuvres religieuses, à la prédication, etc. De nombreux dictionnaires manuscrits, soit autonomes soit figurant dans l'annexe d'un livre, furent rédigés dans les bibliothèques de couvents de Dubrovnik, celui des Frères mineurs, des Dominicains et de Collegium Ragusinum Jésuite.

Ce corpus manuscrit témoigne du fait que les sept décennies qui divisent le dictionnaire *Blago jezika slovinskoga* (1649–1651) de Mikalja et le *Dizionario italiano-*

*latino-illirico* (1728) de Ardelio Della Bella abritent une activité lexicographique très vive et surtout originale. Nous considérons que tous les dictionnaires, imprimés ou manuscrits, qu'ils furent rédigés à Dubrovnik ou que l'une de ses composantes fut l'idiome de Dubrovnik, forment le cercle lexicographique (bicentenaire) de Dubrovnik.

Les auteurs d'œuvres lexicographiques plus importantes (publiés et imprimés) de l'époque ont sans doute puisé dans des recueils manuscrits, plus ou moins grands, dans différentes paires de langues ou combinaisons linguistiques, créés au sein du cercle lexicographique de Dubrovnik. Dès lors, si les dictionnaires manuscrits n'avaient pas le potentiel de communication des ouvrages imprimés, on ne peut pas leur nier une part quasi certaine dans l'émergence de prestigieux grands dictionnaires imprimés.

Outre l'idiome de Dubrovnik en tant que langue au patrimoine linguistique et culturel très riche que Muljačić désigne de langue de Dubrovnik « *dubrovački jezik* »<sup>3</sup>, nous nous intéressons également à la langue vivante (vernaculaire), ce parler que les lexicographes ont enregistré dans leurs manuscrits et qui est utilisé dans sa continuité historique à Dubrovnik. Dans ses études récentes sur l'idiome de Dubrovnik, LOVRIĆ JOVIĆ (2014) prétend que cet idiome comprend la « langue littéraire de Dubrovnik » et le « parler de Dubrovnik ». Tout cela pris en compte, nous mettrons en avant, dans notre analyse des dictionnaires manuscrits, un corpus représentatif d'unités lexicales qui relèvent de ces deux groupes. Ces unités sont encore appelées les « mots ragusains » ou « raguséismes » (GABRIĆ BAGARIĆ 2002).

#### 4.1 Dictionnaire croate-italien (1597) de Bartol Kašić

C'est le Dictionnaire croate-italien de Kašić, rédigé entre 1597 et 1599, qui témoigne d'une activité lexicographique bien vive à Dubrovnik. Le manuscrit, conservé à la Bibliothèque des Frères mineurs à Dubrovnik, fut étudié et publié en 1990 par père Vladimir Horvat, SI.<sup>4</sup> *Dizionario illirico-dalmatino italiano – Slovoslovje dalmatinsko italijansko*, ou le dictionnaire chakavien, comme le désigne MIROVIĆ (1909) est souvent associé, dans un nombre d'études, au rapport du Ragusain Marin Temperica, fait à l'intention de général Acquaviva, qui ordonna à Kašić de rédiger une grammaire.<sup>5</sup> Comme le prétend Horvat, Kašić avait déjà

<sup>3</sup>Ensemble de réalisations écrites de la langue littéraire de Dubrovnik (MULJAČIĆ 2001).

<sup>4</sup>L'ouvrage s'intitule *Hrvatsko-talijanski rječnik* 'Dictionnaire croate-italien', son *editio princeps* fut préparée par père Vladimir Horvat, SI, publié dans les éditions de Kršćanska sadašnjost et de l'Institut de la langue croate (HORVAT 1990).

<sup>5</sup>Comme le prétend PUTANEC (2000), Temperica propose, en 1582, de confectionner un dictionnaire « croate ».

achevé son dictionnaire au moment du rapport de *Temperica*. (HORVAT 1990 : XV). De toute évidence, vers les années 1595, Kašić a rencontré Faust Vrančić à Rome qui, à ce moment-là, avait son *Dictionarium quinque nobilissimarum Europae linguarum* (1595) avec lui, imprimé à Venise ; Kašić l'a probablement utilisé lors de la création de son dictionnaire au moyen du procédé dit de « renversement » (HORVAT 1999 : 97).

Par ailleurs, le nombre de lexèmes empruntés à Vrančić, environ 3 000 mots selon GABRIĆ BAGARIĆ (1990) et HORVAT (1999 : 98), parle en faveur de la base chakavienne-ikavienne du dictionnaire, avec un nombre significatif de ragusésismes. Selon GABRIĆ BAGARIĆ (1990), les lexèmes les plus nombreux relèvent d'agronomie et d'élevage du bétail (y compris les noms d'animaux), ainsi que du quotidien. On y trouve également le vocabulaire en lien avec l'église et la liturgie, ainsi qu'un nombre significatif du vocabulaire marin et gastronomique.

Le *Dictionnaire croate-italien* était la source pour les dictionnaires qui ont suivi, surtout pour le *Dizionario* (1728) de Della Bella et *Rječosložje* (1806) de J. Stulli.

Le *Dictionnaire* de Kašić témoigne du fait que l'auteur lexicographe, dans sa tentative de rédiger le dictionnaire ayant le croate dans la colonne de gauche, s'appuie sur le *Dictionnaire* de Vrančić, le prend pour son point de départ et y ajoute une partie importante de son fonds lexical. Une orientation claire sur la base du dialecte shtokavien rend ce dictionnaire encore plus intéressant car il témoigne en outre du fait que dans ce corpus, les éléments chakavien et shtokavien se trouvent côte à côte, ce qui fait de ce dictionnaire un lien entre les régions croates du nord et du sud.

#### 4.2 *Dictionarium Latino-Illiricum* (1715–1716) de Đuro Matijašević

Ce dictionnaire constitue un des maillons les plus importants de la chaîne lexicographique de Dubrovnik. C'est la première fois qu'un locuteur natif fait un recueil lexicographique de l'idiome de Dubrovnik. En plus, c'est la seule trace du projet lexicographique inscrit au programme de l'Académie des Otieux (*Accademia degli Oziosi Eruditi*) fondée à Dubrovnik vers 1690 avec une attention particulière axée sur les questions de langue littéraire et d'orthographe, ainsi que sur le développement culturel général (les travaux des membres de l'Académie portaient également sur les thèmes d'histoire, en particulier la tradition populaire). Le début du 18<sup>ème</sup> siècle à Dubrovnik étant marqué par un retard général dans la création littéraire (et linguistique), on peut constater l'impossibilité de dresser un véritable portrait linguistique de la ville-république. Ainsi le *Dictionarium* vient-il combler le vide créé dans le travail lexicographique de Dubrovnik en tant que centre culturel de l'époque.

Ce dictionnaire fait partie du Codex numéro 194 intitulé *Razlika skladanja slovinska* (abritant également le Dictionnaire croate-italien de Kašić), qui est conservé à la Bibliothèque du Couvent des Frères mineurs à Dubrovnik.<sup>6</sup> Les Figures 1 et 2 montrent la page de titre ainsi qu'une page choisie du *Dictionarium*.

Le latin figure dans la colonne de gauche et l'idiome croate ragusain dans la colonne de droite. Le principe lexicographique est explicatif, il y a beaucoup de synonymes de contact; la colonne de droite comporte bien souvent des définitions ou formules explicatives et la composante syntaxique est assez riche. Pour le lexique, on peut noter les influences significatives de langues romanes (italien) ainsi qu'un nombre de lexèmes d'origine dalmate.

Le *Dictionarium* représente un recueil d'environ 13 000 entrées, il s'agit de la combinaison d'un dictionnaire général explicatif et d'un dictionnaire général à forte démarche traductive avec des éléments encyclopédiques et lexicaux. Dans l'orthographe, Matijašević s'appuie largement sur les solutions de ses précurseurs (R. Džamanjić et B. Kašić), mais introduit également quelques améliorations intéressantes : le phonème /ʒ/ sous forme du digramme *zf*, et le phonème syllabique /r/ sous forme du digramme *ær*.

Dans sa méthode lexicographique, Matijašević s'inspire considérablement de Mikalja. Il convient de mentionner le caractère descriptif de la colonne de droite (BILIĆ & FRANIĆ 2007), qu'elle soit constituée d'un seul ou plusieurs équivalents (synonymes), ou bien des éléments encyclopédiques :

- équivalent en un mot, ex. *primam* 'je reçois' s. v. *accipio*;
- bloc syntagmatique, ex. *cinim gorjet* 'je fais brûler' s. v. *igneo* ou locution *npr. mala snaga* 'manque de force' s. v. *anxietas*;
- série de deux ou trois synonymes, ex. *kantun, nugao* 'coin' s. v. *angulus, gredem, idem, hodim* 'je marche' s. v. *eo*;
- équivalent d'un mot + définition, ex. *mencetta, sgradda tvarda* 'bâtiment fort' s. v. *aro*; *fcilla fcivotna, duscnik* 'larynx' s. v. *arteria*;
- définition à deux ou plusieurs éléments : ex. *cistim klacinatu, i sovarnu od fida staroga* 'j'enlève les décombres' s. v. *erudero*.

---

<sup>6</sup>Ce Codex manuscrit fut décrit par P. Mitrović en 1909, dans son étude *Četiri nepoznata dubrovačka rječnika* 'Quatre dictionnaires ragusains inconnus'. Outre le *Dictionarium* de Matijašević, le codex abrite le *Vocabolario italiano-illirico* de L. Cekinić, (cf. § 4.3.), le *Dictionnaire croate-italien* de B. Kašić (cf. § 4.1.), et un quatrième dictionnaire italien-slovène dont l'auteur est inconnu.

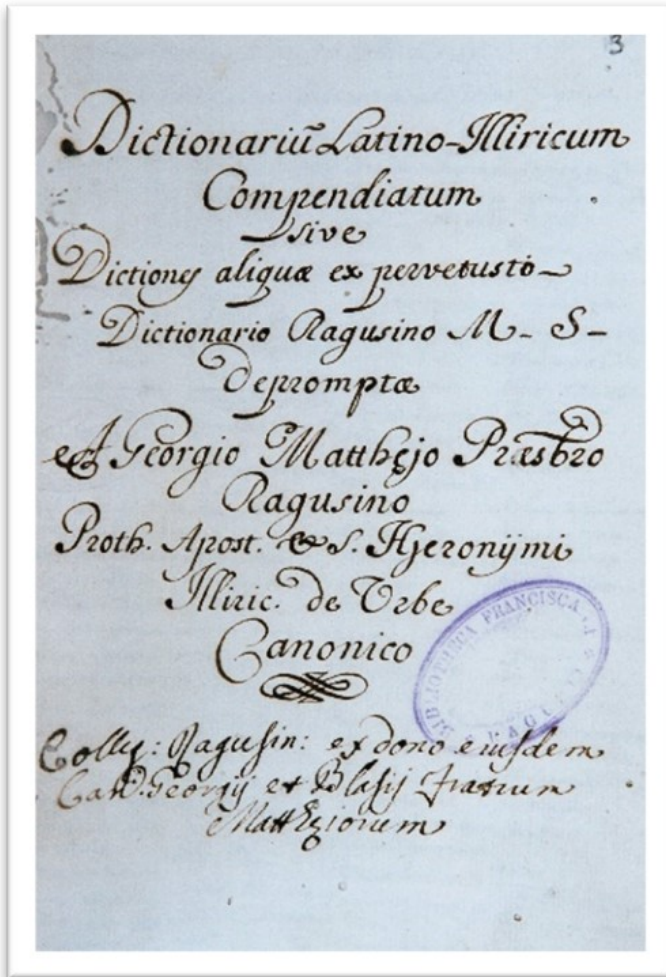


FIGURE 1 : *Dictionarium Latino-Illicum* (1715–1716) de Đuro Matijašević, page de titre, manuscrit 194 conservé à la Bibliothèque du Couvent des Frères mineurs à Dubrovnik



Il convient de souligner ici une autre structure, notamment la structure « A de B ou de + génitif » (ou expression de la possession par le génitif, selon GABRIĆ BAGARIĆ (2002), ayant différentes fonctions : attribut, origine, qualité, apposition, CC de but. En voici quelques exemples :

- *Ambrices*. Kupe od kucchje 'tuiles de la maison' (attribut)
- *Laurinum*. Ulje od lovorikke 'huile de laurier' (origine)
- *Litus*. Kraj od morra 'bord de la mer' (qualité)
- *Boreas*. Vjettar od burra 'vent de bura (nord)'
- *Epidaurus*. Stari Grad od Zaptata 'vienne ville de Cavtat' (apposition)
- *Cretera*. Romjenca od vode 'pot à l'eau' (CC de but)

Citons également la présence de nombreux éléments encyclopédiques, ce qui ne surprend pas, vu l'objectif de l'auteur d'expliquer un certain nombre d'éléments encyclopédiques :

- *Scabellum*. Podnošcje, illi ono na cem Biskup kleci, alli Knef 'plancher ou pied où l'évêque s'agenouille ou le prince'
- *Septentrio*. Strana od svjeta kudse obratisc put istoka kojati na ljevu ruku ostanne 'le côté quand tu tournes vers l'Est puis le côté qui reste sur ta main gauche'
- *Styx*. Jedan kladenaž vode u Arkadii od tolike studeni svaki kojese napie od tolike studeni svake umre illiti scivina, illi covjek, i svaki sud pužza od gne 'une source en Arcadie tellement froide que chacun qui en boit cette eau froide meurt, soit animal, soit humain, et tout récipient en est cassé'
- *Tigris*. Scivina bærfa kako strjela i rjeka koja istjece if Raja femaglskoga 'animal rapide comme une flèche et la rivière qui prend sa source dans le Paradis terrestre'

Pour ce qui est de la structure du lexique, *Dictionarium* comprend de nombreux éléments alloglottes (italiens et dalmates) ainsi qu'un nombre impressionnant de mots nouveaux, fruit de la créativité de l'auteur.

Le vocabulaire de spécialité semble assez riche, si l'on considère les phytonymes par exemple. De cette catégorie nous avons choisi une dizaine d'éléments qui d'ailleurs font partie du lexique habituel de nos jours à Dubrovnik, et dont l'origine est clairement romane (soit les emprunts à l'italien, soit au dalmate).

- *bosilak* s. v. *acapnos* ‘basilic’
- *cempres* s. v. *cuprenus* ‘cypès’
- *cesvinna* s. v. *ilex* ‘houx’
- *dub cedar* s. v. *citrus* ‘cèdre’
- *fjerla* s. v. *ferula* ‘fèrùle’
- *garofao* s. v. *gariophilum* ‘girofle’
- *kukumar* s. v. *cucumer* ‘concombres’
- *lovorika* s. v. *laurus* ‘laurier’
- *mjendeo* s. v. *amigdala* ‘amande’
- *morac* s. v. *feniculum* ‘fenouil’
- *mærkatugna* s. v. *cotoneum* ‘coing’
- *smokva* s. v. *ficus* ‘figue’
- *sparoga* s. v. *asparagus* ‘asperge’
- *smokva suha* s. v. *carica* ‘figue sèche’
- *fcukka* s. v. *juncus* ‘jonc’
- *Zærnika* s. v. *ilex* ‘houx’

Dans ce qui suit, nous présentons un répertoire choisi de lexèmes qui tirent leur origine du dalmate, groupe de parlers romans autochtones répandus le long de la côte adriatique, dont les influences persistent de nos jours en croate, surtout au niveau lexical. En voici des attestations recueillies par Matijašević dans le *Dictionarium* (pour l’explication des origines de mots, cf. FRANIĆ 2014 : 146) :

- *arkuo* s. v. *lecythus* ‘flacon’
- *bumbak* s. v. *bambacium* ‘coton’
- *fjerla* s. v. *ferula* ‘fèrùle’
- *kacuo* s. v. *cacabus* ‘pot, chaudron’

- *kelomna* s. v. basis 'bas de pilier'
- *kimak* s. v. cimex 'punaise'
- *klacina* s. v. calcaria 'four à chaux'
- *kupjerta* s. v. doma 'toit'
- *lukjerna* s. v. lucerna 'lucerne, lanterne'
- *mjendeo* s. v. amigdala 'amande'
- *pikat* s. v. ficatum 'figue'
- *plakjeri* [cas nominatif, pluriel] s. v. apkrodisia 'plaisir'
- *prigam* s. v. frigo 'frire'
- *rusa* s. v. rosa 'rose'

L'auteur emploie fréquemment le verbe *činiti* 'faire' qui peut servir de semi-auxiliaire comme dans l'exemple *cinim rat* s. v. belligero 'je fais la guerre' ou *cinim gosbu* s. v. convivor 'je fais la fête'. Ce même verbe peut introduire la diathèse causative comme dans l'exemple *ranu činju* s. v. exulcero 'je provoque une blessure' ou *cinim gorjet* s. v. igneo 'je fais brûler'.

L'analyse du lexique du *Dictionarium*, faite dans FRANIĆ (2014) montre qu'il existe un nombre relativement significatif de lexèmes que Matijašević emprunte au *Blago* de Mikalja, cependant, on observe dans le *Dictionarium* un nombre important de lexèmes qui ne figurent pas dans le *Blago* de Mikalja. Par ailleurs, Della Bella a sûrement dû prendre en compte le lexique de Matijašević lors de la confection de son *Dizionario*. On trouve également un nombre de lexèmes qui ne sont pas répertoriés chez Della Bella. Néanmoins, il existe un répertoire général significatif qui est commun aux *Dictionarium* (Matijašević), *Blago* (Mikalja) et *Dizionario* (Della Bella).

D'une part, le *Dictionarium* regroupe de nombreuses attestations lexicographiques d'un nombre de lexèmes ragusains qui ne figurent pas dans le *Blago* et le *Dizionario* : *art* 'art', *bankjer* 'banquier', *barun* 'baron', *batessa* 'abbesse', *duplavam* 'je double', *febra* 'fièvre', *kalamar* 'calamar', *kantun* 'coin', *klacinata* 'chaux', *konao* 'canal', *kortigl* 'cour', *kukumar* 'concombre', *kupjerta* 'toit', *libro* 'livre', *mencetta* 'bâtiment fort', *merit* 'mérite', *modeo* 'modèle', *segnavam* 'je fais signe', *tenero* 'tendre', *lukjerna* 'lucerne, lanterne'.

D'autre part, le *Dictionarium* comprend de nombreux éléments allogottes (italiens et dalmates) ainsi que des premières attestations des lexèmes ragusains typiques : *bankjer* 'banquier', *ciret* 'onguent obtenue en mélangeant l'huile et la cire', *kamelja* 'chameau', *klacinata* 'chaux', *romjenciza* 'petit pot', *segnavam* 'je fais signe', *moskar* 'éventail'. Si on le compare avec le dictionnaire *Rječosložje* (1806) de Stulli,<sup>7</sup> publié presque un siècle après le *Dictionarium*, on peut supposer que Stulli aurait pu emprunter au *Dictionarium* les raguséismes *kacuo* 'pot, chaudron', *duplavati* 'doubler' ou certains autres, ayant des formes intéressantes : a) substantifs *okuscaj* 'goût', *samos* 'solitude', *skafan* 'exemple', *skupgljevina* 'pile, amas', *vacellanje* 'hallucination', *vosctanizza* 'chandelle', et b) verbes *bludnujem* 'je suis surabondant', *kakochjem* 'je caquette', *vacellam* 'j'hallucine', *spjegnam* 'j'écume' (cf. FRANIĆ 2014).

Désireux de proposer des équivalents en un mot pour les termes latins, l'auteur créé des formes « pures » en s'appuyant sur la langue littéraire de Dubrovnik, mais aussi sur sa propre intuition : ainsi propose-t-il plusieurs néologismes, dont nous citerons des substantifs *nomina agentis* : *bdilaž* 'celui qui veille', *ctioz* 'lecteur', *daronosca* 'celui qui apporte des cadeaux', *duggofcivaž* 'celui qui vit longtemps', *ifjelaž* 'celui qui mange beaucoup', *ifostalaž* 'celui qui est absent', *ifgovaralaž* 'celui qui dit quelque chose', *kuhacižza* 'cuisinière', *kupilaž* 'acheuteur', *velegovorilaž* 'qui parle beaucoup', *tesckogovorilaž* 'celui qui parle avec difficulté', *vinonos* 'qui porte du vin'.

Sur l'exemple des vocabulaires de spécialité en rhétorique, philosophie et philologie on observe les efforts de l'auteur de proposer des explications originales de plusieurs notions abstraites :

- *spodobnos, kad jedno drugomu odgovara* 'ressemblance, quand une chose correspond à une autre' s. v. analogia ;
- *reccenja u kratko nu spametno* 'mots courts mais intelligents' s. v. dicteria ;
- *govorilaž koga se govor na dva nacina rafumje* 'locuteur dont les paroles sont compris de deux manières' s. v. flexiloquus ;
- *flamenje u pismu na nacin male svjefdizze* 'signe écrit à la façon d'une petite étoile' s. v. asteriscus ;
- *pismo od svoje ruke* 'écriture de sa propre main' s. v. autographum ;
- *tko ispiscuje, alli if jednoga pisma na druggo stavglja* 'qui écrit ou transmet d'une écriture à une autre' s. v. ascriptor ;

---

<sup>7</sup>*Rječosložje iliričko-italijansko-latinsko* (1806).

- *art fa dobro govorit i pisat* 'art de bien parler et écrire' s. v. grammatica.

Les résultats de l'analyse du lexique montrent que le *Dictionarium* de Matijašević, bien qu'il soit resté manuscrit, a servi de modèle ou de source pour d'autres manuels lexicographiques, en particulier les dictionnaires imprimés, surtout ceux de Della Bella et de Stulli. En outre, ce dictionnaire est unique en ce qui concerne les efforts de l'auteur en matière de pureté linguistique, en d'autres termes, de purification du lexique figurant dans la colonne de droite des influences venant d'autres langues, surtout de l'italien.

### 4.3 Vocabolario italiano-illirico (~1745) de frère Lovro Cekinić

Frère Lovro Cekinić (1692–1752), franciscain de l'Ordre des Frères mineurs, installé à Dubrovnik, collectionneur assidu d'œuvres littéraires et historiques, excellent prédicateur, est bien fréquemment mentionné dans les chroniques de l'époque. Les biographes en parlent souvent, pour la plupart ils font mention de ses activités dans la rédaction des chroniques, des transcriptions et de la prédication.<sup>8</sup> Bien que son activité lexicographique soit rarement mentionnée, c'est en 1909 que P. Mitrović décrit ses efforts en lexicographie.<sup>9</sup>

Le *Vocabolario* figure parmi quatre dictionnaires manuscrits reliés dans le *Codex Razlika skladanja slovinska*, conservé à la Bibliothèque des Frères mineurs. De par son étendue, il est assez petit : il regroupe environ 5 080 entrées, dont un nombre significatif de lacunes (absence d'équivalent), presque 600.<sup>10</sup> Les Figures 3 et 4 montrent la page de titre du *Vocabolario* ainsi qu'une page de dictionnaire :

Le *Vocabolario* fut rédigé presque un siècle après le *Blago* de Mikalja et quelques années après la parution de *Dizionario* de Della Bella. En conséquence, on aurait pu s'attendre à une influence significative des deux grands dictionnaires sur le corpus lexical de Cekinić. Malgré l'existence de ces deux grands dictionnaires et leur riche fonds lexical, on peut observer un haut degré de créativité lexicale de Cekinić. En effet, un nombre relativement important de mots ont leur première attestation lexicographique dans le *Vocabolario* (ces mots trouveront plus tard leur place dans le *Rječosložje* de Stulli sans aucune mention de source), nous n'en citerons que quelques-uns : *boghoneuiernik* 'infidèle', *duorkigniza* 'demoiselle', *neinako* 'non autrement', *neufati* 'désespérer' *nochiascni* 'nocturne', *odredba*

<sup>8</sup>Cf. FRANIĆ 2007 et FRANIĆ 2008.

<sup>9</sup>Cf. MIROVIĆ 1909 : 528–533.

<sup>10</sup>Il est fort probable que le travail sur la traduction de l'œuvre *Duhovne zabave duše bogoljubne* fut la principale motivation pour la création du dictionnaire (vers 1745), ce dont nous fait part D. Berić 1956.

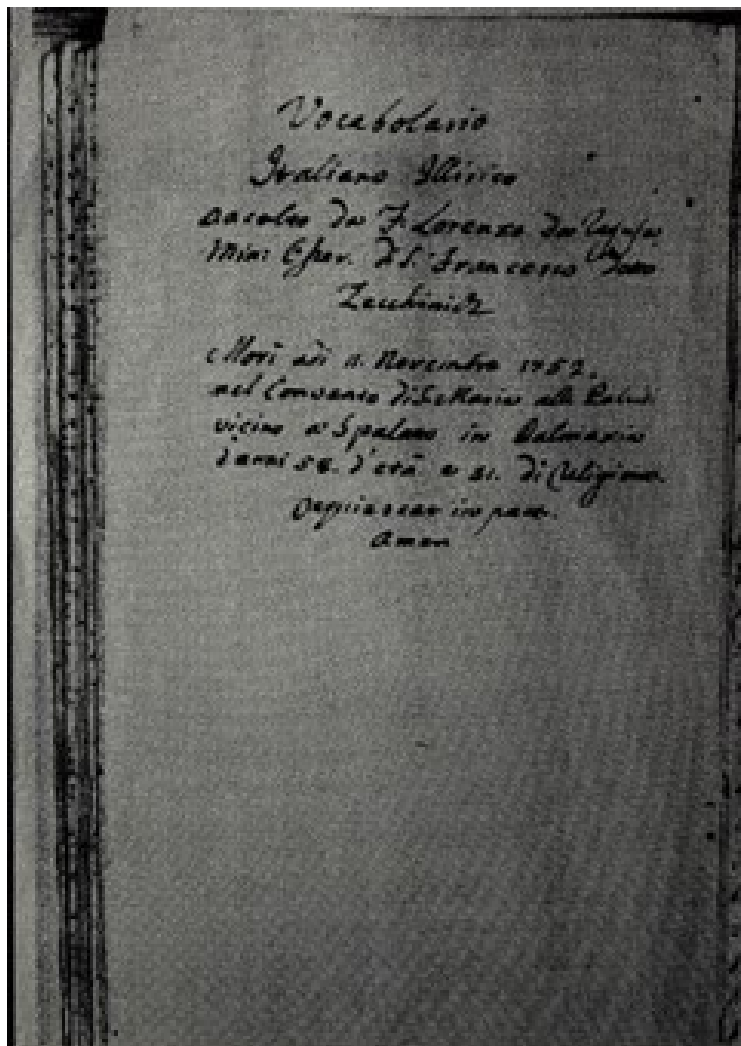


FIGURE 3 : Vocabolario italiano-illirico (~1745) de Lovro Cekinić, page de titre, manuscrit 194 conservé à la Bibliothèque des Frères mineurs de Dubrovnik



FIGURE 4 : Vocabolario italiano-illirico (~1745) de Lovro Cekinić, page 203, manuscrit 194 conservé à la Bibliothèque des Frères mineurs de Dubrovnik

‘disposition’, *okorusciati* ‘écorcer’, *posliednizi* ‘descendants’, *posnasgiti* ‘renforcer, encourager’, *prasnoruk* ‘aux mains vides’, *rasiediniti* ‘séparer’, *raslauiti* ‘diluer’, *rasrok* ‘louche’, *suietostuornik* ‘qui crée le monde’ (FRANIĆ 2007 et FRANIĆ 2008).

Par ailleurs, on observe rarement, parmi les solutions lexicographiques de Cekinić, des lexèmes d’origine étrangère. Les emprunts à l’italien sont les plus nombreux, par exemple : *aier* ‘air’ s. v. *aere*; *casctio* ‘forteresse’ s. v. *fortezza*, *castello*; *kollaina* ‘collier’ s. v. *collana*; *legatti* ‘lire’ s. v. *leggere*; *presciti* ‘se dépêcher’, être pressé’ s. v. *affrettare*, puis des descendants lexicaux (ou reliques) du dalmate : *cersa* ‘blanchisseur’ s. v. *biaxa*; *kaipa* ‘cage’ s. v. *gabbia*; *pedepsa* ‘punition’ s. v. *flagello*; *podumienta* ‘fondement’ s. v. *fondamento*; *prigati* ‘frîre’ s. v. *lessare*.

Malgré sa finalité pratique et une ampleur moins importante, le *Vocabolario* reflète des tentatives de Cekinić de chercher de bons équivalents dans le vernaculaire. En effet, bien qu’il propose un emprunt, Cekinić ajoute à titre égal, son équivalent local, bien souvent un raguséisme : *pouietarze* à côté de *aier* ‘air’ s. v. *aere*; *bielilos* à côté de *cersa* ‘blanchisseur’ s. v. *biaxa*; *smai* à côté de *drokun* ‘dragon’ s. v. *dragone*; *pustignak* à côté de *remeta* ‘ermite’ s. v. *eremita*; *temegl* à côté de *podumienta* ‘fondement’ s. v. *fondamento*.

Sa recherche des équivalents issus de l’idiome de Dubrovnik est particulièrement évidente en ce qui concerne les notions de divers domaines scientifiques, de sorte que le *Vocabolario* répertorie pour la première fois certains termes de philosophie, rhétorique et théologie : *boghoneuiernik* ‘infidèle’, *daunoroditegl* ‘qui est né avant’, *boghociastiti* ‘honorer Dieu’, *istomacnik* ‘interprète’, *kgnigostuorni* ‘académique’, *mudrogouornos* ‘éloquence’, *saglauglie* ‘conclusion’, *snanosakonik* ‘législateur’. Ainsi, les recherches lexicales de Cekinić suivent-elles les traces de tentatives entamées par Matijašević et ses contemporains au sein de l’Académie des Otieux, vers la fin du 17<sup>ème</sup> siècle. Un taux peu significatif d’éléments alloglottes parle en faveur du purisme de Cekinić dans la recherche d’équivalents. Ceci fait du *Vocabolario* un maillon important dans l’histoire de la construction de la langue littéraire croate et marque un pas en avant dans le développement de la lexicographie bilingue.

#### 4.4 Dizionarietto italiano-slavo-moscovitico (1751) de Ivan Marija Matijašević

Ce fut le chercheur croate M. Deanović qui a publié, dans les années 1950, le dictionnaire manuscrit trilingue italien-croate-russe *Dizionarietto italiano-slavo-moscovitico*<sup>11</sup> (DEANOVIĆ 1951, 1952) du jésuite et Ragusain, père Ivan Marija

<sup>11</sup>Ce dictionnaire fait partie du recueil R 4091 qui est conservé à la Bibliothèque nationale et universitaire de Zagreb

Matijašević<sup>12</sup> (1714–1791) (Johannes Maria Mattei, Gian Maria Mattei, Giammaria Mattei), auteur d'œuvres religieuses et prédicateur, gardien du patrimoine jésuite à Dubrovnik et neveu du père Đuro Matijašević (DOLCI 1767 : 34, BRLEK 1952, LUČIĆ 1997).

Ce dictionnaire trilingue est bien spécifique : il fut créé au bord d'un navire, en trois semaines, entre le 6 et le 28 octobre 1751, lors du trajet de l'auteur de Ancona à Dubrovnik. Ce voyage fut décrit dans un journal, publié aussi par Deanović en 1952.

Le dictionnaire regroupe 1087 entrées italiennes, avec 35 entrées en annexe. Compte tenu des circonstances de l'élaboration du dictionnaire (le texte fut écrit lors d'un trajet), les entrées sont dispersées sans aucun critère, la plupart des entrées étant des substantifs et des adjectifs, avec quelques expressions. (DEANOVIĆ 1951 : 578) Avec une tendance à la pureté de la langue, l'auteur ne note pas d'éléments alloglottes, c'est pour cela que l'on ne trouve pas d'équivalents pour les mots italiens *arte, artigiano, bagatella, bandiera, bestimiere, brocca, camera, canale, carico, carro, candeliere, cassa, cipresso, consolazione, collana, colonna, cre-sima, coniglio, danari, disperato, disgrazia, differenza, dispiacere, doppio, eccellente, elemosina, erede, esame, esperienza, evangelio, favore*. Néanmoins, Deanović observe que tous ces mots italiens ont leurs équivalents dans le parler ragusain et qu'il est peu probable que l'auteur ait oublié ces mots de son enfance (nous ne citons que *bagatela, kamara, kono, kandilijer, kasa, čempres, konsolacion, kolana, kolona ou kelomna, dizgracija, diferenca, lemozina, ered, esperjenca, vandelje, fa-vor*, DEANOVIĆ 1951 : 581) . Dès son retour à Dubrovnik, Matijašević continue à travailler sur le dictionnaire en ajoutant certains mots.

L'importance de ce petit recueil manuscrit repose surtout sur le lien familial entre le neveu Ivan Marija et l'oncle Đuro Matijašević. De fait, Ivan Marija a dû disposer, à un certain moment, après la mort de son oncle Đuro, du dictionnaire *Dictionarium latino-illiricum*, car son oncle l'a laissé, ainsi que tout son héritage, aux jésuites ragusains. Par ailleurs, la remarque de Deanović disant que

l'idée même de cette improvisation pourrait être plus importante du point de vue historico-culturel, que son exécution peu réussie, qui d'ailleurs a peu de signification linguistique (DEANOVIĆ 1951 : 581, notre traduction).

fournit un bon résumé sur ce manuscrit. Il convient de constater, malgré tout, que ce dictionnaire s'inscrit dans la continuité des activités lexicographiques à

<sup>12</sup>Il obtient son doctorat en 1733 à Rome, où il passe une période en compagnie de Ruđer Bošković, chercheur, philosophe, mathématicien, astronome et diplomate croate (1711–1787). Bošković, Ruđer Josip. Hrvatska enciklopedija, mrežno izdanje. Leksikografski zavod Miroslav Krleža, 2021. Accédé 27. 4. 2023. <http://www.enciklopedija.hr/Natuknica.aspx?ID=8948>

Dubrovnik et que Matijašević nous laisse un contact bien intéressant de trois langues.

#### 4.5 *Vocabula italico-illyrica* (1760) de Mato Klačić

Ce dictionnaire manuscrit n'a pas encore attiré une attention particulière. C'est M. Brlek qui le décrit dans son étude de 1952, intitulée *Rukopisi knjižnice Male braće u Dubrovniku* 'Les manuscrits de la Bibliothèque des Frères mineurs'. Plus précisément, il cite le dictionnaire « italien-illyrique » de Mato Klačić *Vocabula italico-illyrica*.<sup>13</sup> Par endroits, l'auteur n'a pas inscrit de mots, de sorte que certains passages dans le texte sont restés vides. Après avoir terminé la dernière lettre Z (feuille 115), l'auteur recommence sur la feuille 116 et continue jusqu'à la feuille 162.

Pour les entrées de l'italien (colonne de gauche) on observe parfois plusieurs équivalents :

- *Filosofo, mudrosnanaz, mudaraz, rasumnik* 'philosophe', 'sage' ;
- *Asino, touar, osso, magarre, poslenik* 'âne' ;
- *Parlare, Gouoriti, Bessiediti, Isrizati, Rieti, Sborrit, Slouit* 'parler'.

Klačić emprunte une partie de lexèmes aux dictionnaires déjà publiés, tout en formant certains mots lui-même :

- *Canale, legno cauato, Vodotocie, Vodeni prouod* 'canal' ;
- *Pittore, Slikouaç, Pengatur* 'peintre' ;
- *Prossimo, Iskargni, Sussied, Blixnik* 'proche', 'voisin'.

Parfois, l'auteur apporte des locutions intéressantes et typiques pour la langue parlée : nous citons les entrées *lingua* 'langue' et *ora* 'heure' :

*Lingua, Jesik; Lingua cauata fuori, Isplafen iesik; Auer sula punta della lingua, Imat scto na iesiku; Lingua /fauella/, Jesik; Lingua paesana, Materin Jesik; Lingua di cane /erba/ Tarput; La punta della lingua, Varh od iesika.*

Voici quelques traits spécifiques morphologiques et lexicaux de l'idiome de Dubrovnik :

---

<sup>13</sup>Il s'agit du manuscrit no 237, intitulé Clasci Matthaesus, *Vocabula italico-illyrica*. D'après la datation sur la page de titre, le dictionnaire fut créé en 1760. BRLEK (1952) (236–237) fournit plus de détails sur ce recueil (1952 : 236–237). Une courte description est donnée dans FRANIĆ (2014).

- *Cardinale, Kardeno* [langue standard croate : kardinal];
- *Cantilena, Popieuka, Colenda*;
- (sortes de figue) : *Fico albero, Smokva, Braxetta, Petrouaccia, GlusciŽa, Prizrella, Mletkigna, NeŽrella, OŽimiŽa, Suha smokva*;
- *Sapone, Midlo, Saplun* [langue standard croate : sapun].

Ces exemples ne sont qu'un petit extrait du trésor de l'idiome de Dubrovnik en contact avec l'italien, tel qu'il fut enregistré par Klačić. Bien que ce dictionnaire soit créé après le grand *Dizionario* de Della Bella, ce petit recueil mérite une analyse plus détaillée qui pourrait lui attribuer sa véritable place dans la chaîne ininterrompue de la lexicographie ragusaine et partant la lexicographie croate.

#### 4.6 Dictionnaire anonyme, manuscrit n° 278

Dans son étude sur les manuscrits ragusains, BRLEK (1952) aborde, entre autres, le dictionnaire anonyme italien-croate, conservé à la Bibliothèque des Frères mineurs à Dubrovnik, numéro 278. Le texte du dictionnaire commence par la lettre D et le mot « dissimile, nesličan ». En effet, Brlek constate que le manuscrit contient plusieurs écritures, faites par différentes mains. Les lettres A, B et C viennent à la fin du dictionnaire, ce qui fait penser soit au fait que ces pages ont été déchirées, soit que quelqu'un a ajouté ces entrées plus tard. (BRLEK 1952 : 283).

Parfois, l'auteur ajoute plusieurs équivalents correspondant aux entrées italiennes, p. ex. :

- Barcha 'navire' : lagia, brod, plov;
- Legare 'lire' : sctiti, legatti;
- Malizioso 'malicieux' : himben, priuarni, sciarovit;
- Monte 'montagne' : bardo, planina, gora, glauiza.

On y trouve des explications et définitions, p. ex. :

- logica 'logique' : nauk za razložiti dobro;
- cerbero 'cerbère' : pas is pakla od tri glaue;
- idolatrare 'idolâtrer' : klagnatse Bogovom lafciviem;

- lido ‘sable’ : igalo, kraj od morra;
- metafora ‘métaphore’ : prinescenie od rieci;
- metropolli ‘métropole’ : paruostolni grad.

A la page 192 du manuscrit se trouve un court inventaire de toponymes ragusains, ce qui fait BRLEK (1952) constater que le dictionnaire fut conçu à Dubrovnik, par un locuteur natif, probablement dans le Couvent des franciscains. Etant donné que ce recueil abrite un bon nombre d’éléments lexicaux d’origine ragusaine et vu la liste de toponymes ragusains, ce dictionnaire constitue un apport significatif aux recherches des contacts du croate et de l’idiome ragusain avec la langue italienne.

## 5 Conclusion

Les dictionnaires manuscrits constituent un maillon très important dans l’histoire de la lexicographie en général, ainsi que dans l’histoire de la lexicographie croate. Sur l’exemple de six dictionnaires manuscrits du cercle lexicographique de Dubrovnik et du corpus représentatif d’éléments de l’idiome de Dubrovnik, en particulier les raguséismes lexicaux qui y sont répertoriés, nous constatons que ces dictionnaires ont joué un rôle fort important dans la création de grands dictionnaires imprimés, en particulier aux 17<sup>ème</sup> et 18<sup>ème</sup> siècle.

Seule une étude approfondie des dictionnaires manuscrits peut permettre une meilleure évaluation de la production lexicographique dans sa continuité historique au sein d’une communauté linguistique. En outre, une telle évaluation donnera un aperçu au niveau horizontal et témoignera des contacts linguistiques et culturels entre les langues et cultures européennes de l’époque (DUBOIS 1970). Ce n’est que si les dictionnaires manuscrits côtoient les dictionnaires imprimés qu’il sera possible de mieux mettre en évidence la continuité des modèles et influences (SAMARDŽIJA 2013), de sorte que chaque dictionnaire représente une somme de dictionnaires qui le précèdent ainsi qu’un modèle pour les dictionnaires qui suivent.

## Références

BILIĆ, Josip & Ivana FRANIĆ. 2007. Contact linguistique croate-italien : l’exemple de la lexicographie ancienne de Dubrovnik. *Chroniques slaves*. 175-190.

- BRLEK, Mijo Ivan. 1952. *Rukopisi Knjižnice Male braće u Dubrovniku*. Zagreb : JAZU.
- COLLIGNON, Lucien & Michel GLATIGNY. 1978. *Les dictionnaires. Initiation à la lexicographie*. Paris : CEDIC.
- DEANOVIĆ, Mirko. 1951. Talijansko-hrvatsko-ruski rječnik iz godine 1751 / Un Dizionarietto italiano-slavo-moscovitico del 1751, Zagreb. *Zbornik radova Filozofskog fakulteta u Zagrebu* I. 567-612.
- DEANOVIĆ, Mirko. 1952. Dnevnik Iva M. Matijaševića. *Anali Historijskoga instituta JAZU* 1. 279-330.
- DOLCI, Sebastian. 1767. *Fasti litterario-Ragusini sive virorum litteratorum Qui usque ad annum MDCCLXVI in Ragusina claruerunt Ditione*. Venetiis : Gaspar Stort.
- DUBOIS, Jean. 1970. Dictionnaire et discours didactique. *Langages* 5. 35-47. DOI : 10.3406/lgge.1970.2590. [https://www.persee.fr/doc/lgge\\_0458-726x\\_1970\\_num\\_5\\_19\\_2590](https://www.persee.fr/doc/lgge_0458-726x_1970_num_5_19_2590).
- FRANIĆ, Ivana. 2007. Rukopisni Vocabolario italiano-illirico (~1745) fra Lovre Cekinića. *Rasprave Instituta za hrvatski jezik i jezikoslovlje* 33. 91-105.
- FRANIĆ, Ivana. 2008. Fra Lovro Cekinić (1692-1752). *Anali Dubrovnik* 46. 125-136.
- FRANIĆ, Ivana. 2014. *Tragovima hrvatskih rječnika. Dubrovački Dictionarium Latino-Illiricum (1715 – 1716) Đure Matijaševića*. Zagreb : FF Press.
- FREY, Claude. 2007. Les structures lexicographiques dans les dictionnaires francophones, une rencontre symbolique des mots et des cultures. In Agence universitaire de la FRANCOPHONIE (éd.), *Actes du colloque Penser la francophonie : Concepts, actions et outils linguistiques, Ouagadougou (Burkina Faso), du 31 mai au 1er juin 2004*, Éditions des Archives contemporaines, 197-210. Paris : AUF.
- FURNO, Martine. 2003. Qui et Que dans le Dictionnaire François-Latin de Robert Estienne. *Langue française* 139. 28-46.
- GABRIĆ BAGARIĆ, Darija. 1996. Kašićeva rukopisna Biblija i Blago jezika slovinskoga (1651.) Jakova Mikalje. *Rasprave Instituta za hrvatski jezik i jezikoslovlje* 22. 27-34.
- GABRIĆ BAGARIĆ, Darija. 1998. Nazivi bolesti u rječniku „Blago jezika slovinskoga“ (1651.) Jakova Mikalje. *Radovi Zavoda za slavensku filologiju* 32. 81-90.
- GABRIĆ BAGARIĆ, Darija. 2000. Dijalektna podloga rječnika Blago jezika slovinskoga (1649.–1651.) Jakova Mikalje. *Rasprave Instituta za hrvatski jezik i jezikoslovlje* 26. 45-58.
- GABRIĆ BAGARIĆ, Darija. 2002. Blago jezika slovinskoga J. Mikalje – ishodište hrvatske leksikografije. *Forum* 7-9. 1069-1077.
- GABRIĆ BAGARIĆ, Darija. 2010. *Na ishodištu hrvatske leksikografije*. Zagreb : Institut za hrvatski jezik i jezikoslovlje.

- GABRIĆ BAGARIĆ, Darija. 1990. *Leksička analiza Kašićeva Hrvatsko-talijanskoga rječnika' (uz pretisak Kašićeva rječnika)*, Bartol KAŠIĆ (éd.). Zagreb : Kršćanska sadašnjost i Zavod za jezik Instituta za filologiju i folkloristiku. 443-449.
- GALISSON, Robert. 1988. Cultures et lexicultures. Pour une approche dictionnaire de la culture partagée. *Annexes des Cahiers de linguistique hispanique médiévale* 7. 325-341. DOI : 10.3406/cehm.1988.2133.
- HORVAT, Vladimir. 1990. Bartol Kašić kao leksikograf : Autor Konverzacijskog priručnika (1595) i Hrvatsko-talijanskog rječnika (1599). In *Bartol Kašić, Hrvatsko-talijanski rječnik, V-XXII*. Zagreb : Kršćanska sadašnjost, Zavod za jezik.
- HORVAT, Vladimir. 1999. *Bartol Kašić - otac hrvatskoga jezikoslovlja*. Sveučilište u Zagrebu, Hrvatski studiji Studia Croatica.
- KATIČIĆ, Radoslav. 1981. Gramatika Bartola Kašića. *Rad JAZU* 20. 5-129.
- KATIČIĆ, Radoslav. 2013. *Povijest hrvatskoga jezika. 3. Knjiga : 17. I 18. Stoljeće*, t. 3. Zagreb : Društvo za promicanje hrvatske kulture i znanosti Croatica.
- KRASIĆ, Stjepan. 2009. *Počelo je u Rimu. Katolička obnova i normiranje hrvatskoga jezika u XVII. stoljeću*. Dubrovnik : Matica hrvatska, ogranak Dubrovnik.
- LANDAU, Sidney I. 2001. *Dictionaries : The art and craft of lexicography*. 2<sup>e</sup> éd. New York/Cambridge : Cambridge University Press.
- LOVRIĆ JOVIĆ, Ivana. 2014. *O starome dubrovačkom govoru nazbilj*. Zagreb : Institut za hrvatski jezik i jezikoslovlje.
- LUČIĆ, Josip. 1997. *Dubrovačko povijesno iverje*. Dubrovnik : Matica hrvatska, ogranak Dubrovnik.
- MATASOVIĆ, Ranko. 2013. Hrvatski jezik i ostali europski jezici u 17. i 18. stoljeću. In Katičić RADOSLAV & Josip LISAC (éd.), *Povijest hrvatskoga jezika. 3. knjiga : 17. i 18. stoljeće*, 511-532. Zagreb : Društvo za promicanje hrvatske kulture i znanosti Croatica.
- MEŠTROVIĆ, Zrnka. 1994. Onimijska komponenta jednojezičnog rječnika. *Filologija* 22-23. 169-174.
- MEŠTROVIĆ, Zrnka & Nada VAJS. 1995. Blago jezika slovinskoga Jakova Mikalje – jedan od leksikografskih izvora za *Lexicon latino-illyricum* Pavla Vitezovića. *Rasprave Zavoda za hrvatski jezik* 21. 139-155.
- MIROVIĆ, Pavo. 1909. Četiri nepoznata dubrovačka rječnika. *Nastavni vjesnik* XVII, 7. 522-539.
- MULJAČIĆ, Žarko. 2001. Može li se sredinom 17. stoljeća govoriti o istosti „dubrovačkog“ i „bosanskog“ jezika? *Dubrovnik* 2. 133-148.
- MURANO, Michela. 2013. Les noms propres dans le Grand dictionnaire français-italien, italien-français de Constanzo Ferrari et Joseph Caccia. *ELA. Études de linguistique appliquée* 170(2). 179-192.

- PRUVOST, Jean. 2021. *Les dictionnaires français, outils d'une langue et d'une culture*. Nouvelle édition actualisée. Paris : Ophrys.
- PUTANEC, Valentin. 1991. Tekstološka analiza rukopisnoga Hrvatsko-talijanskoga rječnika Rkp. 194 (Franjevački samostan u Dubrovniku). *Rasprave Zavoda za jezik* 17. 153-209.
- PUTANEC, Valentin. 2000. Prilozi za proučavanje povijesti hrvatske leksikografije i standarda u 16. i 17. stoljeću. *Rasprave. Časopis Instituta za hrvatski jezik i jezikoslovlje* 26(1). 215-243.
- REY, Alain. 1991. Divergences culturelles et dictionnaire bilingue. In *Wörterbücher, Dictionaries, Dictionnaires. An International Encyclopedia of Lexicography, Third Volume*, 2865-2870. Berlin, New York : Walter de Gruyter.
- REY, Alain. 2007. Le pont interculturel. In Giovanni DOTOLI (éd.), *L'Architecture du dictionnaire bilingue et le métier du lexicographe. Actes des premières Journées Italiennes des Dictionnaires*, 7-21. Capitolo-Monopoli, Fasano : Schena.
- REY, Alain. 2008. *De l'artisanat du dictionnaire à une science du mot*. Paris : Armand Colin.
- SAMARDŽIJA, Marko. 2013. Hrvatski leksik i leksikografija u 17. I 18. stoljeću. In Katičić RADOSLAV & Lisac JOSIP (éd.), *Povijest hrvatskoga jezika. 17. i 18. stoljeće*, t. 3, 429-461. Zagreb : Croatica.
- SAMARDŽIJA, Marko. 2019. *Hrvatska leksikografija. Od početaka do kraja XX. stoljeća*. Zagreb : Matica hrvatska.
- SIRONIĆ BONEFAČIĆ, Nives. 1992. Ardelio Della Bella i prvo izdanje njegova trojezičnog rječnika iz 1728. Godine. *Rad HAZU* 446. 1-69.
- TALLARICO, Giovanni. 2013. Les apports de la lexicographie bilingue à l'interculturel. *ELA. Études de linguistique appliquée* 170. 139-152. DOI : 10.3917/ela.170.0139.
- VAJS, Nada. 2003. *Hrvatska povijesna fitonimija*. Zagreb : Institut za hrvatski jezik i jezikoslovlje.
- VAXELAIRE, Jean Louis. 2005. Nom propre et lexicographie française. *Corela HS-2*. DOI : 10.4000/corela.1239. (2 avril, 2021).
- VONČINA, Josip. 1992. Jakov Mikalja u povijesti hrvatskoga književnog jezika. *Fluminensia* 4(2). 59-66.
- VONČINA, Josip. 2003. Hrvatski rječnici. In *Hrvatska i Europa, kultura, znanost i umjetnost*, t. III, 461-472. Zagreb : HAZU.
- WIERZBICKA, Anna. 1988. *The Semantics of Grammar*. Amsterdam : John Benjamins.
- ZGUSTA, Ladislav. 1971. *Manual of lexicography*. Berlin/New York : De Gruyter Mouton.



# Chapitre 10

## Évaluer la dette : L'étendue de la présence de Richelet dans le *Dictionnaire universel* de Basnage (1701)

Clarissa Stincone

Université Sorbonne Nouvelle – Paris 3

Although he does not quote Richelet in the preface to his edition of the *Dictionnaire universel* and attributes exclusively to the Académie française the role of guide and guarantee of his lexicographical choices, Basnage refers to Richelet even more than to the Académie française. Basnage not only transcribes entries or senses, syntactic constructions and authorial quotations from the *Dictionnaire françois* (1693), but also takes up its didactic and descriptive intentions. As both lexicographers are familiar, for different reasons, with the doubts of those learning a foreign language, they make similar lexicographical choices. The didactic intentions can be seen in the spelling chosen as well as in the author citations and grammatical elements provided in their dictionaries. Descriptive intentions are evident in the fact that Richelet and Basnage record all the linguistic usages of which they are aware and at most suggest their preference by relying on the authority of writers and linguists who have used or spoken in favour of a specific usage. The aim of this article is to assess the extent to which Basnage drew inspiration from Richelet by analysing the common features of the two lexicographers' dictionaries, while seeking to understand the reasons why Basnage neglected to refer to Richelet.

### 1 Introduction

Henri Basnage de Beauval révisé en 1701 le *Dictionnaire universel* d'Antoine FURETIÈRE (1690), dorénavant [DU1690]. Le travail lexicographique de Basnage



évolue par rapport à celui de Furetière sous plusieurs égards. En plus d'augmenter le nombre d'entrées et d'enrichir les entrées existantes de précisions au niveau du contenu, Basnage introduit dans son édition nombre d'informations lexicographiques que Furetière n'avait pas prises en compte. Il s'agit de citations d'écrivains, de marques d'usage contenant les jugements du lexicographe sur la fréquence d'utilisation des mots et les axes de la variation linguistique, d'éléments grammaticaux comme les flexions verbales et les constructions syntaxiques, d'indications concernant l'orthographe et la prononciation, de discussions des experts en matière de langage, de collocations. L'analyse de ces éléments permet de découvrir les sources de Basnage, parmi lesquelles le Dictionnaire de l'Académie française (ACADÉMIE FRANÇAISE 1694) et le *Dictionnaire françois* (RICHELET 1693), (première édition : RICHELET 1680) de César-Pierre Richelet occupent une place charnière.

Bien que Basnage s'inspire fortement des deux dictionnaires, il leur réserve un traitement bien différent. Le rôle de l'Académie française est reconnu dès la préface où Basnage affirme que « dans le partage des sentimens, l'Académie Française est l'autorité suprême à laquelle j'ai toujours deféré » (Préface, FURETIÈRE 1701) dorénavant [DU1701]. Il est confirmé dans la suite de l'ouvrage où les citations de l'Académie sous les formes « Ac. », « L'Acad. » et « Mrs. de l'Acad. » sont au nombre de 630 environ. En revanche, le nom de Richelet n'apparaît pas dans la préface tout comme il n'apparaît pas dans la liste de 130 abréviations d'auteurs et d'œuvres au début du dictionnaire. Dans le corps de l'ouvrage, Richelet n'est mentionné que 66 fois, dont deux sous la forme abrégée « Rich. » et souvent de manière critique. Il n'est pas rare que Basnage associe au nom de Richelet le verbe *pretendre* en invitant à préférer la variante orthographique proposée par l'Académie française :

« Richelet pretend qu'il faut prononcer, & même écrire *extrordinaire*; mais l'Academie dit positivement qu'il faut prononcer *extraordinaire*; & par consequent, c'est ainsi que les honnêtes gens prononceront toujours, au moins dans les discours graves » (EXTRAORDINAIRE, DU1701).<sup>1</sup>

Même s'il ne le reconnaît pas, Basnage emprunte à Richelet plusieurs éléments lexicographiques, voire des acceptions et des entrées entières, ainsi que deux

---

<sup>1</sup>Les mots en majuscules entre parenthèses indiquent l'entrée où se trouve la partie de texte citée. S'ils sont suivis de « DU1690 », cela signifie que la portion de texte ne se trouve que dans le *Dictionnaire universel* de Furetière et non dans celui de Basnage. S'ils sont suivis de « DU1701 », cela signifie, au contraire, que la portion de texte ne se trouve que dans le *Dictionnaire universel* de Basnage et non dans celui de Furetière. S'ils ne sont suivis de rien, cela signifie que le texte se trouve dans les deux éditions. Dans ce cas, l'orthographe adoptée est celle de Basnage.

grandes orientations de son travail, à savoir l'orientation didactique et l'orientation sociolinguistique<sup>2</sup>. Comme Richelet, Basnage s'efforce de répondre aux besoins d'apprentissage de son public. Richelet est un professeur de français<sup>3</sup> et s'adresse, entre autres, aux étrangers désireux d'apprendre sa langue, alors que Basnage est un émigré aux Pays-Bas, et s'adresse, parmi d'autres, à la communauté française exilée qui ne peut pas parler le français dans la vie quotidienne. Pour des raisons différentes, les deux lexicographes sont familiers avec les doutes de ceux qui (ré)apprennent une langue. Il n'est donc pas surprenant qu'ils fassent des choix lexicographiques similaires. Les intentions pédagogiques ressortent à partir de l'orthographe. Tous deux suppriment les lettres qui ne se prononcent pas, en veillant à ne pas *défigurer* les mots, c'est-à-dire à en préserver l'étymologie. En ce sens, la position de Basnage est très forte car elle implique la réécriture de l'ensemble du dictionnaire de Furetière, dont l'orthographe est, par ailleurs, quasiment la même que celle des académiciens. Quant aux citations d'auteurs, seuls Richelet et Basnage les utilisent, parmi les grands lexicographes de l'époque classique ; si d'une part elles garantissent l'exhaustivité du travail lexicographique, d'autre part elles constituent un support valable pour les apprenants car elles offrent des exemples d'utilisation ainsi que des modèles linguistiques à imiter. Toujours dans le but d'aider les apprenants, les deux lexicographes enregistrent un grand nombre de constructions syntaxiques, de flexions verbales, des collocations et de règles grammaticales. L'approche descriptive est un autre trait commun aux deux lexicographes : ni Richelet ni Basnage ne s'érigent en prescripteurs de certaines formes ou constructions. Plutôt que de proposer uniquement leur forme préférée, ils enregistrent tous les usages linguistiques dont ils ont connaissance et suggèrent tout au plus leur préférence en s'appuyant sur l'autorité d'écrivains et des experts en langues qui ont utilisé ou se sont prononcés en faveur d'un usage spécifique. Tout en prêtant attention aux éléments qui facilitent l'apprentissage de la langue, les deux lexicographes cherchent donc à répondre à leur besoin d'exhaustivité dans la description linguistique.

Dans ce qui suit, après avoir montré les analogies entre le *Dictionnaire françois* de RICHELET 1693 et le *Dictionnaire universel* de Basnage [DU1701] 1701, notamment en ce qui concerne l'orthographe, les renseignements sur la prononciation, les citations, les marques de domaine et d'usage, le traitement des mots grammaticaux, les flexions verbales, les constructions syntaxiques et les collocations, nous tenterons de trouver les raisons de la dissimulation de la source.

<sup>2</sup>Pour plus de détails sur les orientations didactique et sociolinguistique du *Dictionnaire françois* (1680) de Richelet cf. QUÉMADA 2004 : 7.

<sup>3</sup>Richelet « a donné des cours particuliers de français à des étrangers distingués, principalement allemands, anglais ou scandinaves, qui séjournèrent à Paris et avec lesquels il s'est lié d'amitié » (QUÉMADA 2004 : 2).

## 2 Orthographe

En feuilletant le *Dictionnaire universel* de Furetière et celui de Basnage, la première chose qui saute aux yeux est que l'orthographe adoptée par les deux lexicographes est différente. Furetière était un académicien, dont l'avis a certainement influé sur les choix orthographiques du *Dictionnaire de l'Académie française* (1694)<sup>4</sup>. Par conséquent, il n'est pas étonnant que l'orthographe adoptée par Furetière est *grosso modo* la même que celle des académiciens. Basnage est le rédacteur en chef d'un journal ouvert sur le monde et l'actualité, qui donne la parole aux contemporains et qui défend souvent des positions novatrices, y compris en matière d'orthographe. Basnage affirme s'être tenu à l'orthographe « de Furetière, qui est aussi celle de Mrs. de l'Académie française » (Préface, DU1701). Il fait ainsi référence à la macrostructure du dictionnaire qui présente « les mots dans le même ordre, & la même disposition » (ibid.) que le *Dictionnaire universel* de Furetière<sup>5</sup>. Néanmoins, il reconnaît avoir supprimé less « en beaucoup de mots où elle ne se prononce point » (ibid.). Il fait ainsi référence à la microstructure qui présente une orthographe assez différente par rapport à celle de son prédécesseur. Bien que le seul changement envisagé dans la préface soit la suppression de *s*, Basnage applique trois types de modifications à l'orthographe de la première édition du dictionnaire : la suppression, le remplacement et, plus rarement, l'ajout de caractères. Les changements concernent non seulement les lettres, mais aussi les signes diacritiques. Lorsque Basnage commence la révision du dictionnaire, la discussion sur l'orthographe en France est ouverte.

Conscient des débats en cours, Basnage se déclare attentif à préserver l'étymologie des mots qu'il craint de défigurer ou déguiser. Furetière, après la définition du lemme ORTHOGRAPHE, précise que ceux qui ont essayé, à la suite de « Jacques Pelletier du Mans, qui soutint qu'il falloit écrire comme on parle » (ORTHOGRAPHE, DU1690), de modifier l'orthographe du français ont été mal reçus. Il s'agit de « Louïs Maigret, Pierre la Ramée dit Ramus, Jean Antoine de Baïf, et de notre temps l'Esclache » (ibid.). Basnage efface la référence à la réception des avis en matière de changement d'orthographe ainsi que celle au contemporain l'Esclache<sup>6</sup>. Louis de L'Esclache est un professeur de philosophie qui dénonce l'exclu-

---

<sup>4</sup>Les choix orthographiques du Dictionnaire de l'Académie sont décrits dans les Cahiers de Mézeray. Bien que l'Académie déclare dans la préface du dictionnaire qu'elle souhaite maintenir l'ancienne orthographe « par tout », les Cahiers contiennent « en partie, une condamnation de l'orthographe traditionnelle » (CATACH & ROSIER-CATACH 2001 : 186).

<sup>5</sup>Il s'agit d'une décision de convenance : changer l'orthographe des mots vedettes aurait entraîné une perturbation de leur agencement macrostructurel.

<sup>6</sup>Dans « les systèmes extrêmes de phonéticiens » (CATACH & ROSIER-CATACH 2001 : 120), chaque son correspond à un et un seul signe.

sion des femmes de la société et demande que la langue française soit émancipée du joug de l'étymologie. Le titre de son traité, *Les Véritables règles de l'orthographe Franceze, ou l'art d'aprandre an peu de tams à écrire còrectemant*, donne déjà une idée de certaines des positions qu'il défend comme le remplacement des *en* et *em* par *an* et *am*. Basnage, se prononce en faveur d'une orthographe qui, bien que novatrice, respecte l'étymologie des mots. C'est précisément au titre du traité de L'Esclache qu'il semble se référer lorsqu'il affirme que les auteurs qui écrivent « tans » ou « ortographe » semblent défigurer les mots « temps » et « orthographe » en empêchant de reconnaître leurs origines latine et grecque respectivement. Il n'est donc pas étonnant que Basnage ait supprimé la mention de cet auteur. Basnage écrit que « la chose est encore en contestation, & on n'a pu convenir d'une manière uniforme d'orthographe » (ORTHOGRAPHE, DU1701). Il est conscient de l'existence d'au moins trois positions : celle, dépassée, de ceux qui entendent écrire comme on parle ; celle qui reste ancrée dans la tradition, comme celle des académiciens, qui « ont retenu l'ancienne orthographe, & ont gardé les lettres mêmes qui ne se prononcent point » (ibid.), et enfin celle qui adopte « un tempérament entre l'une, & l'autre orthographe, en retranchant les lettres inutiles, comme le s, dans une infinité de mots, & en conservant pourtant dans les mots l'orthographe qui marque leur origine » (ibid.). De caractère conciliant, après avoir affirmé qu'« on peut donc choisir entre les manières d'orthographe » (ibid.), Basnage déclare préférer la dernière position puisque « le plus sûr est de ne point trop s'éloigner de l'ancien usage pour ne pas trop changer les mots aux yeux, en les voulant accomoder à l'oreille » (ibid.). Parmi ceux qui ont adopté une position intermédiaire entre l'ancienne et la nouvelle orthographe, il y a Richelet, que Basnage ne mentionne pas dans l'entrée ORTHOGRAPHE. En décrivant l'orthographe adoptée dans son dictionnaire, Richelet explique sa position dès la première édition de son *Dictionnaire françois* (RICHELET 1680) :

Touchant l'Ortographe, on a gardé un milieu entre l'ancienne, & celle qui est tout à fait moderne, & qui défigure la Langue. On a seulement retranché de plusieurs mots les lettres qui ne rendent pas les mots méconnoissables quand elles en sont otées, & qui ne se prononçant point, embarassent les Etrangers, & la plu-part des Provinciaux. » (RICHELET 1680 : Avertissement, *Dictionnaire François*)

La position de Basnage est près de celle de Richelet qui propose d'éliminer les lettres qui ne se prononcent pas, tout en faisant attention à ne pas *défigurer* la langue. Le souci de Richelet pour son public est explicite puisqu'il affirme que les lettres qui ne sont pas prononcées « embarassent les Etrangers » (Avertissement,

*Dictionnaire françois*, 1680). Bien que Basnage ne mentionne pas explicitement ses lecteurs, il est possible de discerner dans l'adoption d'une orthographe simplifiée la volonté de leur rendre service. A différence de Richelet, Basnage accorde plus d'attention à la préservation des étymologies. Par exemple, là où Richelet écrit *ortographe*, Basnage invite à conserver le *h* car cette lettre est la marque de la dérivation grecque du mot.

Au-delà des affirmations des préfaces, outre la suppression des *s* pré-consonnes Richelet et Basnage partagent au moins deux autres positions orthographiques dans les corps de leurs ouvrages. La première est la syncope de *t* des pluriels des adjectifs, voire substantifs, en *-ant* et *-ent* qui dérivent du participe présent des verbes latins, ainsi que des substantifs composés du suffixe *-ment*. Alors que Furetière garde la dentale entre nasale et sibilante (ex. : *absents, compliments, correspondants*)<sup>7</sup>, Richelet et Basnage l'effacent systématiquement (ex. : *absens, complimens, correspondans*). La deuxième position que Richelet et Basnage partagent concerne la préférence d'*i* à *y* à la fin des mots que Furetière orthographie avec un *y*, comme les pronoms personnels complément et leurs dérivés (ex. : *moi, toi, celui* au lieu de *moy, toy, celuy*). Basnage déclare qu'« il n'y a point de raison de preferer l'y Grec à un *i* simple » (Y, DU1701) au moins à la fin des mots où « ce n'est que pour observer la vieille orthographe qu'on écrit, *Roy, Loy, & [...] j'ay, j'allay* » (ibid.). Cette position est essentiellement celle de Richelet qu'« à l'imitation de l'illustre Monsieur d'Ablancourt, *Preface de Thucidide, Apophtegmes des Anciens, Marmol*, etc. et de quelques Auteurs célèbres, [...] change presque toujours l'y grec en *i* simple » (RICHELET 1680 : *Avertissement*).

### 3 Prononciation

L'indication de la prononciation des mots augmente fortement dans le *Dictionnaire universel* de Basnage 1701 par rapport à l'édition précédente. Furetière enregistre la prononciation d'environ 50 mots (cf. REY 2020) tandis que Basnage d'environ 800, tout comme Richelet dont il s'inspire souvent.

La plupart des prononciations enregistrées par Furetière concernent des mots contenant un groupe de consonnes commençant par la lettre *s*. Comme il adopte l'ancienne orthographe qui enregistre les lettres d'origine latine, et notamment *s*, même lorsqu'elles ne sont pas prononcées, Furetière sent parfois le besoin de fournir des indications sur les endroits où cette lettre est ou n'est pas prononcée. Cependant, le lexicographe enregistre également des informations sur d'autres

---

<sup>7</sup>Furetière n'est pas toujours irréprochable dans cette position. Par exemple, le DU1690 enregistre 88 occurrences d'« amants » et 37 d'« amans ».

lettres, notamment *c* et *d*. En général, Basnage maintient les informations fournies par Furetière en en introduisant de nombreuses d'autres. Les formulations de Furetière sont impersonnelles ou ont pour sujet la lettre en question tandis que celles de Basnage s'adressent souvent au lecteur par le biais de la deuxième personne du pluriel de l'impératif du verbe *prononcer*. Basnage indique la prononciation de différentes manières, parmi lesquelles on peut reconnaître trois schémas qui se combinent parfois entre eux :

- L'invitation à lire ou à ne pas lire une des lettres contenues dans le mot-vedette, par exemple : *CHRIST [...] Prononcez l's*.
- L'indication de la prononciation du mot entier, par exemple : *AIGUADE [...] Prononcez EGADE*.
- L'indication de la quantité syllabique d'une syllabe du mot, par exemple : *PATE [...] La première syllabe de ce mot est breve, & c'est pour cela qu'il est mieux d'écrire avec l'Academie, Patte*.

Les trois exemples donnés ci-dessus sont tirés de Richelet, y compris le dernier dans lequel l'Académie française est mentionnée pour confirmer ce qui vient de Richelet qui est occulté par Basnage. Bien qu'il s'inspire souvent de Richelet, Basnage modifie fréquemment les formulations introduisant les renseignements sur la prononciation. Alors que les formulations de Richelet sont presque toujours introduites par « Prononcez »<sup>8</sup>, l'alternance de formulations personnelles et impersonnelles montre chez Basnage<sup>9</sup> une plus forte tension entre la volonté d'offrir un outil qui guide le lecteur au moyen des prescriptions et une autre volonté, tout aussi forte, d'offrir une image objective de la langue telle qu'elle émerge des pratiques des locuteurs.

#### 4 Mots grammaticaux

Bien que Furetière prenne en compte le domaine de la grammaire<sup>10</sup>, ses définitions des mots grammaticaux sont concises et fournissent peu d'exemples illustratifs. Basnage, pour sa part, au-delà de l'augmentation quantitative des articles

---

<sup>8</sup>Richelet utilise « Prononcez » dans 636 entrées, « se prononce » dans 82 entrées, « on prononce » dans 93 entrées.

<sup>9</sup>Basnage utilise « Prononcez » dans 333 entrées, « se prononce » dans 258 entrées, « on prononce » dans 108 entrées.

<sup>10</sup>Furetière utilise 84 fois la marque de domaine « Terme de Grammaire ». Il décrit également 86 mots plurisémiqiques qui trouvent une application « en termes de Grammaire » (37) ou « en Grammaire » (49).

consacrés à la grammaire - 206 au total<sup>11</sup> contre les 170 de Furetière -, accorde une large place aux règles grammaticales, notamment dans les entrées des mots grammaticaux. Basnage est conscient, comme il le dit dans la préface de son édition, du fait que « les minuties de Grammaire dont ce Dictionnaire est rempli, ne sont pas fort essentielles pour parler, quand on ne parle que pour se faire entendre » (Préface, *DU1701*). Par contre, elles sont précieuses pour ceux qui cherchent à parler « exactement, poliment, & noblement » (ibid.). Les règles grammaticales introduites par Basnage dans le *Dictionnaire universel* de Furetière proviennent souvent du *Dictionnaire françois* de Richelet ainsi que de grammaires et recueils de remarques sur la langue française. Par exemple, dans l'entrée JE, fortement augmentée, Basnage explicite par le moyen de règles et d'exemples les informations que Furetière avait données sous forme implicite, et les complète par d'autres informations de caractère phonétique et morphologique. Il introduit des remarques de MÉNAGE (1672 : 115) et de RICHELET (1693) (JE), reprises à leur tour dans la *Grammaire générale et raisonnée* de ARNAUD & LANCELOT (1660 : 35). Basnage insère des renvois extratextuels à « Ménage, & la Grammaire raisonnée » (JE, *DU1701*), qu'il invite à consulter mais pas à Richelet, dont il occulte encore une fois la consultation.

## 5 Citations d'écrivains

Les citations d'écrivains sont marginales dans le dictionnaire de Furetière et « y répondent avant tout à des intentions littéraires » (LEHMANN 1995 : 49). Au contraire, dans le dictionnaire de Basnage, elles jouent un rôle de premier plan, puisqu'elles figurent dans toutes les entrées pour lesquelles le lexicographe a pu en trouver.<sup>12</sup>

Les citations de Basnage ne sont pas toutes de première main. Le lexicographe transcrit parfois les citations que d'autres lexicographes avaient incluses dans leurs dictionnaires. Richelet est la source majeure des citations de Basnage. À l'exception de quelques-uns, tous les auteurs cités dans la « table alphabétique de la plupart des Auteurs et des Livres citez dans ce dictionnaire » au début du *Dictionnaire françois* de RICHELET (1693) sont présents dans le *Dictionnaire universel* de Basnage. Toutefois, seuls certains d'entre eux – cinquante-un plus précisément -

---

<sup>11</sup>Basnage utilise 114 fois la marque de domaine « Terme de Grammaire ». Il décrit également 92 mots plurisémiques qui s'appliquent « en termes de Grammaire » (41) et « en Grammaire » (51).

<sup>12</sup>La plupart des entrées terminologiques manquent de citations car les termes sont rarement utilisés par les écrivains.

figurent dans la liste des abréviations de Basnage. Cela ne s'explique que partiellement par le fait que Basnage inclut dans la liste les seuls auteurs qu'il cite au moyen d'une abréviation tandis qu'il ne mentionne pas les auteurs dont il donne le nom complet. En fait, bien que beaucoup d'auteurs de la table alphabétique de Richelet soient cités par Basnage au moyen d'une ou plusieurs abréviations, ils ne figurent pas dans sa liste des abréviations. Probablement, il s'agit d'un oubli. Au moment de la compilation de la table des abréviations, le manque de recherches personnelles empêche Basnage de se souvenir de tous les auteurs et de toutes les œuvres pour lesquels il s'est appuyé sur les recherches d'autres. Alors que Richelet indique souvent ses sources de manière très précise, fournissant en plus du nom de l'auteur également le titre de l'œuvre et la page où la citation se trouve, Basnage est beaucoup plus hâtif et, en même temps, systématique, en ne fournissant que l'abréviation du nom de l'auteur. Néanmoins, bien que « la dette de Basnage envers Richelet [soit] grande, [...] il a lu tous ces auteurs ou presque (une exception semble Voiture) » (FENNIS 1988 : 86). Le fait que les citations identifiées par Basnage et ses collaborateurs s'additionnent à celles de Richelet et de ses collaborateurs explique le nombre très élevé de citations de d'Ablancourt et de Boileau<sup>13</sup> ainsi que de Saint-Évremond<sup>14</sup> et de ses *Œuvres mêlées*. Il est très probable que, s'inspirant de la méthode de travail de Richelet qui se faisait aider par des collaborateurs dans la recherche de citations,<sup>15</sup> Basnage se soit fait envoyer les citations par les mêmes écrivains qu'il cite.

L'entrée consacrée au pronom relatif *que*, réécrite à partir du *Dictionnaire françois* (1693) de Richelet, montre que Basnage fait aveuglément confiance aux informations de sa source qu'il ne vérifie pas toujours :

QUE. Pronom relatif & indeclinable, qui perd son e devant une voyelle, & qui s'emploie diversement. Par exemple, c'est un mechant que le Ciel châtie. Arn. *Que* est mis là pour *lequel*. Ce n'est pas à toi que je voudrois deguiser la verité. Ablan. *Que* est mis là pour *auquel*, ou à qui. Les mechans sont semblables à ces petites pailles que le vent emporte. Port-R. *Que* est mis là pour *lesquelles*.

Après avoir remplacé la définition de Furetière<sup>16</sup> avec celle de Richelet qu'il ne cite pas, Basnage transcrit même les citations du *Dictionnaire françois* (1693).

<sup>13</sup>Richelet le cite comme « Dépreaux ».

<sup>14</sup>Charles de Marguetel de Saint-Denis, seigneur de Saint Evremond (1616–1703), était un correspondant de Basnage et deux de ses lettres de Londres ont été même publiées dans le HOS en décembre 1698 et mars 1699 (cf. GRAVELEAU 2018 : 743).

<sup>15</sup>Cf. QUEMADA 1967 : 543.

<sup>16</sup>« Pronom relatif & interrogatif » (QUE, DU1690).

En général, Basnage utilise l'abréviation « Port. R. » pour se référer au traité de Claude Lancelot, *Nouvelle Méthode pour apprendre facilement la langue latine* (LANCELOT 1662), mais dans le cas en espèce l'abréviation désigne ladite « Bible de Port Royal » ou « Bible de Sacy »<sup>17</sup>, d'où proviennent les exemples du *Dictionnaire françois* de RICHELET (1693). En somme, Basnage reprend les citations de Richelet sans se demander si le collègue attribue la même valeur à l'abréviation que lui-même.

## 6 Marques de domaines et marques d'usage

Les marques de domaine et d'usage<sup>18</sup> sont utilisées par Furetière plus largement que les citations<sup>19</sup>. Cependant, elles semblent être extrêmement peu nombreuses par rapport à la quantité de marques de domaine et d'usage enregistrées par Basnage. Il déclare que dans son ouvrage la distinction entre *mots* et *choses* « ne subsiste plus » (Préface, *DU1701*). Si Furetière entendait décrire essentiellement les sciences et les arts et « le reste n'étoit qu'accessoire » (ibid.), dans son *Dictionnaire universel* Basnage vise à « apprendre à parler poliment, aussi bien qu'à parler juste, & dans les termes propres à chaque Art » (ibid.). Richelet est le premier lexicographe à avoir appelé le sien un « dictionnaire des mots et des choses » (QUEMADA 1967 : 85).

Les marques de domaine et les marques d'usage permettent aux lexicographes de signaler les mots qui s'écartent du *bel usage* de la langue française. Elles sont appelées *nottes* par les lexicographes de l'âge classique. Comme on lit dans les deux éditions du *Dictionnaire universel*, « dans un Dictionnaire on doit mettre une *notte* à un mot, quand il est vieux, ou particulier à quelque art ou science. Quand il est dans l'usage commun, il n'y faut point de *notte* » (NOTTE). Cet exemple décrit la fonction des marques de domaine et d'usage : les premières, placées après les mots-vedettes, détectent les langues spécialisées en indiquant le domaine auquel appartient l'entrée, et les deuxièmes, placées en fin d'entrée, renvoient aux dimensions de la variation linguistique et à la fréquence d'usage des mots. Basnage augmente considérablement les unes et les autres. Le domaine est introduit par diverses formulations, dont les plus courantes sont « (en) terme(s) de [nom

---

<sup>17</sup>Louis-Isaac Lemaistre De Sacy est le principal auteur de l'ensemble des traductions de la Bible en français, publiées entre 1665 et 1696 par les imprimeurs de Port-Royal.

<sup>18</sup>Par 'marque d'usage', nous entendons exclusivement les indications de nature rhétorique et sociolinguistique placées à la fin de l'entrée, et non les déterminations de domaine qui se trouvent immédiatement après la catégorie grammaticale que nous appelons 'marques de domaine'.

<sup>19</sup>Cf. REY (1990).

du domaine] ». Quant aux marques d'usage, bien que l'entrée NOTTE ne mentionne que des archaïsmes, le *Dictionnaire universel* de Basnage décrit un spectre assez large de la variation linguistique. Le lexicographe ne néglige la description d'aucune dimension du langage, à savoir diachronique, diatopique, diastratique, diaphasique et diamésique (cf. COSERIU (1977)). La description de ces différentes dimensions n'est pas systématique et uniforme. Les marques d'usage sont introduites par « ce mot » dans un certain nombre d'entrées, par le pronom personnel « il » dans d'autres, par des périphrases impersonnelles telles que « on dit » ou « on s'en sert » dans d'autres encore. Parfois, elles se réfèrent à tous les sens du mot, parfois au dernier sens seulement. Dans certains cas, Basnage le précise au moyen du complément de limitation « en ce sens », dans d'autres il ne le fait pas.

Furetière tente de regrouper les entrées sous un nombre plus restreint de domaines par rapport à Basnage, en offrant un texte plutôt systématique. De son côté, Basnage ne se soucie pas de préserver la cohérence du texte de Furetière et multiplie les marques de domaine en introduisant de nombreuses désignations de catégories de personnes auxquelles il n'attribue souvent qu'un seul terme. Elles sont généralement reprises du *Dictionnaire françois* de (RICHELET 1693). Si chez Furetière, qui toutefois enregistre plus de 250 métiers et activités, il y a un « début de systématisation » QUEMADA (1967 : 306), Basnage le brise car son *Dictionnaire universel* contient une quantité bien plus imposante de marques de domaine provenant tout droit du *Dictionnaire françois* RICHELET (1693). Cela implique que, bien que moins systématiques, Richelet et Basnage sont plus précis dans la représentation des catégories de travailleurs et de travailleuses. Par exemple, lorsque Furetière ignore les désignations féminines, Richelet et Basnage les introduisent souvent. Dans la société décrite par les deux lexicographes les femmes travaillent, produisent et vendent, comme en témoignent les termes de *blanchisseuses, bouquetiere, coiffeuse, couturiere, couturiere en linge, faiseuse de point, faiseuse de point & de dentelle, fileuse, fileuse au roüet, vendeuse de sel*. En outre, les domaines de Richelet et Basnage contiennent souvent une référence à Paris à laquelle sont associés de termes liés aux métiers ou aux bâtiments de Paris tels que *batelier de Paris, chaudronnier de Paris, college de Paris, executeur de Paris, hôpital general de Paris, libraires de Paris, marchand de bois de Paris, Nôtre Dame de Paris, oiselier de Paris, rotisseur de Paris, seminariste de Paris, theologien de Paris, Université de Paris*. Basnage, d'origine normande, exilé aux Provinces-Unies, a peu de liens avec Paris, et ne peut pas connaître le vocabulaire utilisé par les différentes catégories de travailleurs de la ville. Toutes les désignations renvoyant à Paris proviennent de Richelet qui, au contraire habite à Paris et « pour préciser ses définitions, va descendre dans la rue, il va consulter les gens de métier » (BRAY 1990 : 49). L'absence de réflexion personnelle dans la sélection des domaines est confirmée par

le fait que pour certaines de ces désignations, il n'existe aucune entrée dans le *Dictionnaire universel* de Basnage.

En ce qui concerne les marques d'usage, Basnage puise le plus souvent dans Richelet, puis dans POMEY (1664), MÉNAGE (1672), de LA TOUCHE (1696) et CORNEILLE (1694). Parfois il copie le texte de la source presque mot à mot, d'autres fois il le reformule. Dans quelques cas, Basnage aborde ses sources de manière critique. Par exemple, alors que Richelet affirme que « Ce mot ne se dit point » (INTOLERANCE, RICHELET (1693), Basnage soutient qu'il est utilisé depuis quelques années et par une catégorie spécifique de personnes, les théologiens<sup>20</sup>. Parfois Basnage ramène dans la même entrée des informations rencontrées dans des sources différentes. Dans un cas, après avoir affirmé qu'aucun bon auteur ne s'est servi du mot, ce qui reprend les mots de de LA TOUCHE (1696) qui à son tour cite BOISROBERT (1689), Basnage écrit qu'« à la vérité Mazeray l'a employé » (AFFLUER, DU1701) comme l'indique Richelet. De même, pour presque toutes les marques renvoyant aux styles burlesques, comiques et satyriques, Basnage s'appuie sur les mots de Richelet. Voici quelques exemples :

BARILLET. s. m. [...] Il est hors d'usage en ce sens.

LANTIPONAGE. s. m. Mot bas & burlesque, [...]

PARFAUTE [...] Ce mot est un peu vieux.

RENGAINE [...] Mot tout-à-fait bas & du petit peuple de Paris.

Toutes ces marques de fréquence et d'usage sont tirées de Richelet. Néanmoins, Richelet n'est pas la seule source des marques d'usage de Basnage qui ressent le besoin de fournir également les prescriptions de l'Académie française ainsi que d'autres lexicographes, grammairiens et experts en langue dans le but d'offrir un outil qui décrit la langue telle qu'elle est, dans toutes les situations et contextes de communication possibles. Parfois Basnage fait aussi confiance à sa propre conscience linguistique pour créer des marques d'usage.

## 7 Flexions verbales

Alors que Furetière ne donne des conjugaisons verbales qu'implicitement à travers des exemples illustratifs, Basnage donne des conjugaisons verbales de manière explicite pour environ 240 verbes. Elles sont le plus souvent tirées du *Dictionnaire françois* de RICHELET (1693). Si l'on considère le coût du papier, il est très

---

<sup>20</sup>Cf. INTOLERANCE, DU1701.

probable que les lexicographes aient sélectionné les seules formes irrégulières de verbes qui, à leurs avis, ne sont pas déductibles de l'infinitif lemmatisé. Même si fonctionnel, ce choix se traduit par une présentation chaotique des personnes, des temps et des modes verbaux. Alors que Richelet n'introduit pas les flexions verbales au moyen de formules, chez Basnage la conjugaison est précédée quelquefois par « Conjuguez » qui peut être abrégé en « Conj. », quelquefois par « On conjugue » qui peut être abrégé en « On conj. », quelquefois par « Ce verbe se conjugue ainsi ».

En général, Basnage tend à réduire le nombre de formes verbales qu'il trouve chez Richelet. Voir l'exemple où les formes du *Dictionnaire françois* (1693) de Richelet qui ne se trouvent pas dans le *Dictionnaire universel* de Basnage (FURETIÈRE 1701) sont indiquées entre crochets. :

ABSTENIR v. n. [...] Conjuguez Je m'*abstiens*; je m'*abstenois*; je m'*abstins*; [je me suis abstenu]; je m'*abstiendrai*; [abstiens toi; qu'il s'abstienne]; je m'*abstiendrois* [Je m'abstinsse, que je me suis abstenu] &c.

## 8 Constructions syntaxiques

Contrairement à Furetière, qui ignore quasiment les constructions syntaxiques, Basnage leur donne une large place. Elles sont quelquefois introduites par le verbe *joindre*<sup>21</sup>, dont s'était sporadiquement déjà servi Furetière. Cependant, Basnage utilise plus fréquemment la forme « se construit » pour indiquer principalement des constructions verbales. Un autre introducteur est la forme « regit », utilisée dans la plupart de cas pour décrire le régime des conjonctions et prépositions qui sont lemmatisées par Basnage de manière beaucoup plus systématique que par Furetière. Exceptionnellement, le régime des autres parties du discours comme les adjectifs, les adverbes et les substantifs est donnée par Basnage. En général, en ce qui concerne les verbes, Basnage propose parfois des verbes auxiliaires, parfois la construction pronominale, parfois la préposition qui suit le verbe, parfois la diathèse. En ce qui concerne les prépositions, il indique le cas modulé sur la déclinaison latine (génitif, datif, accusatif, ablatif). Concernant les conjonctions, Basnage précise généralement le mode verbal qui les accompagne (indicatif, infinitif, subjonctif). Bien qu'il adopte une approche plus descriptive que son prédécesseur, Basnage indique aussi des constructions à éviter. Le régime des parties du discours est repris presque toujours du *Dictionnaire françois* de RICHELET (1693) et, plus rarement, du *Dictionnaire de l'Académie française*

<sup>21</sup>Ex. « Il se joint souvent avec le pronom relatif » (ACCREDITER, DU1701)

(ACADÉMIE FRANÇAISE 1694). Dans les entrées suivantes nous avons souligné les mots qui proviennent du *Dictionnaire françois* de Richelet :

AFIN. Conjonction qui denote l'intention, & signifie, Pour, à dessein. Elle est toujours suivie d'un de ou d'un que. Quand elle est suivie d'un de, elle regit l'infinif

AUPRÉS. Preposition qui marque un lieu proche, & qui regit le genitif.

PASSIONNER [...] Le mot de passionner [...] dans le regime actif est assez nouveau.

## 9 Collocations

Les lexicographes de l'âge classique sont très conscients du fait que l'usage de certains mots dépend de la présence d'autres et que le sens d'un mot peut varier en fonction des mots qui l'entourent. Furetière indique quelques collocations que Basnage garde en en précisant d'autres. En général, il informe le lecteur que l'utilisation d'un mot tout seul est sporadique, tout en assignant à des exemples et des citations la tâche de montrer ses collocations. Basnage reprend souvent les collocations de Richelet. Voir quelques exemples de collocations transcrites du *Dictionnaire françois* de Richelet :

BATTEUR. s. m. Ce mot ne se dit presque jamais seul. [...] Les batteurs de gens trouvent souvent de plus mechans qu'eux qui les battent.

EMBLÉE. s. f. [...] Ce mot ne se dit pas seul. [...] Ce Capitaine étoit heureux à prendre les villes d'emblée. La ville étoit trop bien munie pour l'emporter d'emblée. Vaug.

GAGNEUR. s. m. Ce mot [...] ne se dit pas seul, & même il ne se dit guere. Ce gagnneur de tant de batailles. Voit.

## 10 Conclusions

Basnage puise dans le *Dictionnaire françois* de RICHELET (1693) de nombreuses informations afin d'enrichir son *Dictionnaire universel* (1701). En plus de transcrire des sens et des entrées entières que Furetière avait négligés, Basnage reprend des informations concernant la prononciation, la grammaire, les citations d'écrivains, les marques d'usage, les flexions verbales, les constructions syntaxiques, les collocations. Même l'orthographe simplifiée de Basnage ressemble celle de Richelet.

Tous ces éléments sont inhérents à un outil conçu pour soutenir l'apprentissage du français. En même temps, ils permettent aux deux lexicographes d'offrir une description large et objective de la langue. Néanmoins, Basnage ne reconnaît pas sa dette envers le collègue. Il est plausible qu'afin de plaire aux membres de l'Académie française qui, comme le dit Furetière dans le premier *Factum*, « traitent de misérable » (FURETIÈRE 1859 : 9), Richelet, Basnage ne reconnaît pas le rôle de Richelet en évitant de le citer et en le dépréciant lorsqu'il le cite. Par ailleurs, le choix de l'édition de 1693 n'est pas fortuit. D'abord, elle n'avait pas été utilisée par Furetière et permettait donc à Basnage de mettre à jour le *Dictionnaire universel* avec les informations d'une source plus récente que celui-là. Par ailleurs, le choix de l'édition de 1693 n'est pas fortuit. En n'étant pas utilisée par Furetière elle permet à Basnage de mettre à jour le Dictionnaire universel en peu de temps avec une grande quantité d'informations récentes, sélectionnées et ordonnées.

## Acknowledgement

Ce travail a été réalisé dans le cadre du projet ANR-18-CE38-0003 *Numérisation et analyse du Dictionnaire universel de Basnage de Beauval : lexicographie et réseaux scientifiques – BASNUM*

## Références

- ACADÉMIE FRANÇAISE. 1694. *Dictionnaire de l'Académie*. Paris : Coignard.
- ARNAUD, Pierre. & Claude LANCELOT. 1660. *Nouvelle méthode pour apprendre facilement... La langue italienne [par C. Lancelot]*. Paris : P. Le Petit.
- BOISROBERT, André. 1689. *Réflexions, ou Remarques critiques sur l'usage present de la langue françoise*. Paris : D'Houry.
- BRAY, Laurent. 1990. Les marques d'usage dans le Dictionnaire françois (1680) de César Pierre Richelet. In *Lexique 9/ Les marques d'usage dans les dictionnaires (XVIIe, XVIIIe siècles)*. Villeneuve d'Ascq : Presses Univ. Septentrion.
- CATACH, Nina & Irène ROSIER-CATACH. 2001. *Histoire de l'orthographe française*. Paris : Champion.
- CORNEILLE, Thomas. 1694. *Le dictionnaire des arts et des sciences*. Paris : Coignard.
- COSERIU, Eugenio. 1977. *Lezioni di linguistica generale (Didattica)*. Torino : Boringhieri.
- de LA TOUCHE, Pierre. 1696. *L'art de bien parler françois*. Amsterdam : Henri Desbordes. <https://gallica.bnf.fr/ark:/12148/bpt6k7816g> (17 mai, 2023).

- FENNIS, Jan. 1988. Les sources du vocabulaire maritime dans le Furetière de 1701. In Höfler M von BEMMINGEN B (éd.), *Travaux de linguistique et de philologie*, 75-93. Strasbourg-Nancy : Librairie Klincksieck.
- FURETIÈRE, Antoine. 1690. *Dictionnaire universel contenant généralement tous les mots françois tant vieux que modernes et les termes de toutes les sciences et des arts ... Par... Antoine Furetière...* 3 t. Amsterdam : Leers. <http://gallica.bnf.fr/ark:/12148/bpt6k3413126b> (17 mai, 2023).
- [DU1701]. 1701. *Dictionnaire universel, contenant généralement tous les mots françois tant vieux que modernes & les termes des sciences et des arts,.... Tome 1 / . Recueilli & compilé par feu messire Antoine Furetière,.... Seconde édition revue, corrigée & augmentée par Monsieur Basnage de Beauval.* Amsterdam : Leers. <https://gallica.bnf.fr/ark:/12148/bpt6k5841680f> (17 mai, 2023).
- FURETIÈRE, Antoine. 1701. *Dictionnaire universel, contenant généralement tous les mots françois tant vieux que modernes & les termes des sciences et des arts,.... Tome 1 / . Recueilli & compilé par feu messire Antoine Furetière,.... Seconde édition revue, corrigée & augmentée par Monsieur Basnage de Beauval.* Amsterdam : Leers. <https://gallica.bnf.fr/ark:/12148/bpt6k5841680f> (17 mai, 2023).
- FURETIÈRE, Antoine. 1859. *Recueil des factums d'Antoine Furetière, de l'Académie française, contre quelques-uns de cette Académie; suivi des preuves et pièces historiques données dans l'édition de 1694. Tome 1 / avec une introduction et des notes historiques et critiques, par M. Charles Asselineau,...* Charles ASSELINEAU (éd.). Paris : Poulet-Malassis et de Broise. <https://gallica.bnf.fr/ark:/12148/bpt6k58321422> (19 mai, 2023).
- GRAVELEAU, Sara. 2018. "Les hérésies sont d'utiles ennemies" : *Itinéraire d'Henri Basnage de Beauval (1656-1710), avocat de la République des Lettres et penseur de la tolérance civile.* Université d'Angers. (thèse de doct.). <https://theses.hal.science/tel-02157922> (17 avril, 2023).
- LANCELOT, Claude. 1662. *Nouvelle méthode pour apprendre... La langue latine [par C. Lancelot], 4e édition... Avec un traité de la poésie latine et une brève instruction sur les règles de la poésie française.* Paris : Pierre Le Petit. <https://gallica.bnf.fr/ark:/12148/bpt6k1510380q> (19 mai, 2023).
- LEHMANN, Alise. 1995. La citation d'auteurs dans les dictionnaires de la fin du XVIIe siècle (Richelet et Furetière). *Langue Française* 106. 35-54. DOI : 10.3406/lfr.1995.6442. <https://www.jstor.org/stable/41558721> (2 août, 2021).
- MÉNAGE, Gilles. 1672. *Observations de Monsieur Ménage sur la langue française.* Paris : C. Barbin.
- POMEY. 1664. *Dictionnaire Royal des langues française et latine.* Molin.
- QUEMADA, Bernard. 1967. *Les dictionnaires du français moderne : 1539-1863.* Paris : Didier.

- QUÉMADA, Bernard. 2004. Pierre Richelet (1626-1698) et le Dictionnaire françois de 1680. *Présentation de l'édition électronique du Dictionnaire françois de Richelet sur le site Classiques Garnier Numérique*. <https://num-classiques-garnier-com.janus.bis-sorbonne.fr..>
- REY, Alain. 1990. Le marques d'usage et leur mise en place dans les dictionnaires du XVIIe siècle : Le cas Furetière. In *Lexique 9/ Les marques d'usage dans les dictionnaires (XVIIe, XVIIIe siècles)*, 17-29. Villeneuve d'Ascq : Presses Univ. Septentrion.
- REY, Christophe. 2020. Du Dictionnaire universel de Furetière au Dictionnaire de l'Académie Française : deux traitements du marquage de la prononciation ? Journée d'études de l'Académie française - Rencontre avec Antoine Furetière (1619-1688) à l'occasion du quadricentenaire de sa naissance, Paris, Institut de France. 21 janvier 2020.
- RICHELET, Pierre. 1680. *Dictionnaire François. Contenant ... plusieurs nouvelles remarques sur la langue Française : ses expressions propres, figurées et burlesques, la prononciation des mots ... difficiles, le genre des noms ; le regime des verbes : avec les termes les plus connus des arts et des sciences*. Genève : Jean Herman Widerfold.
- RICHELET, Pierre. 1693. *Dictionnaire François. Contenant ... Plusieurs nouvelles remarques sur la langue Française : Ses expressions propres, figurées et burlesques, la prononciation des mots ... Difficiles, le genre des noms ; le regime des verbes : Avec les termes les plus connus des arts et des sciences*. Genève : D. Ritter.



# Chapitre 11

## De Félibien à Boutard : L'évolution du dictionnaire artistique entre le XVIIème et le début du XIXème siècle

Rosa Cetro

Université de Pise / PLIDAM, INALCO

Artistic terminology displays a hybrid and complex character, due to several factors (among others, the vastness of the field, the coexistence of erudite and ordinary technologists, the status of the arts over the centuries). The first lexicographic description of this terminology in the French language was written by the historiographer André Félibien. This chapter aims to show the evolution and the different variations of the artistic dictionary on a corpus of five monolingual French dictionaries covering a century and a half (1676–1826).

### 1 Introduction

Le dictionnaire monolingue français naît vers la fin du XVIIe siècle, avec la publication de trois ouvrages fondateurs, ce que PRUVOST (2003 : 463) appelle la « trilogie fondatrice », à savoir les dictionnaires de RICHELET (1680), de FURETIÈRE (1690) et la première édition du *Dictionnaire de l'Académie française* (1694). Cela est bien connu. Ce que l'on sait peut-être moins est que d'autres tentatives lexicographiques monolingues ont vu le jour avant ces ouvrages fondateurs, notamment dans des domaines spécialisés. C'est le cas, par exemple, du dictionnaire de termes d'art compilé par André Félibien (1619–1695) et annexé au traité *Des Principes de l'Architecture, de la Sculpture et de la Peinture et des autres Arts qui en dépendent* (1676). Signé par Félibien, cet ouvrage est cependant le fruit du travail de toute une institution, celle de l'Académie royale d'Architecture, fondée en 1671 et dont Félibien était secrétaire.



Malgré ses limites, ce dictionnaire s'avère être un observatoire intéressant pour l'étude des dynamiques qui accompagnent la mise en place du « chantier lexicographique en français » du XVIIe siècle, ainsi qu'un modèle de référence dans la lexicographie spécialisée sur l'art. Dans ce chapitre, nous voudrions retracer l'évolution – et les déclinaisons – du genre dictionnaire artistique en langue française de la fin du XVIIe siècle jusqu'au début du XIXe siècle, en nous appuyant sur un corpus de cinq dictionnaires représentatifs. Il s'agira, d'un côté, de relever les questionnements qui sous-tendent la description lexicographique mise en place par les différents auteurs et qui sont influencés à la fois par leur conception du vaste domaine artistique et par les publics cibles auxquels ces ouvrages s'adressent. De l'autre côté, une confrontation constante sera menée entre ces ouvrages spécialisés et la lexicographie généraliste.

Après avoir présenté les dictionnaires composant le corpus ainsi que leurs auteurs, nous analyserons les objectifs poursuivis par chaque ouvrage et les destinataires auxquels ils s'adressent, avant de passer à une comparaison entre leurs macrostructures et microstructures.

## 2 Présentation du corpus et méthodologie

Le corpus est composé des cinq dictionnaires suivants : le précité Dictionnaire annexé au traité *Des Principes de l'Architecture, de la Sculpture et de la Peinture et des autres Arts qui en dépendent* (1676) d'André Félibien, le *Dictionnaire abrégé de peinture et d'architecture* (1746) de François Marie de Marsy, le *Dictionnaire portatif des beaux-arts* (1752) de Jacques Lacombe, le *Dictionnaire portatif de peinture, sculpture et gravure, avec un traité pratique sur les différentes manières de peindre* (1757) d'Antoine-Joseph Pernety et le *Dictionnaire des arts du dessin, la peinture, la sculpture, la gravure et l'architecture* (1826) de Jean-Baptiste-Bon Boutard.

Un rapide survol des titres nous montre la diversité du format de ces ouvrages et la variété des sous-domaines artistiques sur lesquels se focalisent les auteurs. Par ailleurs, il ne nous semble pas superflu de donner quelques informations biographiques concernant ceux-ci.

André Félibien (1619–1695) a été un homme de pointe du Grand Siècle. Nommé Historiographe des Bâtiments du Roi Soleil par le ministre Colbert, il a fait partie de trois académies royales (Académie des Inscriptions et des belles lettres, Académie de Peinture et de Sculpture et Académie royale d'Architecture). Spécialiste reconnu des beaux-arts et écrivain prolifique, il est surtout connu pour son cycle des *Entretiens sur les vies et les ouvrages des plus excellents Peintres* (1666–1688)<sup>1</sup>.

---

<sup>1</sup>Pour plus d'informations sur André Félibien et son dictionnaire, nous nous permettons de renvoyer à notre ouvrage (CETRO 2022).

Le Jésuite François Marie de Marsy (1710–1763), auteur du *Dictionnaire abrégé de peinture et d'architecture* (1746) en deux tomes, a été professeur de grammaire, d'humanités et de rhétorique au collège Louis-le-Grand de Sommervogel (SGARD 1999 : 690).

Jacques Lacombe (1724–1811), avocat de formation, a été un important libraire puis imprimeur dans la France des Lumières, proche des Philosophes (notamment de Voltaire). Il s'est occupé de la compilation des dictionnaires de remplissage de l'*Encyclopédie méthodique* pour Panckoucke (MARTIN & CHARTIER 1984 : 188). Son *Dictionnaire portatif des beaux-arts* (1752), un de ses premiers ouvrages et fruit de son intérêt pour la critique d'art, a connu un grand succès (SGARD 1999 : 550–551).

Antoine-Joseph Pernety (1716–1796) dit Dom Pernety a été un bénédictin mauriste défroqué, alchimiste et écrivain prolifique. Il est connu surtout pour avoir fondé le mouvement maçonnique des *Illuminés d'Avignon*. Comme on peut le deviner du titre, son *Dictionnaire portatif de peinture, sculpture et gravure, avec un traité pratique sur les différentes manières de peindre* (1757) reprend la formule adoptée par Félibien d'un dictionnaire associé à un traité. Cependant, ici la focalisation est sur la peinture, dont l'auteur retrace le statut dans l'Antiquité d'après des sources anciennes (par ex. Pline) avant d'aborder les différentes techniques picturales.

Pour finir, Jean-Baptiste-Bon Boutard (1771–1838), architecte, publie son *Dictionnaire des arts du dessin, la peinture, la sculpture, la gravure et l'architecture* (1826)<sup>2</sup> après une longue collaboration (1800–1823) avec le *Journal des Débats* en tant que publiciste sur les beaux-arts.

À l'exception de FÉLIBIEN (1676), que nous avons numérisé pour des travaux précédents (CETRO 2022), les dictionnaires du corpus ne sont pas exploitables de façon automatique, ce qui nous oblige à un dépouillement manuel. La comparaison entre les cinq ouvrages a été donc menée à partir d'un échantillon. En raison de la présence de termes-clés de ce domaine (comme *colonne* et *couleur*), le choix est tombé sur la lettre C.

### 3 Buts et destinataires des ouvrages

L'analyse des préfaces et des avertissements de ces ouvrages – là où ils sont présents – nous renseigne sur les objectifs poursuivis par leurs auteurs et sur les

---

<sup>2</sup>Dorénavant, les dictionnaires du corpus seront indiqués par le nom de l'auteur suivi de l'année de publication entre parenthèses : FÉLIBIEN (1676), MARSY (1746), LACOMBE (1752), PERNETY (1757) et BOUTARD (1826).

destinataires potentiels auxquels ils s'adressent. S'il est vrai que tous ont pour cible un public cultivé, appartenant aux élites françaises, il est tout aussi vrai que l'on peut assister à un élargissement du nombre de lecteurs touchés par ce genre de publication au fil des décennies, grâce aussi aux progrès de l'alphabétisation et à l'engouement pour les dictionnaires au XVIII<sup>e</sup> siècle. Comme le soulignent MARTIN & CHARTIER (1984 : 190)

le dictionnaire devient une des formes canoniques, et la plus dynamique, de l'accumulation du savoir

Ainsi passe-t-on de Félibien 1676, conçu pour un public très restreint – les commanditaires d'œuvres d'art (MAËS 2016 : 178) –, à un nombre de plus en plus vaste de lecteurs-amateurs qui désirent s'instruire sur ce sujet et faire usage des connaissances acquises dans la conversation mondaine. Les dictionnaires d'art deviennent ainsi de véritables « vecteurs d[e] sociabilité » (MAËS 2016 : 182).

Le seul ouvrage du corpus à ne pas offrir une préface ou un avertissement est MARSY (1746). FÉLIBIEN (1676), PERNETY (1757) et BOUTARD (1826) offrent une préface, tandis que LACOMBE (1752) propose au lecteur un avertissement précédé d'une épître aux beaux-arts. Dans le cas de (BOUTARD 1826), la préface n'est pas rédigée par l'auteur du dictionnaire mais par ses éditeurs. L'objectif de ces pages est de présenter les contenus des ouvrages et de justifier les choix opérés par les auteurs.

On peut identifier quatre parties principales dans la préface de FÉLIBIEN (1676). Dans le paragraphe d'ouverture, l'auteur explique que la terminologie artistique employée dans ses ouvrages précédents n'a pas été tout à fait comprise par ses lecteurs. Il a donc décidé d'en entreprendre la description pour faciliter l'accès au sens de ces mots. Puisqu'il est bien plus aisé de comprendre l'utilisation des termes lorsqu'ils sont insérés dans un discours, le projet initial de l'auteur de décrire ces termes à l'intérieur d'un dictionnaire se tourne vers la forme d'un traité suivi d'un dictionnaire. De ces deux premières parties ressort clairement la vocation pédagogique de l'ouvrage. Dans la troisième, l'accent est mis sur le service que cet ouvrage peut rendre au public visé dans la communication avec les artisans. Cette attention à la dimension communicationnelle justifie aussi la présence de nombreux termes appartenant aux technoclectes ordinaires (MESSAOUDI 2013), employés dans les ateliers. Pour finir, dans la quatrième se manifeste la volonté de l'auteur de fournir à la postérité un témoignage précieux sur les arts figuratifs pendant le Grand Siècle.

LACOMBE (1752) résume dans le premier paragraphe de son *Avertissement* les raisons d'être de l'ouvrage :

Le goût que le Public témoigne pour les BEAUX-ARTS, l'empressement avec lequel il saisit tout ce qu'on lui présente sur cet objet, enfin l'utilité, je dirai même la nécessité d'un Livre qui renferme les Recherches & les Connoissances d'un *Amateur*; tels sont les principaux motifs qui m'ont poussé à entreprendre cet Ouvrage.

Bien qu'il existe déjà des textes sur le même sujet, il estime que ceux-ci sont lacunaires à plusieurs égards, notamment dans le traitement de l'ensemble des beaux-arts. De plus, il choisit d'inclure aussi dans son dictionnaire – genre textuel qu'il qualifie de « commode » – des articles sur les plus grands artistes de tous les temps et de tous les pays ainsi que sur leurs chefs-d'œuvre, sans pour autant délaissier la description de la terminologie des beaux-arts. Comme nous le verrons dans la section suivante, sous cette dénomination Lacombe inclut aussi la musique et la poésie. Ces deux arts sont évoqués dans l'avertissement pour souligner la richesse des contenus du dictionnaire, ainsi que l'exactitude des informations fournies :

[...] En un mot, on n'a rien négligé pour s'assurer de l'exactitude des recherches; & on a tâché surtout de concilier la netteté avec la précision, & l'utile avec l'agréable.

L'auteur passe ensuite à donner quelques-unes des références bibliographiques qu'il a consultées pour étayer ses recherches. L'avertissement se clôt sur la présentation de quelques musiciens illustres qui, faute de documentation suffisante, ne se voient pas attribuer un article à l'intérieur du dictionnaire.

La préface de PERNETY 1757 s'ouvre par une longue section de défense du genre dictionnaire, dans laquelle l'auteur fait référence au succès de ce genre d'ouvrages publiés au XVIII<sup>e</sup> siècle et aux critiques de ceux qui considèrent qu'ils ne fournissent qu'une érudition superficielle :

Les Dictionnaires sont [...] nécessaires & plus aujourd'hui que jamais, parce que le goût des Sciences & des Arts a gagné tous les États. On veut sçavoir tout, ou plutôt parler de tout, & ne paroître ignorer de rien : il faut donc se prêter à ce goût du siècle.

Il passe ensuite à présenter les spécificités du lexique de la peinture, qui peut être de difficile compréhension car il emploie beaucoup de termes usuels de façon imagée. Dans la dernière partie de la préface, Pernety explique les raisons qui l'ont poussé à rédiger cet ouvrage : d'un côté, un souci franchement pédagogique envers les amateurs et les élèves d'art – « la volonté de me rendre utile »,

comme il le dit –, qui justifie entre autres le choix de faire précéder d'un traité le dictionnaire et, de l'autre, l'envie de pallier aux manques des ouvrages précédents.

La confrontation avec d'autres ouvrages du même genre est aussi un *leitmotiv* de la préface des éditeurs de BOUTARD 1826 qui, contrairement aux autres auteurs, en font expressément mention. Après avoir souligné l'importance des arts comme sujet de conversation et l'engouement des « gens du monde » – destinataires cibles du texte – pour ceux-ci, ils rappellent la réticence de l'Académie française à insérer ces termes spécialisés dans son dictionnaire. Les autres ressources lexicographiques citées – l'*Encyclopédie* de Diderot et d'Alembert, l'*Encyclopédie méthodique*, le dictionnaire de Watelet et celui de Lunier –, bien qu'elles contiennent des articles consacrés aux termes artistiques, affichent néanmoins des défauts. En particulier, les préfaciers déplorent le fait que les arts soient traités parmi une multitude de sujets, par le biais d'abondantes dissertations pouvant fatiguer le lecteur. Ce qui plus est, dans quelques cas ces articles ne sont même pas originaux, mais se limitent à copier d'autres dictionnaires. D'une façon générale, toutes ces ressources sont datées pour les contemporains de Boutard, ne prenant pas en compte les avancées techniques survenues au XIXe siècle.

L'attention accordée aux néologismes constitue donc un atout de BOUTARD (1826). Ceux-ci peuvent désigner de nouvelles réalités dues aux avancées des procédés mécaniques ou encore de nouvelles idées. Cet argument n'est cependant pas le seul pilier sur lequel les préfaciers bâtissent leur discours promotionnel de l'ouvrage de Boutard. Ils en louent l'équilibre :

[le] milieu entre l'aridité des lexicographes qui n'ont eu à rendre compte que de la signification des mots et la prolixité déclamatoire, jusqu'ici trop ordinaire aux littérateurs qui ont écrit sur les beaux-arts,

l'originalité et le caractère de nouveauté qu'il représente dans le panorama éditorial français de l'époque. L'autorité reconnue de l'auteur dans le domaine ne peut être qu'un gage de succès :

[...] Il nous a semblé qu'un écrivain habitué depuis longues années à manier la langue de l'art, dans un commerce en quelque sorte familier avec le public, qui souvent a dû éprouver par lui-même la difficulté de se faire comprendre de ses lecteurs, et qui a plus qu'aucun autre l'occasion d'observer sur quels mots, quelles expressions, quelles idées, porte particulièrement cette difficulté, seroit aussi plus qu'un autre habile à composer le dictionnaire de cette langue, tel que nous l'avions conçu.

## 4 La macrostructure

Examiner la macrostructure de ces dictionnaires revient à identifier à la fois la façon dont ces auteurs ont délimité le vaste domaine des arts et les choix opérés dans la description de la nomenclature.

### 4.1 Le découpage du domaine

La terminologie artistique affiche un caractère hybride et complexe, dont les causes sont à rechercher dans le statut que ces arts ont recouvert au fil des siècles. Dans la hiérarchisation aristotélicienne des disciplines en théorétiques, pratiques et poïétiques, les arts figuratifs occupaient cette dernière catégorie (EUSEBI 2013), en tant qu'arts mécaniques et opposés aux arts libéraux comme la géométrie ou la musique. Maintenu tout au long du Moyen Âge, cette distinction ne sera remise en cause que pendant l'Humanisme et la Renaissance italiens, suite à la redécouverte de la culture classique et à la diffusion de sources textuelles jusque-là inconnues par le biais de traités spécialisés rédigés exclusivement en latin (BIFFI 2007). Cela provoque une révolution épistémologique et culturelle qui a pour résultat l'élévation de l'architecture, de la sculpture et de la peinture au rang des arts libéraux. Pour l'architecture, un travail remarquable de systématisation est accompli par l'architecte italien Leon Battista Alberti à partir de la refonte de la terminologie de Vitruve, source classique de référence pour cette discipline, et de l'alignement entre ce lexique classique et le lexique artisanal d'origine régionale. Faute de références de l'Antiquité, la terminologie de la peinture subit en revanche l'influence de la littérature en vulgaire (MOTOLESE 2012 : 8), en particulier des *Vite* du peintre arétin Giorgio Vasari. Les quatre composantes du lexique vasarien identifiées par LE MOLLÉ (1988 : 12) – théorique, critique, artistique et artisanale – témoignent de la coprésence de différents registres au sein de cette terminologie.

À partir du XVI<sup>e</sup> siècle, les ouvrages de Vasari, Alberti et d'autres artistes/experts d'art font l'objet de traductions et se diffusent ainsi dans d'autres pays, faisant entrer bon nombre de termes italiens dans d'autres langues européennes, dont le français.

La compilation d'une ressource lexicographique spécialisée est réalisée à partir de l'identification préalable du secteur de la connaissance – le « domaine » selon WÜSTER (1981) – qui fait l'objet de la description. Cette tâche peut ne pas être aisée, notamment pour les terminologies riches et hétérogènes comme celle de l'art, pour lesquelles la notion de domaine s'avère être rigide et peu fonctionnelle.

Or, la comparaison entre ces cinq dictionnaires met en exergue des différences dans le découpage du domaine décrit qui sont dues aussi bien aux connaissances de l'auteur qu'aux choix opérés en fonction des destinataires-cibles. Même les titres des ouvrages contiennent des informations importantes à cet égard. FÉLIBIEN (1676) s'intéresse aux trois arts principaux – architecture, peinture et sculpture – et aux arts qui en dépendent, à savoir les arts mécaniques comme la charpenterie, la menuiserie, la serrurerie, l'orfèvrerie, la maçonnerie, à travers la description des technoclectes ordinaires (MESSAOUDI 2013), typiques de la communication orale et liés aux objets-référents *in situ*. En tant qu'expert du domaine, Félibien adopte une approche onomasiologique, qui va du sens aux mots.

MARSY (1746) se limite aux sous-domaines de la peinture et de l'architecture, tandis que LACOMBE (1752) ajoute aux trois arts figuratifs principaux la musique et la poésie, se focalisant ainsi sur les beaux-arts, exprimés par les technoclectes savants.

Si PERNETY (1757) ne prend en considération que la peinture, la sculpture et la gravure, BOUTARD (1826) ajoute à celles-ci l'architecture, dont il est spécialiste, et indique ce regroupement de disciplines par l'étiquette « arts du dessin ». Cependant, à la différence de Félibien, Boutard exclut de son dictionnaire les termes techniques employés dans les ateliers. Sans doute Boutard choisit-il sciemment de ne pas employer l'appellation « beaux-arts » qui inclut aussi la musique après la constitution officielle de l'Académie des Beaux-Arts en 1816, héritière des Académies royales de Sculpture et Peinture (1666), de Musique (1669) et d'Architecture (1671).

## 4.2 La nomenclature

Avant de rentrer dans les détails de la microstructure, voyons comment les données lexicographiques sont présentés au lecteur. Tous les auteurs optent pour l'agencement alphabétique des articles, qui sont disposés sur une seule colonne occupant la page dans FÉLIBIEN (1676), MARSY (1746) et BOUTARD (1826) ou sur deux colonnes dans LACOMBE (1752) et PERNETY (1757).

Dans tous les dictionnaires la vedette est affichée en lettres capitales ou en petites capitales. Comme dans les dictionnaires de langue, les verbes sont enregistrés à l'infinitif, les adjectifs au masculin singulier et les noms au singulier. Cependant, il n'est pas rare de trouver la forme au pluriel comme vedette si le nom est employé au pluriel et ce choix concerne tous les dictionnaires du corpus. FÉLIBIEN (1676) adopte le critère du dégroupement homonymique (PRUVOST 2006) dans la description des différentes acceptions d'un même terme : chaque

acception occupe une ligne différente, comme ce sera aussi le cas pour les dictionnaires de RICHELET (1680) ou de l'Académie française (1694). Ce critère est repris aussi par les autres lexicographes avec quelques exceptions : LACOMBE (1752) n'opte pour le dégroupement homonymique que lorsqu'un même terme est employé dans des sous-domaines différents (tel est le cas, par exemple, du terme *camayeux* utilisé en peinture et en gravure), BOUTARD (1826) n'utilise pas souvent les capitales ou petites capitales pour indiquer le terme-vedette dans une acception différente.

Il en va de même pour le traitement des collocations, qui sont disposées sur une ligne séparée à l'intérieur de l'article du nom-tête à partir duquel elles sont formées. Elles peuvent aussi constituer des entrées indépendantes, comme par exemple le terme *amitié des couleurs* que FÉLIBIEN (1676) décrit séparément et non pas au sein de l'article *couleur*, ou tous les termes composés à partir du nom-tête *colle* dans le dictionnaire de PERNETY (1757), auxquels l'auteur consacre des entrées individuelles. Dans ce dernier cas, il nous semble que ce choix est fonctionnel aux enjeux pédagogiques de l'ouvrage. De son côté, BOUTARD (1826) décide d'adopter un critère morphologique, selon lequel les mots composés par trait d'union – comme *clair-obscur* ou *contre-allée* – constituent des entrées indépendantes, tandis que les collocations sont traitées à l'intérieur de l'article du nom-tête. Ainsi tous les types de colonnes sont-ils décrits sous l'article *colonne*, conformément à la démarche encyclopédique de l'auteur. Contrairement aux autres lexicographes, LACOMBE 1752 décrit souvent les collocations en faisant apparaître le collocatif en position de vedette et le nom-tête entre parenthèses : à titre d'exemple, nous citerons *crayon (dessin au)* et *chinoise (peinture)*.

Il peut arriver aussi que deux ou plusieurs vedettes ouvrent un même article : il s'agit de mots formés à partir d'un même radical, comme *conduire* et *conduite* ou *correct*, *correction* et *correctement*. C'est MARSY (1746) qui opte pour ce regroupement par racines, caractérisant la première édition du *Dictionnaire de l'Académie* (ACADÉMIE FRANÇAISE 1694) et puis abandonné par les Académiciens dès la deuxième. Ce procédé permet à l'auteur de traiter ensemble des mots pouvant être employés pour exprimer un même concept, sans avoir pour cela à créer deux articles séparés. À titre d'exemple, nous reportons ici l'article *conduire, conduite* :

CONDUIRE, CONDUITE.

*Conduire* signifie diriger, ménager, distribuer.

Des jours & des ombres *conduits* judicieusement. Félibien a dit : un tableau bien *conduit* de couleurs, c'est-à-dire, où les couleurs sont ménagées, & distribuées avec Art. il y a beaucoup de *conduite* dans les compositions de Poussin, c'est-à-dire, beaucoup d'entente & d'ordonnance.

## 5 La microstructure

La comparaison entre les cinq dictionnaires du corpus sera approfondie dans cette section, dans laquelle nous nous focaliserons sur leur microstructure, à savoir les types d'informations contenues au sein des articles. Parmi celles-ci, notre attention portera en particulier sur les catégories syntaxiques et sémantiques faisant partie de la nomenclature, sur les types de définitions employées et sur les marques de domaine. La comparaison sera étayée par l'analyse d'un même article.

### 5.1 Catégories syntaxiques et sémantiques

Comme on peut le déduire à partir de ce qui a été présenté dans les sections précédentes, ces dictionnaires ne contiennent pas les mêmes lemmes. Si une partie de la nomenclature est commune aux cinq ouvrages, force est de constater qu'ils affichent des différences aussi bien au niveau quantitatif – nombre de lemmes traités – que qualitatif. S'il est vrai que FÉLIBIEN (1676) contient le plus grand nombre d'entrées (372) pour la lettre C, il ne faut pourtant oublier qu'une partie de celles-ci ne constituent que des renvois au traité ou aux planches contenues dans celui-ci et sont donc dépourvues d'une définition.

Dans les ressources terminographiques, les noms constituent la catégorie syntaxique la plus représentée. Les cinq dictionnaires du corpus ne font pas exception dans ce sens, mais ils accueillent également des verbes, des adjectifs (et des participes passés employés comme adjectifs) et quelques adverbes. Parmi ces dictionnaires, le seul à fournir une indication de la catégorie grammaticale au sein des articles est BOUTARD (1826). Deux dictionnaires du XVIII<sup>e</sup> siècle – MARSY (1746) et LACOMBE (1752) – font la part belle aux noms propres, qui désignent surtout des artistes<sup>3</sup>, français ou internationaux. Il en résulte alors un genre hybride, combinant les caractéristiques du dictionnaire spécialisé sur l'art et celles du recueil de biographies d'artiste. Ce dernier est en effet très en vogue dans les mêmes années (1745–1755), notamment pour les vies des peintres (MAËS 2016 : 171). À côté des artistes, les noms propres peuvent aussi désigner des divinités ou des figures de la mythologie représentées dans des œuvres célèbres. Dans les articles consacrés à ces noms propres on peut aussi trouver le récit de ces mythes. Le dictionnaire abrégé/portatif accomplit ainsi sa mission :

Aux côtés de l'*Encyclopédie*, le portatif obéit à un besoin profond de l'époque : celui de vulgariser le savoir, et aussi de le rendre accessible,

---

<sup>3</sup>Dans MARSY (1746) bon nombre de noms propres désignent aussi des monuments.

à un prix relativement modique, à une clientèle cultivée mais élargie. [...] Mais le portatif est aussi et surtout le domaine de la liberté, de la sollicitation permanente des goûts du public par les compilateurs de toute sorte. [...] Le portatif ressort dans ce contexte global comme sous-dictionnaire à tout faire, mais il est lui-même ouvert aux influences encyclopédiques. (MARTIN & CHARTIER 1984 : 191)

Voyons à présent les catégories sémantiques représentées dans la nomenclature des dictionnaires du corpus. Les verbes désignent principalement des techniques ou des actions, tandis que les adjectifs se réfèrent à des qualités ou à des caractéristiques. Pour les substantifs, il est possible d'identifier des catégories sémantiques communes à tous les dictionnaires du corpus – bien que les nomenclatures ne se recouvrent pas parfaitement – comme celles des ornements et des décorations, des styles, des techniques, des couleurs, des parties de la peinture, des objets. À l'exception de PERNETY 1757, tous les dictionnaires contiennent des noms désignant des types et des parties d'édifices. En raison de l'importance accordée aux technoclectes ordinaires, les noms décrits par FÉLIBIEN (1676) désignent surtout des matériaux, des outils, des objets, comme le feront aussi MARSY (1746) et PERNETY 1757, bien qu'en moindre mesure. LACOMBE (1752) et BOUTARD (1826) sont davantage tournés vers le vocabulaire conceptuel de la peinture. Puisque LACOMBE (1752) s'intéresse aussi à la poésie et à la musique, bon nombre des noms répertoriés appartiennent aux catégories sémantiques des genres poétiques et musicaux et des instruments de musique.

## 5.2 Les types de définitions

Venons-en maintenant aux types de définitions que l'on peut observer dans ces dictionnaires à partir de l'ouvrage fondateur de FÉLIBIEN (1676). Celui-ci se sert de plusieurs types de définitions, à commencer par les définitions logiques, d'origine aristotélicienne et très répandues dans la pratique lexicographique. Appelées aussi « définitions en compréhension », elles visent une catégorisation logique du monde (PRUVOST 2006). Il s'agit de procédés définitoires directs qui, à un premier niveau de classification, mettent en relation un hyponyme avec son hyperonyme par le biais de définisseurs génériques. À un deuxième niveau, les définisseurs spécifiques fournissent plus de détails sur l'objet décrit – notamment, sa fonction ou sa finalité. Suivant la quantité d'informations fournies, les définitions logiques peuvent être qualifiées de hypospécifiques, suffisantes ou hyperspécifiques. Au vu de la complémentarité entre le dictionnaire et le traité *Des Principes*, la taille des définitions logiques fournies par Félibien est variable,

pouvant aller de définitions très réduites pour les concepts déjà abordés dans le traité à de véritables développements encyclopédiques pour d'autres unités terminologiques.

À côté des définitions logiques de taille variable, FÉLIBIEN (1676) propose aussi des définitions plus courtes, comme les définitions synonymiques, typiques des dictionnaires de poche et consistant à définir un mot par le biais d'un synonyme<sup>4</sup>, et des définitions par rattachement morphologique, employées surtout pour expliquer le sens de verbes spécialisés, comme par exemple

- CALER, Mettre une cale.
- CONTRESCARPER, Faire une contrescarpe.

Cependant, parmi les définitions utilisées par l'historiographe celles qui ont particulièrement retenu notre attention sont les définitions finalisées. Celles-ci sont des énoncés définitoires ayant la forme d'une phrase relative du type *N qui* ou *N que...* / *N dont* et qui sont employés pour exprimer les finalités sociales ou instrumentales d'un objet par le biais d'un complément circonstanciel. Relevées par COLLINOT & MAZIÈRE (1990) dans le *Dictionnaire Universel de Furetière* (FURETIÈRE 1690) et dans la première édition du *Dictionnaire de l'Académie française* (ACADÉMIE FRANÇAISE 1694), les définitions finalisées semblent caractériser la lexicographie monolingue au XVIIe siècle. Les schémas définitoires de ce type employés par Félibien sont les suivants :

1. N qui sert à Nhum<sup>5</sup> pour / à
2. N servant à
3. N qui est propre à / pour
4. N dont on se sert pour / à
5. N dont les Nhum se servent pour / à.

Les prépositions *à* et *pour* peuvent être suivies par un verbe à l'infinitif ou par un complément d'objet, comme dans

- CABLES, ou Chables, ce sont de grosses cordes servant à monter des fardeaux.

---

<sup>4</sup>Les définitions synonymiques peuvent parfois indiquer des régionalismes.

<sup>5</sup>Nhum : nom humain. Dans le cas de ces types de définitions, les Nhum désignent surtout des noms de métiers (ex. les serruriers, les charpentiers, etc.).

- COYERS, ce sont des pièces de bois servant à la couverture d'un bâtiment.
- BEC D'ASNE croche, dont les Serruriers se servent pour ferrer les fiches dans le bois.

Comme nous avons déjà eu l'occasion de le remarquer ailleurs,

[...] La présence des définitions finalisées dans le dictionnaire de Félibien est un indice incontestable du ferment du chantier lexicographique monolingue dans la seconde moitié du XVII<sup>e</sup> siècle en France et de la circulation des savoirs entre les différentes académies royales. (CETRO 2022 : 162).

Un dernier type de définitions relevées dans FÉLIBIEN (1676), bien que rarement employées, sont les définitions en extension (BÉJOINT 1993). Dans ces rares cas, une classification de l'objet décrit est fournie à la place de sa définition. D'habitude, il s'agit d'objets communs<sup>6</sup>, d'un usage très répandu, que le lecteur est censé connaître : Félibien exploite donc une connivence avec celui-ci pour donner des détails sur ces objets sans passer par une définition. Ce procédé définitoire est parfois employé aussi par MARSY 1746<sup>7</sup> ou par PERNETY 1757. Pour revenir aux autres types de définitions passés en revue plus haut, tous les dictionnaires du corpus fournissent des définitions en compréhension, qui sont souvent hyperspécifiques, conformément à leur vocation encyclopédique. Les trois dictionnaires du XVIII<sup>e</sup> siècle enregistrent aussi des définitions synonymiques aussi bien pour des raisons d'espace que pour signaler des variantes régionales<sup>8</sup>, tandis que BOUTARD (1826) s'en sert très rarement. Sur l'échantillon analysé, ce procédé définitoire est utilisé pour indiquer un nécrologisme :

- COMMISSURE, s.f. Archit. Quelques vieux auteurs, comme Philibert de Lorme et le père Deran, appellent ainsi les joints de pierre. (BOUTARD 1826 : 167).

Cet exemple montre l'emploi d'un énoncé métadiscursif qui sert d'appui à la définition. Ce procédé, que BOUTARD 1826 utilise occasionnellement puisque ses définitions se basent la plupart des fois sur un hyperonyme, est en revanche très fréquent chez ses prédécesseurs, à partir de FÉLIBIEN (1676), qui antépose à ses définitions un présentatif ou un énoncé métadiscursif formé sur un verbe du dire

---

<sup>6</sup>C'est le cas par exemple des entrées *balance*, *gomme* et *porte*.

<sup>7</sup>Voir par exemple l'article *colle*.

<sup>8</sup>Voir pour cela l'article *cocagne* dans MARSY (1746 : 131) : Cocagne, c'est le Pastel du Languedoc. Voyez Pastel.

– *appeler, nommer, dire*, etc. Dans le sillage de Félibien, MARSY (1746) et LACOMBE (1752) exploitent souvent ces procédés, tandis que PERNETY (1757) élimine les présentatifs de ses énoncés définitoires, se limitant parfois à employer un énoncé métadiscursif.

### 5.3 Les marques de domaine

FURETIERE (1690), qui enregistre de nombreux termes spécialisés, exprime les marques de domaine de deux façons. Une première, au début de l'article, qui est *terme de* suivie du domaine dans lequel le terme est employé (ex. *terme de géographie*)<sup>9</sup> et une deuxième qui passe par l'emploi d'un énoncé métadiscursif (ex. *Les Medecins appellent isthme, cette partie étroite de la gorge qui est située entre les deux amygdales*). Il en va de même pour FÉLIBIEN (1676), qui se sert de trois schémas récurrents pour indiquer l'appartenance d'un terme à un sous-domaine artistique. Leur but est donc de souligner la dimension fonctionnelle des termes. Il s'agit des schémas suivants :

- *En terme de* + nom de domaine : ex. en terme de fortification, de peinture, etc. ;
- Un énoncé métadiscursif avec le pronom impersonnel *on* suivi d'un verbe du dire, de la préposition *en* et du nom de domaine : ex. L'on dit en peinture [...];
- Un énoncé métadiscursif ayant pour sujet le nom désignant une catégorie professionnelle, suivi d'un verbe du dire et d'un complément d'objet direct : ex. Les Joüaillers et les Lapidaires nomment Camayeus les Onices, Sardoines et autres pierres taillées de relief.

Si l'emploi des marques de domaine n'est pas systématique, il est néanmoins très fréquent chez Félibien. Il n'en est pas ainsi chez MARSY (1746), qui n'y recourt que rarement. Sur le corpus analysé, les quelques cas relevés font usage du schéma félibien *en terme de* suivie d'un nom de domaine (ex. en terme d'architecture) ou simplement de la préposition *en* suivie du nom de domaine (ex. en architecture) ou encore d'un énoncé métadiscursif calqué sur le troisième schéma employé déjà par Furetière et Félibien (ex. les peintres appellent X...). De son côté, LACOMBE (1752), chez qui l'emploi des marques d'usage reste assez limité, privilégie la marque *terme de* suivie du nom de domaine à l'emploi d'un énoncé métadiscursif impersonnel. Quant à PERNETY (1757), il reprend les deux premiers

---

<sup>9</sup>Ce même procédé est employé dans la première édition du *Dictionnaire de l'Académie française*.

schémas employés par Félibien, auxquels il ajoute une autre formulation impersonnelle (ex. se dit en terme de peinture). Si un même terme est employé dans deux domaines différents, les noms de ceux-ci sont donnés entre parenthèses (ex. peinture et gravure). Avec BOUTARD 1826, l'emploi des marques de domaines devient systématique et homogène. Les énoncés métadiscursifs disparaissent au profit d'abréviations, et cette convention est désormais un acquis de la lexicographie.

#### 5.4 Comparaison d'un même article

Dans cette dernière section, nous nous appuyerons sur le traitement d'un même mot au sein des cinq dictionnaires pour procéder à une comparaison entre leurs microstructures. Pour cela, nous avons reporté dans le tableau suivant (voir (1)) les articles relatifs au mot *chargé* – *charge* dans LACOMBE (1752) et BOUTARD (1826) :

- (1) L'article chargé/charge dans les dictionnaires du corpus.
  - a. FÉLIBIEN (1676)
 

CHARGÉ. Les Peintres appellent un portrait chargé, lorsqu'on représente un visage avec des traits marquez avec excès, et de telle manière qu'avec trois ou quatre coups de crayon ou autrement on connaît une personne, quoy que ce ne soit pas un véritable portrait, mais plutôt des défauts marquez. Aussi quand une Figure est trop marquée on dit qu'elle est chargée.
  - b. MARSY (1746)
 

CHARGE, CHARGER. On appelle proprement *charges* ou *tableaux chargés*, des représentations où l'on exagère les choses en bien ou en mal, mais plutôt ordinairement en mal.

Un Peintre satyrique en trois ou quatre coups de pinceau, fait un portrait ridicule, mais fort ressemblant, quoiqu'en laid, en exagérant la difformité de la personne qu'il représente. On appelle encore *charges* certains caprices & certaines figures grotesques : par exemple, des animaux avec une figure humaine, des hommes avec des pieds d'animaux, des femmes sous la figure d'un pot ou de quelqu'autre vase. C'est une espèce de genre burlesque que le mauvais goût, qui corrompt tous les Arts, a introduit dans la Peinture [...].
  - c. LACOMBE (1752)
 

CHARGE. On donne en général ce nom dans le Dessein, à tout ce

qui est outré et hors de vraisemblance. Mais on appelle particulièrement ainsi, en terme de Perspective ou de Dessin, une exagération burlesque des parties les plus marquées du visage, de façon que la ressemblance soit conservée, & que l'on puisse reconnoître la personne dont on fait la *Charge*. Voyez *Carricature*.

d. PERNETY (1757)

CHARGÉ, se dit d'un portrait, dont les traits sont marqués avec excès ou exagérés. Ce terme est employé pour définir un contour qui mérite le blâme & la critique, pour être dur et prononcé. Un portrait chargé, se fait en trois ou quatre coups de crayon, selon M. Félibien; & ils suffisent pour présenter la ressemblance d'une personne, quoique le portrait ne soit pas achevé. Les mauvais Peintres chargent les portraits par ignorance de leur art, & les bons Artistes par la science qu'ils en ont, aidée d'une main hardie, ferme & décidée.

e. BOUTARD (1826)

CHARGE, s.f. *Peint.* Exagération dans les formes, dans l'inflexion des lignes, dans les traits du visage, les attitudes, le mouvement d'une figure. On appelle aussi charge l'ouvrage exécuté de cette manière. La charge diffère de la caricature avec laquelle on la confond quelquefois, en ce qu'elle n'admet point d'êtres fantasques, mais qu'elle a toujours son type dans la nature animée. La charge n'est pas une fiction, mais une exagération de la vérité. [...] La charge est aussi l'artifice ou plutôt la ressource des peintres du portrait de second ordre [...].

Comme on peut l'observer, jusqu'à PERNETY (1757) la définition s'ouvre par un énoncé métadiscursif, tandis qu'avec BOUTARD (1826) l'objet est immédiatement qualifié à l'aide du définisseur *exagération*. Par rapport à ses prédécesseurs, ce dernier auteur affiche une tendance à la nominalisation, qui lui permet de condenser les informations et qui caractérisera la lexicographie contemporaine. Les informations données sur l'objet décrit diffèrent quelque peu selon le dictionnaire. Cependant, il est possible de relever des motifs récurrents et la reprise ou la reformulation de certaines informations du dictionnaire fondateur, celui de Félibien.

## 6 Conclusion

La terminologie artistique – et plus en général le monde des beaux-arts – a suscité un véritable engouement au sein de la société française de l'Ancien Régime,

comme en témoignent les nombreuses publications consacrées à ce sujet, dont les dictionnaires. Ce genre textuel va particulièrement à l'encontre du « goût du siècle » pendant les Lumières et se poursuit même après la Révolution, se perfectionnant et anticipant les tendances de la lexicographie contemporaine.

Rédigés par des amateurs, des connaisseurs ou des spécialistes et destinés à des publics cibles diversifiés, les ouvrages pris en examen dans ce chapitre mettent en avant les difficultés dans le découpage et la description d'un si vaste secteur de connaissance(s). Malgré leurs limites, ils constituent un témoignage inestimable des questionnements qui sous-tendent à la rédaction d'un dictionnaire spécialisé et des réponses-solutions apportées par les auteurs.

## Références

- ACADÉMIE FRANÇAISE. 1694. *Dictionnaire de l'Académie*. Paris : Coignard.
- BÉJOINT, Henri. 1993. La définition en terminographie. In Pierre J. L. ARNAUD & Philippe THOIRON (éd.), *Aspects du vocabulaire* (Travaux du CRTT), 19-36. Lyon : Presses universitaires de Lyon.
- BIFFI, Marco. 2007. La terminologia tecnica dell'Alberti tra latino e volgare. In R. CARDINI & M. REGOLIOSI (éd.), *Alberti e la cultura del Quattrocento : Atti del convegno internazionale* (Strumenti 5), 655-682. Firenze : Polistampa.
- BOUTARD, Jean-Baptiste-Bon. 1826. *Dictionnaire des arts du dessin, la peinture, la sculpture, la gravure et l'architecture*. Paris : Le Normant & Gosselin.
- CETRO, Rosa. 2022. *La démarche terminologique d'André Félibien. La systématisation du lexique artistique en français*. Torino : L'Harmattan Italia.
- COLLINOT, André & Francine MAZIÈRE. 1990. Les définitions finalisées dans le Furetière. Où il est montré que l'hétérogénéité syntaxique des définitions a du sens. In Jacques CHAURAND & Francine MAZIÈRE (éd.), *La définition*, 237-249. Paris : Larousse.
- EUSEBI, Cristina. 2013. *Contributo dell'italiano alla formazione del lessico architettonico rinascimentale inglese*. Università di Trento. (thèse de doct.).
- FÉLIBIEN, André. 1676. *Des Principes de l'Architecture, de la Sculpture et de la Peinture et des autres Arts qui en dépendent avec un Dictionnaire propre à chacun de ces Arts*. Paris : J. B. Coignard.
- FURETIÈRE, Antoine. 1690. *Dictionnaire universel contenant généralement tous les mots françois tant vieux que modernes et les termes de toutes les sciences et des arts ... Par... Antoine Furetière,...* 3 t. Amsterdam : Leers. <http://gallica.bnf.fr/ark:/12148/bpt6k3413126b> (17 mai, 2023).
- LACOMBE, Jacques. 1752. *Dictionnaire portatif des beaux-arts*. Paris : Estienne & Herissant.

- LE MOLLÉ, Roland. 1988. *Georges Vasari et le vocabulaire de la critique d'art dans les «Vite»*. Grenoble : Ellug.
- MAËS, Gaëtane. 2016. Les dictionnaires des Beaux-Arts au XVIIIe siècle : Pour qui et pourquoi ? In Jessica L. FRIPP, Amandine GORSE, Nathalie MANCEAU & Nina STRUCKMEYER (éd.), *Artistes savants et amateurs : Art et sociabilité au XVIIIe siècle (1715-1815)*, 171-184. Paris : Mare & Martin.
- MARSY, François-Marie de. 1746. *Dictionnaire abrégé de peinture et d'architecture*. Paris : Nyon fils.
- MARTIN, Henri-Jean & Roger CHARTIER (éd.). 1984. *Histoire de l'édition française, T.2. Le livre triomphant 1660-1830*. Paris : Promodis.
- MESSAOUDI, Leila. 2013. Technolectes savants, technolectes ordinaires : Quelles différences ? In Leila MESSAOUDI (éd.), *Sur les technolectes*, 127-135. Kénitra : Publications du laboratoire Langage et société.
- MOTOLESE, Marco. 2012. *Italiano lingua delle arti : Un'avventura europea*. Bologna : Il Mulino.
- PERNETY, Antoine-Joseph. 1757. *Dictionnaire portatif de peinture, sculpture et gravure, avec un traité pratique sur les différentes manières de peindre*. Paris : Bauche.
- PRUVOST, Jean. 2003. Les dictionnaires français : Histoire et méthodes. In Marina YAGUELLO (éd.), *Le grand livre de la langue française*, 457-489. Paris : Seuil.
- PRUVOST, Jean. 2006. *Les dictionnaires français : Outils d'une langue et d'une culture*. Paris : Ophrys.
- RICHELET, Pierre. 1680. *Dictionnaire François. Contenant ... plusieurs nouvelles remarques sur la langue Française : ses expressions propres, figurées et burlesques, la prononciation des mots ... difficiles, le genre des noms; le regime des verbes : avec les termes les plus connus des arts et des sciences*. Genève : Jean Herman Widerfold.
- SGARD, Jean (éd.). 1999. *Dictionnaire des Journalistes 1600-1789*. Oxford : Voltaire Foundation.
- WÜSTER, Eugen. 1981. L'étude générale de la terminologie, zone frontalière entre la linguistique, la logique, l'ontologie, l'informatique et les sciences des choses. In Guy RONDEAU & Henri FELBER (éd.), *Textes choisis de terminologie*, 55-108. Québec : Girsterm.

## Chapitre 12

# La valeur pragmatique des langues dites « orientales » dans le Dictionnaire universel de Trévoux (1721)

Georgios (Yorgos) Kassiteridis

Centre Norbert Elias – Avignon Université

A remarkable feature of the *Dictionnaire universel de Trévoux* is the increased occurrences of non-European languages therein, such as Hebrew, Arabic, and Syriac, referred to as “oriental languages” in the dictionary. A comparative study of some relevant entries illustrates that this cannot be justified by its source text, namely Basnage de Beauval’s *Dictionnaire universel* (1701), where these languages are not equally present. The use of “oriental languages”, particularly in the articles related to religious scriptures, leads us to associate this lexicographical innovation with the tradition of Critical Biblical Studies, represented by Louis Cappel (1585–1658) and Richard Simon (1638–1712). This interpretative approach examines the Bible in a philological-historical manner by comparing different textual versions written in various languages. This chapter aims to identify the role of the so-called “oriental languages” in the second edition of the *Dictionnaire universel de Trévoux* (1721) in order to gain a better understanding of the lexicographic tools applied in the context of a controversy, as well as to capture the ideologies related to these languages in the eighteenth century.

Une particularité marquante du *Dictionnaire universel de Trévoux* est son identité bilingue (français-latin) ainsi que l’amplification de ses pratiques plurilingues au fur et à mesure de différentes éditions qui s’échelonnent sur tout le XVIII<sup>e</sup> siècle (de 1704 à 1771). L’usage du latin<sup>1</sup> et des langues européennes est analysé

---

<sup>1</sup>Il s’agit d’une innovation, puisque les rédacteurs du *Dictionnaire universel de Trévoux* reprennent une pratique déjà ancienne (à l’instar notamment de Robert Estienne et de Jean Nicot), alors que la généalogie sur laquelle ils se fondent, celle de Furetière (puis Basnage de Beauval) et des dictionnaires de langue comme celui de Richelet ou de l’Académie française, avait produit des dictionnaires monolingues.



par Chantal WIONET (1998, 2006). On constate également la présence accrue de langues non européennes telles que l'hébreu, l'arabe, le syriaque, désignées dans l'édition trévoltienne, comme nous le verrons plus loin, comme « langues orientales ». Cet intérêt ne peut pas se justifier par rapport à son texte source, à savoir le *Dictionnaire universel* (1701) de Basnage de Beauval, où ces langues ne sont pas présentes de manière égale, ni même encore précédemment à celui de FURETIÈRE (1690). Le recours aux langues orientales, particulièrement dans les articles relatifs aux écritures religieuses, nous amène à lier cette innovation lexicographique aux études bibliques critiques (LAPLANCHE 1994), comme les travaux de Louis Cappel (1585–1658) et de Richard Simon (1638–1712), qui abordent l'interprétation de la Bible de manière philologico-historique à travers la mise en regard des versions et des langues. La présente étude vise ainsi à approfondir et à identifier les rôles et les finalités des langues orientales dans le *Dictionnaire universel de Trévoux* (1721), afin d'enrichir notre compréhension des outils lexicographiques appliqués dans le contexte d'une controverse de nature religieuse et commerciale, ainsi que pour percevoir les imaginaires linguistiques des langues orientales au XVIII<sup>e</sup> siècle.

Pour ce faire, mon analyse se fonde sur un échantillon d'entrées tirées des lettres A et D de la deuxième édition du *Dictionnaire universel de Trévoux* publié par l'imprimerie de Trévoux en 1721, édition anonyme (comme la première) supervisée par Etienne Souciet (1671–1744), bibliothécaire au collège jésuite Louis-le-Grand (*Mémoires de Trévoux*, avril 1744 : 753). Le choix de cette édition n'est pas arbitraire ; il s'agit d'une édition qui s'écarte davantage de ses origines (du *Dictionnaire universel* de Basnage de Beauval et celui de Furetière) par rapport à son édition antérieure, ce qui illustre l'augmentation de nouvelles entrées surtout dans le domaine religieux, mais aussi géographique, onomastique, botanique (LECA-TSIOMIS 2023 : 137). Dans une première phase, j'ai recueilli, par une lecture rapprochée du dictionnaire, les entrées où s'affiche la structure « langues orientales », afin de comprendre quelles sont les langues concrètes qui appartiennent à ce groupe, selon les rédacteurs du Trévoux, à partir des caractéristiques linguistiques reliées aux langues orientales. Puis, j'ai utilisé cette liste de langues comme mots-clés pour identifier les entrées pertinentes et constituer le corpus de recherche.

Il faut souligner que pour l'identification des langues appartenant aux langues dites « orientales », je me suis d'abord focalisé sur le terme « langues orientales » en excluant les mentions aux peuples « orientaux » et leurs langues, contrairement à l'approche que j'ai suivie pour les langues individuelles (voir ci-dessous), puisque le terme « orientaux » pour les rédacteurs du Trévoux comprend plusieurs peuples, comme les chinois, les japonais, les arméniens, dont les langues

ne partagent pas les mêmes caractéristiques linguistiques que celles attribuées aux langues orientales<sup>2</sup>. Pour illustrer cela, il suffit de consulter la sous-entrée ORIENTAL<sup>3</sup> ou les langues citées comme exemples sont « l'Hébreu, le Chaldéen, le Syriaque, l'Arabe, le Copt ».

Avant même la publication de la première édition du *Dictionnaire universel de Trévoux* (1704), on perçoit quelques traces d'expression d'intentions dans les huit échanges que les collaborateurs et partisans du *Dictionnaire universel* de Basnage de Beauval (à savoir Basnage de Beauval, Gédéon Huet, Jean Le Clerc) ont eus avec leurs critiques trévoltiens dans les pages des *Mémoires de Trévoux* (janvier-février 1701; mars 1701; septembre-octobre 1701), de sa contrefaçon à Amsterdam publiée chez Jean-Louis de Lorme (mai-juin 1701; juillet-août 1701) et du *Journal des sçavans* (juillet 1701). On y trouve notamment des critiques assez élaborées sur des éléments des langues hébraïque (ALEPH), arabe (ACCENT), syriaque (AB-BOT), à savoir de fautes qui viennent de l'ignorance des langues Orientales (dans *Mémoires de Trévoux*, mars 1701 : 104)<sup>4</sup>.

Par ailleurs, la comparaison entre le *Dictionnaire universel* (1701) de Basnage de Beauval et celui de Trévoux 1721 illustre une multiplication de références aux langues orientales (et en plus la transcription dans une écriture non latine) dans le *Dictionnaire universel de Trévoux*. En particulier, dans le corpus d'entrées comportant des références aux langues orientales concrètes (nous voyons ci-dessous les langues concernées), tirées de la lettre A de la deuxième édition du Trévoux (dorénavant sous-corpus A), 53 % n'existe pas au niveau d'entrées dans le *Dictionnaire universel* de Basnage de Beauval et seulement 29 % de celles qui y existent comporte déjà des références aux langues orientales, sous la forme de langues individuelles (par exemple, l'arabe ou l'hébreu). La majorité des références aux langues orientales représente donc une innovation du *Dictionnaire universel de Trévoux* par rapport à sa source. En outre, dans le dictionnaire trévoltien lui-même, les rédacteurs admettent que plusieurs mots religieux proviennent de

---

<sup>2</sup>« c'est le nom qu'on donne aux Prêtres Orientaux, particulièrement à la Chine et au Japon » dans BONZE; « car les Arméniens & les autres Orientaux [...] » dans ARMÉNIEN.

<sup>3</sup>Les références en majuscules renvoient à des entrées du *Dictionnaire universel de Trévoux* (1721).

<sup>4</sup>Nous y trouvons également une compréhension du rôle des langues non indoeuropéennes dans les écritures sacrées, voir « Pour parler exactement il faut dire que les Apôtres & les Evangelistes ont conservé dans leurs écrits de certains mots Syriaques », dans les *Mémoires de Trévoux*, mars 1701 : 109, et une attention à la grammaire et à sa relation avec la traduction, voir « car ces articles changent souvent, le sens du discours, & si l'on n'en sçait la force, on ne peut pas traduire exactement les Livres sacrez en nôtre langue » dans *Mémoires de Trévoux*, mars 1701 : 107; ce qui oriente notre hypothèse sur la relation entre l'intérêt pour les langues orientales et l'interprétation/traduction des Écritures.

langues étrangères<sup>5</sup>, comme l'hébreu, « langue [...] dans laquelle sont écrits la plupart des livres de l'ancien Testament qu'un bon Interprète de l'Ecriture doit savoir » (dans HÉBREU), et ils soulignent que « le rapport qu'ils ont aux Langues étrangères dont ils sont dérivez, [...] sont essentielles, & d'un très-grand secours pour les entendre » (Préface). L'addition de langues et d'entrées, l'intérêt déjà manifesté dans les critiques du dictionnaire de Basnage de Beauval pour les langues orientales, ainsi que l'importance de l'hébreu pour la compréhension des mots religieux - sujet pertinent pour la controverse d'où est issu ce dictionnaire - me semble bien indiquer que la présence des langues orientales dans le Trévoux n'est pas accidentelle, mais un choix conscient dont je me propose ici de saisir la logique.

Tout d'abord, pour éviter d'imposer des taxonomies contemporaines, j'ai créé une liste de caractéristiques de langues orientales à partir du contexte textuel du syntagme « langues orientales » dans le sous-corpus A (voir Annexe). Ensuite, un processus d'élimination a été effectué afin de comprendre ce que les rédacteurs du Trévoux désignent comme « langues orientales ». On trouve, notamment, que l'augment syllabique « pourroit être appliqué à la Grammaire des langues Orientales » (dans AUGMENT). L'augment syllabique existe dans de nombreuses langues, comme le grec, le sanskrit, l'arménien, l'hébreu, les langues bantoues, etc. Mais on lit également que les langues orientales « n'exprimoient point les voyelles » (dans ASPIRATION). Un exemple hors sous-corpus A et D : dans CONSONE, les rédacteurs du Trévoux indiquent que les Grammairiens Hébreux « ont été imitez [...] par les Grammairiens des autres langues orientales », un « autres » qui me paraît être une preuve que l'hébreu appartient, pour nos rédacteurs, aux langues orientales ; ou également dans AIN, qui renvoie à une lettre de « toutes les langues orientales » et qui est signalée comme « Terme de Grammaire Hébraïque & Arabe ». Ces deux langues nous offrent, ainsi, une porte d'entrée pour commencer à examiner l'identité des langues orientales dans le *Dictionnaire universel de Trévoux*. On reconnaît aussi le syriaque et le chaldéen comme les « filles de l'hébreu » (dans AMEN). Il existe d'autres langues qui pourraient être pertinentes dans une analyse qui porterait dans un corpus plus étendu, comme le phénicien, le samaritain ou l'assyrien, mais puisque leur présence dans la lettre A et D n'est pas assez fréquente pour en tirer des hypothèses soutenues, nous insistons sur les quatre langues mentionnées ci-dessus : l'hébreu, l'arabe, le syriaque, le chaldéen.

---

<sup>5</sup>« mais même dans aucun autre Livre, à sçavoir une explication très-curieuse & très-nette de toutes les sectes différentes en fait de Religion. Comme ces mots transferez d'une Langue étrangere dans la nostre, en font maintenant une partie, on n'a pû s'empêcher de les mettre en leur place » dans la Préface du *Dictionnaire universel de Trévoux* (1721).

Si on s'éloigne pour un moment du texte dictionnaire, afin de vérifier cette idée aussi à l'aide des ressources externes, on trouve que l'hypothèse selon laquelle l'hébreu, l'arabe, le chaldéen et le syriaque se cachent derrière l'expression « langues orientales » s'aligne sur l'analyse de STOLZENBERG (2015). En particulier, Stolzenberg a entrepris de déterminer le sujet linguistique de la *philologia orientalis* pendant la période entre 1650 et 1750 à partir des titres d'ouvrages publiés dans cette période. Pour y parvenir, à partir de la base bibliographique de l'OCLC WorldCat, il a tracé les titres qui contenaient des termes relatifs aux langues orientales comme « *linguae orientales* », « *philologie orientale* », « *livres orientaux* », « *érudition orientale* », « *auteurs orientaux* » et « *antiquités orientales* ». Il a ainsi illustré que la *philologia orientalis* coïncidait avec la *philologia sacra*, à savoir une philologie des langues du christianisme ancien et des Écritures – un élément que les rédacteurs du *Dictionnaire universel de Trévoux* n'ignorent pas, comme on le verra ci-dessous. Notamment, Stolzenberg a constaté que les ouvrages qui font référence aux langues orientales sont relatifs à l'arabe (23 titres), l'hébreu (22 titres), le chaldéen (22), le syriaque (22 titres), l'éthiopien (10 titres), le persan (10 titres), le samaritain (9 titres), le turc (6 titres), l'arménien (3 titres), le copte (3 titres), le rabbinique/hébreu mishnaïque (2 titres), le phénicien (1 titre), le grec (1 titre). Les quatre langues les plus fréquentes correspondent à l'hypothèse de cette étude.

À partir de ces entités linguistiques, j'ai identifié des mots-clés pour ancrer la lecture conséquente afin de retrouver les entrées concernées ; ces mots-clés sont soit le nom de la langue au masculin, soit la langue comme adjectif au féminin pour désigner la langue, soit les noms du peuple associé : « hébreu » / (langue) « hébraïque » / (les) « Juifs » (appellent/nomment)/ « Hébreux », « arabe » / « Arabes », « syriaque » / « Syriens », « chaldéen » / « chaldaïque » / « Chaldéens ». En parcourant les lettres A et D du *Dictionnaire universel de Trévoux*, j'ai composé une liste d'entrées dictionnaires où apparaissent ces mots-clés en leur associant des métadonnées comme (i) le domaine de l'entrée, (ii) l'orientation religieuse<sup>6</sup>, (iii) la présence d'une entrée du Trévoux dans le *Dictionnaire universel* de Basnage de Beauval et (iv) l'existence de langues orientales aux entrées de ce dernier. Lorsque le domaine est explicitement indiqué dans l'entrée dictionnaire, il est conservé dans les métadonnées ; lorsque cette information est absente, un domaine général est déduit du contexte. Cette pratique permet une première orientation, mais risque d'imposer des éléments extérieurs au corpus. Cela exige donc la vérification des hypothèses qui s'appuient sur ces domaines avec d'autres métadonnées. Quant à l'orientation religieuse d'une entrée

<sup>6</sup>Puisque parfois le domaine d'une entrée est marqué comme « nom propre », mais on découvre que le nom se réfère à un personnage des Écritures (par exemple ABIBON).

(voir « sujet religieux »), la classification est basée sur les références explicites à des livres d'Écritures saintes (voir Genèse, Livres des Rois) ou à des œuvres de la tradition catholique telles que celles consacrées aux saints de l'Église catholique (*Acta Sanctorum*)<sup>7</sup>. Les entrées sur les sectes et l'histoire comparée des religions, lorsqu'elles ne comportent pas de références explicites aux œuvres mentionnées, ne sont pas prises en considération pour cette étude pour la catégorie « sujet religieux », car elles présentent également un intérêt pour les informations ethnographiques et pas seulement pour les études bibliques critiques, intérêt qui nécessiterait une étude dédiée pour l'approfondir. La classification de langues suit la même méthode en se concentrant sur les références explicites des langues et la présence de mots dans une langue orientale dans l'entrée. Dans le cas où la langue n'est pas indiquée par les rédacteurs, la langue est classifiée comme « non spécifiée ». De surcroît, la comparaison avec le *Dictionnaire universel* de Basnage de Beauval est effectuée au niveau d'entrées et de sous-entrées ; en particulier, j'ai cherché à savoir si une entrée du Trévoux existe en tant que telle (ou en tant que sous-entrée) dans le dictionnaire de Basnage de Beauval. Enfin, il faut prendre en compte que même si l'analyse du corpus signale l'existence de références aux langues orientales pour certaines entrées dans le dictionnaire de Basnage de Beauval, parfois les entrées du savant huguenot n'ont pas le détail de celles du *Dictionnaire universel de Trévoux*.

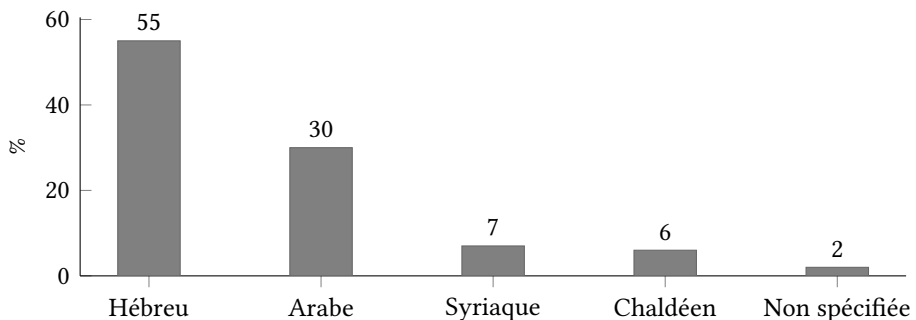


FIGURE 1 : Distributions des langues – lettre A

Dans l'analyse du sous-corpus A on constate, quant à la distribution des langues (voir Figure 1), que 55 % (102 attestations) concerne l'hébreu, 30 % (56 attestations) l'arabe, 7 % (12 attestations) le syriaque et 6 % (11 attestations) le chaldéen. Le sous-corpus D représente une distribution similaire (62 % pour l'hébreu, 21 % pour l'arabe, 8 % pour le chaldéen et 9 % pour le syriaque).

<sup>7</sup>Si la référence à ces œuvres est liée à un exemple d'usage du terme sans commentaires (comme source d'un exemple), l'entrée n'est pas considérée comme référant à un sujet religieux.

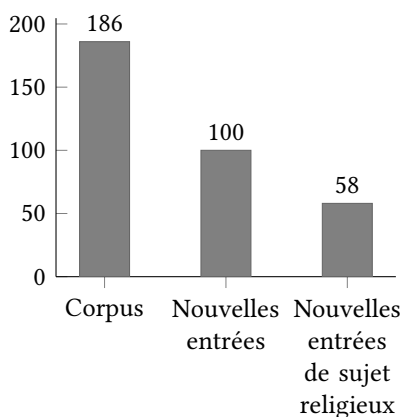


FIGURE 2 : Lettre A

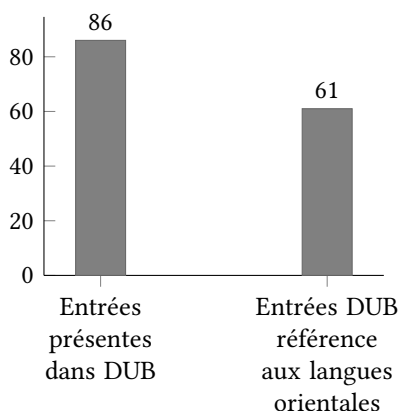


FIGURE 3 : Entrées présentes dans le DUB/sans référence aux langues orientales

Passons à la comparaison entre le *Dictionnaire universel de Trévoux* et celui de Basnage de Beauval : dans le sous-corpus A, les 100 des 186 cas sont ajoutés par les rédacteurs du Trévoux, autrement dit, ils n'existent pas dans le dictionnaire de Basnage de Beauval (voir Figure 2). Parmi ces cas, plus de la moitié (58 cas) comporte un sujet religieux. Pour le sous-corpus D, une tendance similaire est constatée, notamment parmi les 63 entrées qui comprennent des langues orientales, les 41 sont de sujet religieux.

Néanmoins, les rédacteurs trévoltiens ne se limitent pas à ajouter de nouvelles entrées. Ils reformulent les entrées déjà présentes dans le *Dictionnaire universel* de Basnage de Beauval (DUB dans la Figure 3) pour y inclure des références aux langues orientales. Comme représente la Figure 3 ci-dessus, 46 % des entrées du sous-corpus A (86 sur 186) existent déjà dans le dictionnaire de Basnage de Beauval et seulement 13 % (25 sur 86) comprennent déjà des langues orientales.

À partir d'une analyse préliminaire des entrées qui comprennent ces langues, nous pouvons distinguer plusieurs domaines concernés tels que nom propre (ABADDON), géographie<sup>8</sup> (ABARIM), grammaire (ACCENT), religion (ABÉLIENS), mythologie (ABADIR) mais aussi architecture (ABAQUE), botanique (ACACIA), médecine (ABDOMEN), monnaie (ABOKELLE). Cet éventail de sujets semble moins diversifié lorsque l'on approfondit les entrées permettant de trouver des références aux Écritures (57 % du sous-corpus A et 63 % du sous-corpus

<sup>8</sup>Ce domaine n'existe en tant que tel dans le *Dictionnaire universel de Trévoux* de 1721 et il regroupe les villes, villages, rivières, montages, lacs.

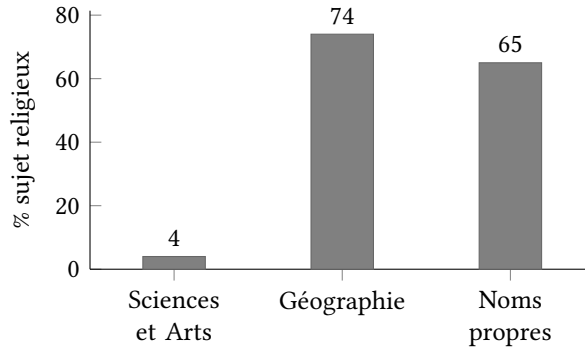


FIGURE 4 : Proportion des sujets réligieux par thématique

D est d'intérêt religieux, même s'il n'est pas toujours évident par le domaine indiqué). Par exemple, 65 % des entrées du domaine « noms propres » et 74 % des entrées géographiques dans le sous-corpus A présente un intérêt religieux. En revanche, cela ne se produit pas dans tous les domaines, puisque pour les entrées de sciences et d'arts ensemble, seule une sur 25 dans le sous-corpus A inclut un sujet religieux (voir Figure 4). Par ailleurs, 95 % des entrées de noms propres et de géographie avec un sujet religieux sont ajoutés par le *Dictionnaire universel de Trévoux*, puisque seules 2 (sur 38) existaient déjà dans le dictionnaire de Basnage de Beauval. Cette augmentation d'entrées sur les noms propres et la géographie ainsi que leur association aux sujets religieux nous amènent à penser que ces informations peuvent avoir un intérêt spécifique pour les rédacteurs du Trévoux. Dans le sous-corpus D, toutes les entrées de noms propres et 83 % des entrées géographiques sont de sujet religieux. Il est donc possible que le texte dictionnaire soit conçu comme un guide ou un glossaire d'interprétation/traduction<sup>9</sup> de textes religieux, en désignant les villes, les rivières, les montagnes et les noms qui apparaissent aux Écritures. Richard Simon, considéré comme le principal rédacteur de la première édition du Trévoux 1704<sup>10</sup>, dans son œuvre *Histoire critique du Vieux Testament*, invite dans une nouvelle traduction de l'Ancien Testament « à joindre un dictionnaire de la langue hébraïque, mais aussi des outils d'histoire,

<sup>9</sup>On constate plusieurs références aux pratiques de traduction : « Cependant quelques nouveaux Critiques ont prétendu que dans une version Française de l'Écriture on ne devoit se servir du mot d'adorer, que lorsqu'il étoit parlé du culte qui se rend à Dieu seul » dans ADORER; « Il est mieux de traduire avec le Pere Amelotte, Abba, mon Pere; ou plutôt avec M. Simon, Abba; c'est-à-dire, mon Père » dans ABBÉ.

<sup>10</sup>Un avis soutenu aussi par les *Mémoires* de Nicéron et le supplément au *Grand dictionnaire historique* de Moréri, GRAVELEAU (2018 : 598).

comme un dictionnaire généalogique, un autre géographique, etc. » (LAPLANCHE 1994), ce qui nous semble coïncider avec les additions d'intérêt exégétique dans le Trévoux. L'exégète dieppois a même remarqué dans sa lettre XXV (SIMON 1730 : 220–226) un tel dictionnaire, à savoir le *Dictionnaire hébreu-italien des mots difficiles de la Bible* (1612) de Léon de Modène (1571–1648), un dictionnaire dont les Juifs « se servent pour apprendre aux enfans à expliquer le Texte Hebreu de la Bible » soulignant l'utilité d'un tel ouvrage pour bien traduire l'Écriture Sainte.

L'importance des noms propres a également été soulignée par nos rédacteurs dans la Préface du dictionnaire :

les noms propres n'ont-ils pas leurs significations, leur étymologie, leur orthographe, leurs variations, leurs nombres, leur usage & leurs difficultez, comme tous les autres? [...] Ne doit-on pas sçavoir comment il faut les traduire des autres Langues, & les exprimer dans la nôtre? [...] Qu'on écrive l'Histoire, sur-tout Ecclésiastique, ou la Vie des Saints, sans être instruit de la manière dont nous avons travesti les noms propres dans notre Langue [...] Combien de ces sortes de mots, qui dans leur origine étant les mêmes, ont néanmoins dans notre Langue presque autant de différentes formes, qu'il y a de différens Saints qui les ont eus, ou de différens lieux, où le culte de ces Saints s'est établi? Quelle confusion l'ignorance de l'usage ne produira-t-elle pas.

Cet intérêt pour les noms propres n'est pas nouveau. Comme le signale Richard Simon dans une autre lettre (« Lettre XXXV » dans SIMON 1730 : 286–294), Robert Estienne (1503–1559) a également rédigé des ouvrages consacrés (i) aux catalogues de noms propres, de villes, de montagnes, etc.<sup>11</sup>, (ii) à la manière dont il faut traduire des noms propres de latin en français<sup>12</sup> (iii) ainsi qu'aux noms hébreux, chaldéens, grecs, latins attestés dans la Bible et leurs interprétations<sup>13</sup>, ce qui résonne avec la pratique constatée dans le *Dictionnaire universel de Trévoux*. Ces éléments nous amènent à considérer qu'il s'agit alors d'un dictionnaire-guide écrit pour reprendre l'autorité d'interprétation (en passant par la traduction) des Écritures, contestée par les huguenots.

Cette autorité tente d'être rétablie par l'expertise en langues orientales, outil indispensable pour « un bon Interprète de l'Écriture » (dans HÉBREU). Dans ce

---

<sup>11</sup>Voir *Dictionarium propriorum nominum virorum, mulierum, populorum, idolorum, urbium, fluviorum, montium caeterorumque locorum quae passim in libris prophanis leguntur* (1541).

<sup>12</sup>Voir *La maniere de tourner toutes especes de noms latins en nostre langue françoise. Reveue et corrigée soigneusement à l'utilité des jeunes enfans* (1546).

<sup>13</sup>Voir *Hebraea, chaldaea, graeca et latina nomina quae in Bibliis leguntur, restituta, cum latina interpretatione. Locorum descriptio ex cosmographi* (1537).

cadre, l'hébreu sert souvent de langue-source d'un mot (étymologie) ou d'élément d'association entre langues (équivalent). Le statut de l'hébreu comme langue de l'Ancien Testament et comme « la plus ancienne langue qu'il ait eue au monde » (dans HÉBREU), ainsi que l'utilité de l'étymologie pour « comprendre mieux la force & la signification des mots »<sup>14</sup> (dans ÉTYMOLOGIE) peuvent justifier la mise en application de la langue hébraïque comme ressource étymologique d'autorité visant à figer une interprétation précise des éléments de l'Écriture. Néanmoins, l'hébreu ne peut pas établir à lui seul une interprétation exacte, puisque dans l'Ancien Testament « il y a même quelques parties qui sont en Chaldaïque, & differens mots Chaldaïques, ou de quelques autres langues, répandus en différens endroits » (dans HÉBREU). En outre, l'hébreu étant une langue source assez obscure (interprétation liée à la verbalisation différente des radicaux ou des consonnes subissant des substitutions lors de la copie), a amené les critiques à s'intéresser à la comparaison des versions existantes<sup>15</sup> afin de reconstituer « le sens véritable ». Ainsi, nous trouvons dans le Trévoux des langues considérées comme des dialectes de l'hébreu<sup>16</sup>, langues aussi pertinentes pour les versions traduites de l'Ancien Testament, dont la présence a un objectif similaire à celui de l'hébreu. Ceci est illustré par la comparaison de différents équivalents<sup>17</sup>, pratique associée à la méthode historico-critique de la Bible liée à l'œuvre de Richard Simon.

Je pense qu'il vaut mieux s'arrêter un instant sur la pensée de ce savant, pour clarifier l'enjeu du dictionnaire. Richard Simon dans son ouvrage *Histoire critique des principaux commentateurs du Nouveau Testament* (1693) a mis en question la validité de l'interprétation biblique de St. Augustin sur laquelle se fonde la tradition catholique (TAMBRUN 2020 : 255). L'interprétation augustinienne - selon Richard Simon dans les *Additions aux « Recherches curieuses sur la diversité des langues et religions » d'Edward Brerewood* (1677) - dérivait de la tradition antérieure en raison de son ignorance de l'hébreu et du grec (LE BRUN 2012 : 261). Ainsi, Richard Simon a tenté, à partir de la comparaison des sources, d'atteindre le sens de l'Écriture avant cette dérivation interprétative TAMBRUN (2020 : 256).

---

<sup>14</sup>Voir « aussi car pour expliquer les termes plus précisément, il faut retourner à la première imposition, afin de parler juste, & de bien entendre ce que l'on dit » dans ÉTYMOLOGIE.

<sup>15</sup>On peut tracer l'intérêt sur les relations entre des langues sémitiques au XVI<sup>e</sup> siècle dans les travaux de Joseph-Juste Scaliger et Isaac Casaubon (LAPLANCHE 1994)

<sup>16</sup>« Les langues Chaldaïque, Syriaque, Ethiopienne, Arabe, &c. sont des dialectes de l'Hébreu, comme les langues Française, Italienne, Espagnole, Portugaise, sont des dialectes du Latin », dans HÉBREU.

<sup>17</sup>À titre d'exemple, « Au reste, quoique le texte Hebreu porte אבנא , Abana, aussi-bien que la Vulgate, cependant le Keri, ou Variante, porte אמנא , Amana, & les Septante l'ont ainsi écrit, Ἀμανὰ » , dans ABANA.

Afin d'être plus précis dans l'interprétation biblique, la connaissance des langues hébraïque et grecque est indispensable pour le recours aux textes originaux et leur comparaison avec la Vulgate (SIMON 1730 : 6). Ceci s'appuie sur la nécessité d'une, disons, professionnalisation de l'exégèse, « le seul Art dont tout le monde se meste » (SIMON 1730 : 3). Richard Simon se garde bien de renoncer à l'autorité de la Vulgate établie par le Concile de Trente (en 1551–1552 et en 1562–1563), mais il soutient que la méthode comparative permet de mieux la comprendre (SIMON 1730 : 4). Cette méthode sert également d'outil de prosélytisme pour les huguenots en démontrant la similitude de la Vulgate, dont ceux-ci ne reconnaissent pas l'autorité, avec les autres versions (SIMON 1730 : 5).

Cet intérêt pour les langues sémitiques s'inscrit dans ce contexte et se manifeste également dans les entrées grammaticales (A, ACCENT) où les rédacteurs nous aident à comprendre les spécificités de ces langues et leurs différences avec le français<sup>18</sup> tout en montrant leur expertise qui justifie leur capacité à corriger d'autres auteurs. Par exemple, à l'entrée AIN, lettre hébraïque et arabe, son absence dans les langues européennes est soulignée « Nous n'avons point [...] dans nos langues d'Europe » ; les rédacteurs du Trévoux tentent de mettre cette entrée en relation avec les différentes méthodes de traduction/translittération des mots hébreux contenant la lettre ain « א » en grec (« les Septante ont rendu cette lettre de deux manières différentes tantôt sans aspiration, comme dans עֵדֶן qu'ils expriment 'Edèn, Eden ; & tantôt par un Γ, c'est-à-dire un G, comme dans עֲמֹרָה qu'ils traduisent Γόμορρα, Gomorrha ») et en latin (« Quelques Grammairiens qui l'expriment par gn, d'autres par gn »).

Reprenons maintenant les noms propres qui servent non seulement de glosaire pour l'interprétation et la traduction des Écritures, comme mentionné ci-dessus, mais aussi de scène pour démontrer et valoriser la méthode étymologique et comparative ; pour les rédacteurs trévoltiens, « c'est principalement par la connoissance des noms propres, que l'on apprend bien les origines, & les principes d'une Langue, & que l'on se rend habile dans la science des étymologies » (Préface) - un élément essentiel pour leur interprétation des éléments religieux. Selon THÉBAUD-SORGER (2015), la mise en scène d'objets et de phénomènes scientifiques, comme les démonstrations publiques d'expériences d'électricité de l'abbé Nollet ou l'expérience de Rouen de Blaise Pascal sur l'existence du vide, peut être considérée comme constitutive de la science moderne. Le public de la cour, puis des salons et des sociétés savantes, a privilégié ce « théâtre

<sup>18</sup>« C'est inutilement que la plupart des Grammairiens comparent la Lettre a des Latins & des François, avec l'aleph des Hébreux & l'eliph des Arabes ; parceque ces deux Lettres n'ont aucun rapport avec notre a, si ce n'est qu'elles sont les premières dans l'Alphabet Hébreu & dans celui des Arabes ; mais elles ne sont pas des voyelles comme dans la Langue Française », dans A.

de la preuve ». Dans le contexte d'un dictionnaire adressé à un grand public, il faut prendre en compte la valeur de ce spectacle textuel (d'une méthode qui tente de s'établir comme autoritaire) comme outil de valorisation et de vulgarisation. Pour illustrer cette démonstration, examinons l'analyse étymologique méticuleusement menée par les rédacteurs du *Dictionnaire universel de Trévoux* dans l'entrée ARAIGNÉE : ils commencent par présenter l'hypothèse erronée que quelques-uns, comme Basnage de Beauval dans son dictionnaire, soutiennent selon laquelle ce mot vient du grec « ἀράχνη » ou même « ἀραιός ». Ils présentent ensuite leur propre hypothèse qui est « bien plus vraisemblable » en recherchant les origines de ce mot dans le mot hébreu « אַרְגָּה » qui signifie « faire un tissu ». Ils mettent en évidence les explications erronées liées à l'hypothèse hébraïque (comme celle dans le dictionnaire de Moréri qui a soutenu que « arag » [אַרְגָּה] signifie araignée en hébreu) et ils soutiennent la leur, à savoir que « arag » signifie faire de la toile, en se référant à l'utilisation de ce mot hébreu chez David de Pomis ainsi que chez un autre rabbin appelé Menahhem. Enfin, ils démontrent en détail les changements phonétiques qui aboutissent du mot hébreu « arag » au mot grec « ἀράχνη », puis au latin « aranea » (« le G Hébreu s'est changé en Grec χ, de même que souvent le χ Grec se change en G Latin ») en soulignant la régularité de ces changements phonétiques par référence à d'autres mots (« comme en χαλβάνη, Galbanum ; λείχω, lingo ; ἄΓχω, ango »), ainsi que par référence à d'autres Grammairiens (Hieroz, P. Thomassin). Ils présentent ainsi leur méthode étymologique qui se fonde sur la régularité des correspondances linguistiques, l'usage ainsi que les travaux d'autres membres de la communauté savante et non sur l'assemblage non critique des idées d'autres auteurs. L'emploi de ces langues dans des entrées sans intérêt religieux comme la Botanique, l'Architecture, la Médecine peut être justifié à la lumière de la mise en scène de la méthode d'interprétation des rédacteurs ; or, une autre interprétation est possible puisque les rédacteurs trévoitiens reconnaissent dans la Préface et dans certaines entrées, l'impact des peuples sémitiques sur les connaissances scientifiques (voir ASTRONOMIE) ; leur présence dans des entrées de domaines scientifiques peut aussi être liée à cela. Les entrées pertinentes de la lettre A et D ne témoignent cependant pas d'une telle interprétation, car il n'y a rien sur l'importance des travaux des peuples sémites en médecine, en botanique ou en architecture.

Ce que semble viser la présence des langues orientales dans le *Dictionnaire universel de Trévoux*, c'est d'offrir un guide de lecture (et par conséquent d'interprétation et de traduction) des textes religieux de façon à valoriser leur méthode philologique. Par conséquent, ce guide sert d'outil de prosélytisme par la reconstitution d'une interprétation « exacte » des Écritures. L'hébreu, l'arabe, le syriaque, le chaldéen étant des langues étroitement associées à l'histoire des Écri-

tures, elles deviennent stratégiques dans le contexte de la méthode exégétique historico-critique. Contrairement aux langues européennes, les langues orientales n'appartiennent pas à la même famille que le français, ces sont des reliques obscures et l'intérêt qu'on leur porte est principalement de nature érudite. Cependant, on constate un intérêt ethnographique par le biais de l'histoire comparée des religions et des entrées du domaine « Relations » qui nécessite une étude plus approfondie.

## **Annexe A Les caractéristiques des « langues orientales » selon le *Dictionnaire universel de Trévoux* (1721), Lettre A.**

### ACHAL

- Extrait : Le mot achal dans les langues orientales signifie tuer & manger.
- Caractéristique linguistique : Existence du mot « achal » avec la signification *tuer & manger*.

### ABYSSINS

- Extrait : Cette langue [le caldéenne] a des caracteres particuliers, & elle n'a pas des points voyalles separez des consones, comme il y en a dans l'Hebreu & dans les autres langues Orientales.
- Caractéristique linguistique : Langues écrites en abjad. Le caldéen et l'hébreu appartiennent aux langues orientales.

### AIN

- Extrait : C'est le nom d'une lettre, qui est une aspiration passée par le nez. Toutes les langues orientales ont le ain.
- Caractéristique linguistique : Lettre aspirée « ain ».

### ASPIRATION

- Extrait : Une seconde raison que ce sont des consones, c'est que les langues Orientales qui n'expirnoient point les voyelles, ont cependant toujours exprimé les aspirations.
- Caractéristique linguistique : Lettres aspirées. Langues qui n'expriment pas de voyelles.

## AUGMENT

- Extrait : Ce terme d'augment syllabique, qui n'est en usage que lorsqu'on parle de la langue Grecque, pourroit être appliqué à la Grammaire des langues Orientales, où la même chose arrive.
- Caractéristique linguistique : Augment syllabique. Grec n'appartient pas aux langues orientales.

## Références

- FURETIERE, Antoine. 1690. *Dictionnaire universel contenant généralement tous les mots françois tant vieux que modernes et les termes de toutes les sciences et des arts ... Par... Antoine Furetière...* 3 t. Amsterdam : Leers. <http://gallica.bnf.fr/ark:/12148/bpt6k3413126b> (17 mai, 2023).
- GRAVELEAU, Sara. 2018. "Les hérésies sont d'utiles ennemies" : *Itinéraire d'Henri Basnage de Beauval (1656-1710), avocat de la République des Lettres et penseur de la tolérance civile*. Université d'Angers. (thèse de doct.). <https://theses.hal.science/tel-02157922> (17 avril, 2023).
- LAPLANCHE, François. 1994. *La Bible en France entre mythe et critique : (XVIe-XIXe siècle)*. Paris : A. Michel.
- LE BRUN, Jacques Le. 2012. L'antiaugustinisme de Richard Simon. *Revue de l'histoire des religions* 2. 257-273. DOI : 10.4000/rhr.7899.
- LECA-TSIOMIS, Marie. 2023. *La guerre des dictionnaires : Le Trévoux aux sources de l'Encyclopédie*. Paris, France : CNRS éditions.
- SIMON, Richard. 1730. *Lettres choisies de M. Simon. Où l'on trouve un grand nombre de faits anecdotes de littérature. Nouvelle édition, revuë, corrigée & augmentée d'un volume; & de la vie de l'auteur par M. Bruzen La Martiniere. Tome premier*. Antoine-Augustin BRUZEN DE LA MARTINIÈRE (éd.). Amsterdam : P. Mortier.
- STOLZENBERG, Daniel. 2015. Les « langues orientales » et les racines de l'orientalisme académique : Une enquête préliminaire. *Dix-septième siècle* 268(3). 409-426. DOI : 10.3917/dss.153.0409. <https://www.cairn.info/revue-dix-septieme-siecle-2015-3-page-409.htm> (17 avril, 2023).
- TAMBRUN, Brigitte. 2020. La Tradition sans l'exclusion. Richard Simon, Bossuet et le socinianisme face aux critères de l'authenticité. *Études théologiques et religieuses* 95(2). 249-273. DOI : 10.3917/etr.952.0249. <https://www.cairn.info/revue-etudes-theologiques-et-religieuses-2020-2-page-249.htm> (17 avril, 2023).

- THÉBAUD-SORGER, Marie. 2015. Spectacles de sciences. In Pestre DOMINIQUE (éd.). Trad. par Bruno PONCHARAL & Agnès MULLER, *Histoire des sciences et des savoirs*, 130-153. Paris, France : Éditions du Seuil.
- WIONET, Chantal. 1998. Du colinguisme dans le Dictionnaire universel dit de Trévoux. *Langage & société* 83(1). 97-111. DOI : 10.3406/lsoc.1998.2815. [https://www.persee.fr/doc/lsoc\\_0181-4095\\_1998\\_num\\_83\\_1\\_2815](https://www.persee.fr/doc/lsoc_0181-4095_1998_num_83_1_2815) (17 avril, 2023).
- WIONET, Chantal. 2006. L'esprit des langues dans le Dictionnaire universel de Trévoux (1704-1771). *Dix-huitième siècle* 38(1). 283-302. DOI : 10.3917/dhs.038.0283. <https://www.cairn.info/revue-dix-huitieme-siecle-2006-1-page-283.htm> (17 avril, 2023).



# Chapter 13

## Musical terms of Greek and Italian origin in Ottoman Turkish lexicography

Agata Pawlina

Jagiellonian University, Poland

Turkish musical terminology is a highly specialized field of inquiry, dealing not only with the linguistic aspects of the contemporary Language for Specific Purposes (LSP) used by Turkish musicians and musicologists, but also with complex socio-cultural issues originating from the history of the reforms of both Turkish music and Turkish language. In this chapter, the author focuses solely on the Ottoman Turkish era. She presents 64 musical terms of Greek and Italian origin collected from 10 Ottoman Turkish dictionaries (17th-19th centuries). She puts the linguistic data into a wider context focusing on two specific diachronic aspects of her broader research on Turkish terminology of art music: 1) how the Greek and Italian languages used by the inhabitants of the Ottoman Empire influenced terminology of the makam music, and 2) to what extent the westernizing reforms of the Ottoman musical culture conducted in the first half of the 19th century are reflected in the lexicography of the period. To answer both questions, she employs a sociocultural understanding of lexicography by seeing the lexicographer as an active observer of not only the language which he describes in a dictionary, but also of the culture of the people who use that language.

### 1 Introduction

Turkish musical terminology is a highly specialized field of the Oriental Studies, which until now has not been researched by linguists. It has gained some interest amongst musicologists (Farmer 1936, Kalender 1981, Bulut 2009, Tohumcu 2009, Yağın 2013) and experts specialized in the Ottoman literature (Çetin 2009, Kadioğlu 2012), however not as a leading research field of their scholarly work.



After spending the last few years examining and analysing the Ottoman Turkish (*Osmanlıca*) and Turkish (*Türkçe*) musical terminology, it has become apparent to me, that to fully understand this interdisciplinary subject one must be well versed in both, the history of Turkish music and the – not less turbulent – history of Turkish language. The socio-cultural changes and the westernizing reforms that took place in Turkey in the 19th and 20th centuries affected both of those disciplines of human creativity. My main goal was to establish whether one can trace a change in Turkish musical culture happening over time by investigating pages of the lexicographical sources from several historical periods. In this chapter, I will solely focus on the Ottoman era, i.e., before 1923.

The turn of the 16th and 17th centuries witnessed the emergence of the *Ottoman idiom* within Middle Eastern art/urban music, distinct from the earlier Arabo-Persian classical canon (Behar 2006, Feldman 1996, Greve 2015, Reinhard & Reinhard 1984, Reinhard et al. 2001, Wright 1992). Later, the modernizing and westernizing reforms of the 19th century led to the introduction of the European (especially Italian) art music in Istanbul (Aracı 2002, 2006, Komsuoğlu & Turan 2007, Özalp 2000, Turan 2004, 2011a,b, Yöre 2011). Since then, both musical traditions have professionally developed concurrently, but independently from each other in Turkey. To describe this phenomenon I use the term *dualism of Turkish art music*. It means that the genre of *art music* in Turkey can be further divided into two sub-genres, which (in oversimplified terms useful for my terminological, not an in-depth musicological research) I titled the *Eastern-style art music* and the *Western-style art music*.

The LSP (Language for Specific Purposes) used by Turkish musicians and musicologists is also *dualistic*. It signifies that a practitioner of the Eastern-style Turkish art music uses a different vocabulary while describing musical pieces, than a musician educated in the Western-style music and who works in an opera theatre or a philharmonic orchestra. However, there is a part of specialized vocabulary that is mutually used by practitioners and theoreticians of both musical genres (e.g., verbs, some adjectives and denominations of such *objective* acoustic phenomena like ‘sound’, ‘pitch’, ‘timbre’, etc.). My recent research (Pawlina 2021) proved that the dualism of Turkish musical terminology can be traced back to the mid-19th century – the same period when European-style music was officially incorporated into the musical culture of Constantinople. This multi-layered complexity makes the lexicon of the LSP used by the Turkish musicians and musicologists, a very interesting, however currently under-researched, field of Turkish linguistics.

In this chapter, I would like to focus on a specific diachronic, etymological aspect of the subject matter. My investigations showed that over the centuries,

the Turkish musical terminology expanded by borrowings mainly, while word-formation and terminologization played a minor role in the process.<sup>1</sup> Thus, it should be regarded as a “lexicon built on the conceptual foundation of a foreign language” (Lukszyn & Zmarzer 2006: 59). To be precise, the two groups of foreign languages: Persian and Arabic – in the case of terminology of the Eastern-style art music; and European languages (Italian in the 19th century, currently mainly French) – in describing theory and practice of the Western-style art music.

It is popularly believed that musical terminology consists mainly of the Italian and Greek (or Latin < Greek) borrowings, hence, the idea for the subject of this paper. It has already been established that this belief is not entirely correct within the European context (Dąbkowski 1997). Here, I would like to ascertain the origin and the role that the Greek and Italian loan-words played in the Ottoman Turkish music vocabulary that was generally dominated by the Arabic and Persian borrowings. To do so, I will present all terminological units of this type, which I have collected from 10 mono-, bi- and multilingual, general and specialized Ottoman Turkish dictionaries. I will provide their lexicological analysis and describe historical and socio-cultural background that led to the introduction of this vocabulary into Turkish language.

## 2 The Ottoman Turkish lexicography: Methodology & sources

The history of the Ottoman Turkish lexicography in the broadest methodological approach encompasses all dictionaries written in the Ottoman lands since the emergence of the dynasty in the late 13th century until the proclamation of the Republic of Turkey in 1923 (Aksan 1998, Ölmez 1994, Prokosch 2011). For centuries, lexicographical works authored by Ottomans focused on Arabic and Persian vocabulary rather than on Turkish lexicon. This situation changed in the 19th century, when the Ottoman intellectuals deemed the reforms necessary and they simplified their written language. At the same time the first monolingual Ottoman Turkish dictionaries and grammars were published in Istanbul (Lewis 1999: 5–32, Prokosch 2013, Siemienieć-Gołaś 2015, Stein 1990: 335). Before that, those who believed essential to compile Ottoman Turkish glossaries for practical and, later on, also scientific purposes, were European merchants, missionaries,

---

<sup>1</sup>Consequently, native Turkish words account for only 25% of terms used to describe the dualistic art music. Such etymological structure – with the minority of vocabulary being native words – is not surprising. Similar structure can be observed in musical terminology of the most of the European languages (Dąbkowski 1997).

diplomats and scholars-orientalists (Rocchi 2011, Schweickard 2014, Stachowski 1995).

Dictionaries have proven to be a reliable source of terminological data (Galinski & Budin 1999, Harras 1989, Miodunka 1989, Sager 1989), therefore, for the purposes of this current research, I have selected 10 Ottoman Turkish lexicographical works as a source of material. I have narrowed the vast period of the history of Turkish lexicography down to the 17th-19th centuries according to the state-of-the-art on the periodization of the Ottoman Turkish music summarized in the previous section. Of the selected dictionaries, six represent the general bilingual type of dictionaries: Italian-Turkish (Siemieniec-Gołaś 2005), French-Turkish (Kieffer & Bianchi 1835, Sāmī 1885, Stachowski 2002) and English-Turkish (Redhouse 1861, *Redhouse Sözlüğü* 2017); two are general multilingual dictionaries (Korabinsky 1788, Meniński 1680); one is a French-Turkish terminological dictionary of arts, science and medicine (Tinghir & Sinapian 1891), and one is a monolingual specialized dictionary of the Ottoman Turkish art music (Uz 1892).

They represent several varieties of sources that are typical for the research on Turkish historical lexicography:

1. dictionaries written in both the Arabic<sup>2</sup> and Latin scripts, for the Ottoman Turkish lexical items, and their definitions in the European languages, respectively (Kieffer & Bianchi 1835, Meniński 1680, Sāmī 1885, Redhouse 1861, *Redhouse Sözlüğü* 2017, Tinghir & Sinapian 1891);
2. dictionaries written in the Latin script for both the Turkish lexical items and their definitions (Korabinsky 1788, Siemieniec-Gołaś 2005, Stachowski 2002);<sup>3</sup>
3. dictionaries written in the Ottoman Turkish in the Arabic script (Uz 1892).

The latter deserves a special attention. *T'alīm-i mūsīki, yahūd, mūsīki iştīlāhātı* written by İsmail Kâzım Uz, printed in Constantinople in 1892–93 (*AH1310*), is the first and (to the best of my knowledge) the only specialized dictionary of the Ottoman Turkish music written by an Ottoman. It is very modern in its

---

<sup>2</sup>I use the term “Arabic script” as a simplification for Ottoman Turkish alphabet based on Arabic and Persian letters, which was used in the Ottoman Empire until its replacement by the Latin-based alphabet in 1928 (Lewis 1999).

<sup>3</sup>These dictionaries can be considered as a part of the vast group of the so-called Turkish *Transkriptionstexte* (Stachowski 1995: 185–86) in which Turkish words are transcribed into Latin script influenced by the orthography of the language native to the author of a given text (in the case of my sources – German, Italian and French).

macro- and microstructure. It describes ca. 500 musical terms presented in an alphabetical order. Given the turbulent history of the Ottoman Turkish art music after the proclamation of the Republic of Turkey, the value of the musicological and linguistic content of Uz's definitions is invaluable. These contain as well as the semantic data that is important for the subject matter of this paper, which I will describe below.

From those selected sources, in total I have collected ca. 1300 musical terms. I analysed them in terms of semantics, morpho-syntactic structure and etymology. Below, in Sections 4 and 5. I present lexical items of the Greek and Italian origin. Etymological information provided in these sections, has been derived from the historical dictionaries of Turkish and the state-of-the-art dictionaries of the Ottoman Turkish (Aksoy & Dilçin 2009, *Kubbealtı Lugati* 2023, Parlatur 2017, *Red-house Sözlüğü* 2017, M. Stachowski 2019, Tietze 2016) and it has been compared with the information on the European terminology and its origins (Dąbkowski 1991, 1997, 2004).

In Section 6, I examine the data in a wider context in order to answer the following questions: 1) how Greek and Italian languages used by the inhabitants of the Ottoman Empire influenced terminology of the Ottoman Turkish art music, and 2) to what extent the westernizing reforms of the Turkish musical culture conducted in the first half of the 19th century are reflected in Turkish lexicography of the period. To answer both of these questions, I employ a sociocultural understanding of lexicography (Doroszewski 1970: 37, Hausman 1989). I regard the lexicographer as an active observer of not only the language that he describes in his dictionary, but also of the culture of people who use that language (Ligara 2017).

### 3 Transcription and abbreviations

For convenience, in the following lists of the Ottoman Turkish vocabulary (Sections 4 and 5) terms are given in the alphabetical order of the Latin script. I employ transcription system from Arabic into the Latin characters developed by the scholars of the Project *Corpus Musicae Ottomanicae* based in Münster, published on the CMO website in October 2022.<sup>4</sup>

Each of the Ottoman Turkish terms is numbered and given in the Latin script followed by all spelling variations in the Arabic script found in the sources. Below, I provide the English meaning of the term or – when there is no English

---

<sup>4</sup>See: *Guidelines for the Transcription of the Ottoman Lyric from Arabic into Latin Characters – Revised edition*: <https://www.uni-muenster.de/CMO-Edition/en/publikationen/publikationen.html> (15-03-2023-03-15).

equivalent – its short definition. In the next line, I present etymology and (if applicable) the morphological and/or syntactic structure of the term. The last part of the description of each numbered lexical item is a list of abbreviated primary sources in which a term has been found. The sources are provided in chronological order, from the oldest, with the volume (if applicable), and page number in brackets ( ), along with the original Latin transcription given by the author of the dictionary (if applicable). For example: Men(144: *arghanūn*) means that Meniški in his *Thesaurus Linguarum Orientalium* transcribed Ottoman Turkish ارغانون [erġanun] as *arghanūn*.

The following is the list of the abbreviated sources:

|     |                         |     |                         |
|-----|-------------------------|-----|-------------------------|
| K-B | Kieffer & Bianchi 1835  | R   | Redhouse 1861           |
| Kor | Korabinsky 1788         | ŞS  | Sāmī 1885               |
| Men | Meniški 1680            | T-S | Tinghir & Sinapian 1891 |
| Mol | Siemienieć-Gołaś 2005   | Uz  | Uz 1892                 |
| NR  | Redhouse Sözlüğü (2017) | Vig | Stachowski 2002         |

When the definition of the given term in the source is particularly interesting, it is quoted in English translation. If the exact etymological root is not confirmed in the sources, the indication [?] is stated next to the origin proposed by me.

## 4 Musical terms of Greek origin

Among the Ottoman Turkish musical terms collected from the sources described earlier, I have found only 10 Greek loan words and 3 derivatives consisting of a Greek stem with a Turkish suffix creating *nomina agentis*: جى [read as: -ci, -ci, -cu or -cü]. These 13 terms belong to the following semantic fields: a) the general musical terms (i.e., those that can be used in the context of both, the Eastern- and the Western-style music), b) the theory of the Ottoman Turkish art music, c) the practice of the Ottoman Turkish art music, d) the organology (i.e., the names of instruments and their parts):

(1) erġanun ارغانون , ارغنون

‘organ’<sup>5</sup>

< Greek ὄργανον [órganon] ‘instrument, tool; musical instrument’

---

<sup>5</sup>Even though my sources translate this word as a name of the specific instrument known as an ‘organ’, other Ottoman sources indicate that it was a rather ambiguous term designating “any wind instrument” (see: Farmer 1931: 3).

Mol(96: *arghanan*); Men(144: *arghanūn*); Kor(4: *arganon*); Vig(95: *èrgha-noun*, *èrghoun*); K-B(I,24: *arghanoun*); R(549); T-S(II,202); ŞS(69: *erghanoûn*); NR(345)<sup>6</sup>

- (2) *erġanuncu* ارغانونجى  
'organist'

< ارغانون [erġanun] + جى [-cu]  
T-S(II,201)

- (3) *fiske* فسكه

a name of a playing technique on the plucked string instruments

< Greek φούσκα [fóuskē] 'to flip, to flick'  
NR(376)

- (4) *katıkofti* قاتاڤوتى

a name of an *usul* (i.e., a rhythmic pattern in the Ottoman Turkish art music)

< Greek κατακόπτω [katakoptō] 'to cut' [?]  
NR(617) "G[ree]k slang swagger; Or[iental] mus[ic] same as *müsemmen*"

- (5) *katıkofti* قاتيقوتى

see above: *katıkofti*

NR(619) "Or[iental] mus[ic] a fast rhythmic pattern of eight beats"

- (6) *komma* قوما

'comma', a name of the smallest interval

< Greek κόμμα [kómma] 'piece'  
Uz(43); NR(672)

- (7) *laġuta* لاڡوته

see below: *lauta*

Uz(46)<sup>7</sup>

<sup>6</sup>*Kubbealtı Lugati* informs that the name of a zither-like *kanun* in its original Arabic form قانون [qānūn] was derived from Greek ὀργάνον, see: <http://lugatim.com/s/kanun>.

<sup>7</sup>This spelling is also confirmed in the, so-called, *Tarama Sözlüğü* – the biggest historical dictionary of Turkish language – based on the Ottoman Turkish sources from the 16th, 18th and 19th centuries (Aksoy and Dilçin 2009, IV,2661; V,3395).

(8) lauta لاوته, lauṭa لاوطه

‘lute’ (of the European lutes family)

< Greek λαούτο [laouto] ‘lute’

Mol(96: *lauta*) “liuto, stromento di sonare”; R(492); ŞS(936: *laouta*); NR(705: *lavta*) “an instrument like *ud*, now out of use”<sup>8</sup>

(9) lauṭacı لاوطه جی

1. ‘lute player’, 2. ‘maker of seller of lutes’

< لاوطه [lauṭa] + جی [-ci]

NR(705)

(10) mandal ماندال

a name of a metal peg of a *kanun* (a zither-like instrument used in the Ottoman Turkish art music)

< Greek μανταλι [mantáli] ‘bolt, lock’

NR(729)

(11) mūsīkī موسیقی, mūsīkī موسیقی

‘music’

< Arabic موسیقی [mūsīqī] ‘music’ < Greek μουσική [mousiké] ‘the art of muses, music’

Mol(150: *musiki*); Men(5024: *mūsykȳ*); Men(5025: *mūsyky*); Vig(204: *mousiqi*); K-B(II,1046: *mouciqy, moucyqy*); R(526); ŞS(1098); T-S(II,153); Uz(49); NR(801)

(12) mūsīkīci موسیقچی

‘musician’

< موسیقی [mūsīkī] + جی [-ci]

Mol(41: *miskgi*); Vig(204: *mousiqidji*)

(13) sirto سیرتو

a name of an instrumental genre in the Ottoman Turkish art music; since the 19th century it was a part of the *fasıl* suite

< Greek σιρτός [sirtós], a name of a folk dance

Uz(30-31); NR(1022)

---

<sup>8</sup>The, so-called, *Kieffer-Bianchi Dictionary* is the only source to list another form of this term: لاوت [laut] transcribed into Latin characters as *laut* and *lävout* – K-B(II,693).

## 5 Musical terms of Italian origin

In the collected material, there are more Ottoman Turkish musical terms borrowed from Italian. In total, I have found 50 lexical items comprising: 41 single-word loan-words, 3 derivatives with an Italian stem and a Turkish suffix creating *nomina agentis*: جى [read as: -ci, -cı, -cu or -cü] and 1 substantivized derivative with an Italian stem and a Turkish suffix creating adjectives لى [read as: -li, -lı, -lu or -lü]. There are also 6 compounds in the form of the *indefinite izafet* (Lewis 1991: 41), i.e., a group of two nouns, where the first noun remains in the absolute form and qualifies the second noun, which has the possessive suffix of the third person: سى [read as: -si, -sı, -su or -sü] or ى [read as: -i, -ı, -u or -ü].

The collected terms belong to the following semantic fields: a) the general musical terms, b) the theory of the Western-style art music, c) the names of the symbols in the musical notation, d) the musical practice of the Western-style art music, e) the musical practice of the Eastern-style art music, f) organology.

(14) *alafranga* آلا فرانغه

a general term for the 'European music' (i.e., the Western-style music)

< Italian *alla Franca* 'in the French style'

NR(42)

(15) *alaturka* آلا تورقه

a general term for the 'Turkish music' (i.e., the Eastern-style music)<sup>9</sup>

< Italian *alla Turca* 'in the Turkish style'

NR(44)

(16) *alaturkacı* آلا تورقه جى

1. a practitioner of Turkish music (i.e. the Eastern-style music), 2. a lover of Turkish music

< آلا تورقه [alaturka] + جى [-cı]

NR(44)

<sup>9</sup>Both terms, *alafranga* and *alaturka*, emerged in the 19th century Ottoman Empire and were used to describe 'Occidental-' and 'Oriental way' in many other elements of culture, e.g., calendar, hour, clothing, etc. However, in my sources and in literature on music from the period they are often used not only as adjectives but also as nouns designating 'Turkish music' and 'European music' (see: Özalp 2000: 10, Öztuna 1969: 2-27).

- (17) *آلتو* *alto*  
'alto, contralto'  
< Italian *alto* 'alto, contralto'  
NR(54)
- (18) *آرمونیوم* *armonyum*  
'harmonium, pump organ'  
< Italian *armonium* (< French) 'harmonium, pump organ'  
T-S(I,394); NR(74)
- (19) *آریا* *arya*  
'air, song'  
< Italian *aria* 'aria, air, song'  
NR(72,76)
- (20) *باندہ* *banda*  
see below: *bando*  
NR(129)
- (21) *باندو* *bando*  
1. 'military band', 2. a name of the 19th-century official Ottoman military band, which replaced the former *mehterhane* after the abolishment of the Janissary corps in 1826; its first conductor was Giuseppe Donizetti, who had been invited by sultan Mahmud II to perform the westernizing reforms in the musical culture of the Ottoman court (Aracı 2002, 2006)  
< Italian *banda (musicale)* '(musical) band'  
NR(130)
- (22) *باصو* *başo*  
'bass'  
< Italian *basso* 'bass'  
T-S(I,53); NR(134)<sup>10</sup>

---

<sup>10</sup>Names of voices are also listed in the first, English-Turkish edition of the so-called *Redhouse Dictionary* (Redhouse 1861) but there are no Ottoman Turkish equivalents yet, only a short description of each voice type.

- (23) başo kemānı باصو کمانی  
'double bass'  
< Italian *basso* 'bass' + Ottoman Turkish (<Persian) کمان [kemān] 'violin'  
+ 3SG.POSS  
T-S(I,148)
- (24) batuṭa باتوطه  
'bar, measure'  
< Italian *battuta* 'bar, measure'  
Uz(11)
- (25) çimbalo چیمبالو , چیمبالو  
'harpsichord, cembalo' or/and 'piano'  
< Italian *cembalo* 'harpsichord, cembalo'  
R(461) "keyboard"; R(576) "piano", "pianoforte";  
T-S(I,120) "clavicorde"
- (26) çinbalo چنبالو  
see above: *çimbalo*
- (27) do دو  
'do (in solmisation)'  
< Italian *do*  
R(256); T-S(I,225) "a name of the first note in singing [sic!]; [same as] *çargâh* [in the scale used in the Ottoman Turkish music]"; Uz(24) "this is a name of the [pitch] *çargâh* in the *alafranga* [music]; its other name is *ut*; it is the first pitch in the musical scale of the *alafranga* music"; NR(306)
- (28) fa فا  
'fa (in solmisation)'  
< Italian *fa*  
R(314); T-S(I,314); Uz(41); NR(357)

- (29) فالصو *falšo*  
'false note'  
< Italian *falso* 'false [note]'  
Uz(41–42) "with this word one can describe a mistake in singing"; NR(359)
- (30) فانفار *fanfar*  
'fanfare, trumpet call'  
< Italian *fanfara* 'fanfare, trumpet call'  
T-S(I,317)
- (31) فانطازيه , فانتازيا *fanṭaziya*, *fantaziya*  
'fantasy, fantasia'  
< Italian *fantasia* 'fantasy, fantasia'  
T-S(I,314); NR(360)
- (32) فلاوته *flauta*  
'flute'  
< Italian *flauto* 'flute'  
R(338); T-S(I,337); Uz(42); NR(377: *flâvta*)
- (33) فلاوته جى *flautacı*  
'flutist'  
< [flauta] + جى [-cı]  
R(338); NR(377: *flâvtacı*)
- (34) اسقله , اسقاله *ısqala*  
'musical scale'  
< Italian *scala* 'scale'  
T-S(I,241) "échelle diatonique"; Uz(6–7); NR(503)
- (35) قرانته [sic!] *qıranṭa*  
'clarinet'  
< Italian *clarinetto* [?]  
Uz(43)<sup>11</sup>

---

<sup>11</sup>It is possible that K. Uz made a spelling mistake here. According to H.G. Farmer a different spelling of this term: قرناطه [qırnaṭa] was used in the Ottoman Turkish sources at least since the 17th century Farmer (1936: 24–25). However, I did not find it in the dictionaries.

- (36) kitara كيتاره  
'guitar'  
< Italian *chitarra* 'guitar'  
Uz(44)
- (37) konçerto كونچرتو  
'concert' (an event, not a genre)  
< Italiano *concerto* 'concert'  
T-S(II,153) "musique de concert"
- (38) kōntrabaşo كونترا باصو  
'double bass'  
< Italian *contrabasso* 'double bass'  
T-S(I,53)
- (39) koro قورو  
'choir'  
< Italian *coro* 'choir'  
NR(675)
- (40) la لا  
'la (in solmisation)'  
< Italian *la*  
R(464); T-S(II,45); Uz(45); NR(700)
- (41) longā لونغه  
an instrumental genre in the Ottoman Turkish art music; emerged in the late 19th century from the Rumelian Gypsy music; sometimes used as the last part of the *fasıl* suite  
< Italian *lungo* 'long'  
R(713); NR(713)
- (42) mayestro مايسترو  
'conductor of the orchestra'  
< Italian *maestro* 'master, teacher'  
NR(739)

(43) mezosoprano مه زوسوپرانو

‘mezzo-soprano’

< Italian *mezzo soprano*

NR(772)

(44) mızıka مزیکه

see below: *muzıka*

NR(774)

(45) mızıkalı مزیکه لی

a member of the Sultan’s military orchestra

< [mızıka] مزیکه + [li] لی

NR(774)

(46) mi می

‘mi (in solmisation)’

< Italian *mi*

R(511); T-S(I,237); T-S(II,124); Uz(49)

(47) a. muzıka موزیکا

1. ‘military band’, 2. ‘military music’

< Italian *musica* < *banda musicale*

ŞS(1097: *mouzyqa*); NR(806)<sup>12</sup>

b. asker muzıkası عسکر موزیکه سی

1. ‘military band’ 2. ‘military music’

< Ottoman Turkish (<Arabic) عسکر [‘asker] ‘soldier’ + [muzıka] موزیکه + 3SG.POSS

R(526)

---

<sup>12</sup>The official name of the *Bando* (see above: 2. meaning of the term no. 21) was موزیکه همایون [muzıka-yı hümāyūn], literally ‘the military band of the sultan’. It is an interesting compound term in the form of the *Persian izafet* combining Italian and Persian loan-words together. It was also a name of the palace music school founded by Giuseppe Donizetti to train musicians in the Western-style art music (Özcan 2020).

- (48) muzıkacı موزیقاجی  
 a member of a military band  
 < موزیقہ [muzıka] + جی [-cı]  
 T-S(II,152) “musicien.enne”
- (49) a. noṭa نوطہ , nota نوتہ  
 1. ‘musical note’ 2. ‘written music’  
 < Italian *nota* ‘musical note’  
 ŞS(1138: *nota*); T-S(I,56); Uz(51); NR(891)
- b. musıki noṭası موسیقی نوطہ سی  
 1. ‘musical note’ 2. ‘musical notation’  
 < [musıki] موزیقہ + [noṭa] نوطہ + 3SG.POSS  
 T-S(II,176)
- (50) nota kitābı نوتہ کیتابی  
 ‘musical score’  
 < [nota] نوتہ + Ottoman Turkish (<Arabic) [kitāb] ‘book’ + 3SG.POSS  
 R(526)
- (51) noṭa miḳyāsı نوطہ مقیاسی  
 ‘diatonic scale’  
 < [noṭa] نوطہ + Ottoman Turkish (<Arabic) [miḳyās] ‘scale’ + 3SG.POSS  
 T-S(I,241)
- (52) opera اوپرا  
 ‘opera’  
 < Italian *opera*  
 R(547); ŞS(162); NR(901)
- (53) opera tiyatrosu اوپرا تیاتروسى  
 ‘opera theater’  
 < [opera] اوپرا + [tiyatrosu] تیاتروسى < Italian *teatro* ‘theater’ + 3SG.POSS  
 R(547)

- (54) orkestra اورکستره  
'orchestra'  
< Italian *orchestra*  
T-S(II,199); NR(902)
- (55) primo پریمو  
a term describing the main melody in a musical piece  
< Italian *primo* 'the first'  
Uz(14) "we call that a melody which is played by such instruments as [a lute] *ut*, *keman* ['violin'], [zither-like] *kanun*, [a lute] *tanbur*, etc."
- (56) piyano پیانو  
'piano'  
< Italian *piano*, *pianoforte*  
R(596); §S(297: *piano*); T-S(II,267); Uz(15); NR(937)
- (57) re ره  
're (in solmisation)'  
< Italian *re*  
R(623); T-S(I,180); T-S(II,353); Uz(28); NR(950)
- (58) si سی  
'si (in solmisation)'  
< Italian *si*  
R(685); T-S(I,390); T-S(II,421); NR(1015)
- (59) şol صول  
'sol (in solmisation)'  
< Italian *sol*  
R(696); T-S(II,428); Uz(37); NR(1026)
- (60) şoprano صوپرانو  
'soprano'  
< Italian *soprano*  
T-S(II,431)

- (61) *tono* طونو  
 ‘tone, pitch’  
 < Italian *tono* ‘tone, pitch’  
 T-S(II,496)
- (62) *viyola* ويوله  
 ‘viola’  
 < Italian *viola*  
 NR(1230)

## 6 Conclusions

The Ottoman Turkish musical terms of Italian origin belong to the newest stratum of that language, the so-called *New Ottoman Turkish* period of its history. The oldest dictionary that recorded them in its definitions is *A Lexicon, English and Turkish Shewing in Turkish, the Literal, Incidental, Figurative, Colloquial, and Technical Significations of the English Terms* written by Sir James W. Redhouse, published in London in 1861. They were absorbed into the Ottoman Turkish during the initial period of the westernization of the Turkish musical culture. Its forerunner was Giuseppe Donizetti (the elder brother of the famous opera composer - Gaetano), who, at the request of Sultan Mahmud II, undertook a complete reorganization of the Ottoman military band in accordance with the European fashion and musical style. His work had begun in 1827, and within a few years, the former “Eastern-style” official military band of the Sultan called *Mehterhâne*, was replaced by the *Muzıka-yı Hümayün* (colloquially called *Bando*) operating within the framework of the new palace music school of the same name (Aracı 2002, 2006, Komsuoğlu & Turan 2007, Pawlina 2014, Sanal 1964, Yöre 2011: 278–95). It seems that, along with the Western instruments and the Italian musicians, that Donizetti had also brought to Constantinople was the basic terms used in the teaching of the European musical theory and practice.

The most prevalent terms incorporated into Turkish during that period are the pitch names (lexical items no. 27, 28, 40, 46, 57, 58, 59). Until today, the solmisation is the main system used by the Turkish musicians to describe the pitches and keys in the Western-style art music (e.g., Tur. *La-majör* ‘A major’, Tur. *la-minör*

‘A minor’).<sup>13</sup> The continuous use of solmisation in the LSP of Turkish musicians and musicologists can be traced back to the reforming and pedagogical efforts of Donizetti and his colleagues. The same can be said about incorporating the names of the voice types (lexical items no. 17, 22, 43, 60) and their use within the names of instruments (no. 23). The derivatives and the compound terms of mixed origin such as *flautacı*, *muzıkacı*, *başo kemāni*, ‘*asker muzıkası* and *noça miğyāsı* prove that in the second half of the 19th century new Italian loan-words were fully assimilated.<sup>14</sup>

Most importantly, the analysis of the Italian vocabulary presented above allows us to conclude that the 19th-century westernizing reforms of the Turkish musical culture are indeed perfectly reflected in the Ottoman Turkish lexicography of the period. The names of the European instruments, instrument players, musical bands and genres mirror the new elements of the cultural life of the Ottoman elite in Constantinople: the state military band (lexical items no. 20, 21, 30, 32, 33, 35, 42, 44, 45, 47, 48), opera (no. 17, 19, 22, 29, 39, 43, 52, 53, 60) and its orchestra (no. 23, 31, 37, 38, 54, 62), and the chamber ensemble instruments that would be played during private gatherings at the palaces (no. 18, 25, 26, 36, 56). Additionally, the theoretical terms recorded in specialized dictionaries of the period reflect the new pedagogical methods (no. 24, 34, 49, 50, 51, 55, 61), especially the five-stave musical notation, which until the 19th century had not been used in Turkey as a method of instruction of music. The new approach to the musical practice and theory can also be seen in the musicological content of definitions published in the specialized French-Turkish dictionary prepared by Tinghir & Sinapian (1891). The Ottoman scholars described the solmisation names of pitches by comparing them to their “equivalents” on the Ottoman Turkish musical scale (see: lexical item no. 27). The same practice had been employed by the musicologist and pedagogue İsmail Kâzım Uz (1892).

In Section 5 there is only one Italian borrowing with a different historical background. The name of an Ottoman Turkish Eastern-style genre: *longa* is enlisted as an Italian loan-word only in the *Kubbealtı Lugatı*<sup>15</sup>. Other historical and ety-

---

<sup>13</sup>This is the system used in the musical terminology of Romance languages and Russian, while in LSP of English, German, Czech and Polish musicians the basic system of pitch/key naming employs first seven letters of the alphabet (Dąbkowski 1997: 82).

<sup>14</sup>It is worth noting that in the 20th century most of them have been replaced by French borrowings and now, in modern Turkish (*Türkçe*) are out of use, for example: Ott.Tur. *flauta* had been replaced by Tur. *flüt* < Fr. *flûte*; Ott.Tur. *konçerto* – Tur. *konser* < Fr. *concert*, Ott.Tur. *çimbalo* – Tur. *klavsen* < Fr. *clavecin*. An investigation of change in Turkish musical terminology after the 20th-century language reform can be found in my unpublished PhD dissertation (Pawlina 2021).

<sup>15</sup>See: <http://lugatim.com/s/longa> (Accessed: 17.03.2023).

mological dictionaries do not confirm this information. This fact, along with the musicological data on the genre, questions the presented etymology. However, if it is correct, it would be an important clue motivating further investigations regarding the impact of the language used by Italian Levantines on the musical LSP of the period.

Musicological research has already proved that the Greek (*Rūm*) inhabitants of the Ottoman Empire played an important role in the history of the Ottoman art music since the late 18th century. The peak of this trend in musical practice happened in the 19th century (Behar 2006, Feldman 1996, Greve 2015). In this context, it does not come as a surprise that the several technical terms denoting the phenomena of musical practice (lexical items no. 3, 4, 5, 10, 13) and the theory of music (no. 6) had only been recorded during the last period of the history of the Ottoman Turkish language (i.e., the *New Ottoman Turkish*). These terms confirm that the Ottoman Turkish urban music of the 19th century was multicultural not only in the sphere of the sound, but also in its terminology, combining elements of traditions of various ethnic groups that lived in the capital of the Ottoman Empire.

On the other hand, among the Ottoman Turkish musical terms of Greek origin there are also borrowings belonging to the oldest stratum of the language (lexical items no. 1, 2, 8, 9, 11, 12). A denomination of ‘music’ – *mūsīqi* – has the most interesting history. Unsurprisingly, it is the only term that has been found in all 10 dictionaries selected for the current research. It was in use in the Ottoman Turkish since the 14/15th century, when the first Turkish translations of the Arabic and Persian treatises on the theory of music began to emerge (Kalender 1981). The Greek term μουσική [*mousiké*] had been included into Arabic lexicon (موسيقى [*mūsīqi*]) much earlier, when in the 9th-10th centuries Arabic scholars adapted the theory of music of ancient Greeks and transformed it into the roots of the theory of Middle-Eastern music (Danielson et al. 2002, Öztuna 1969: 215,290,341, Reinhard & Reinhard 1984: 52–53). In the 20th century, the term *mūsīqi* was removed from the modern Turkish (*Türkçe*) to be – eventually – replaced by a French loan-word: *müzik* (< Fr. *musique*) used today. Ottoman Turkish term in its contemporary spelling, *musiki*, is described as an “obsolete” or omitted in most of the dictionaries written in the 20th century. Interestingly, in the newest dictionaries, it made a come back: it is sometimes enlisted as a name for the ‘Eastern-style music’ only (not ‘music’ in its broadest sense). The investigation of the written and spoken LSP of the contemporary Turkish musicians and musicologists in terms of the use of both terms, *musiki* and *müzik*, is an open research field.

Both groups of loan-words analysed in this paper – Greek and Italian – should be seen as *cultural borrowings* (Bloomfield 1984). In historical periods indicated in this paper, the Ottomans borrowed both the ideas (theory) and the physical objects (instruments) related to music, and along with them, they incorporated their names with a little or no change in their linguistic form. Based solely on the lexicographical data presented here, it is difficult to ascertain to what extent Greek and Italian languages used by the non-Turkish inhabitants of the Ottoman lands influenced the LSP of Ottoman musicians practicing the art/urban music. However, it can be said that such a linguistic influence existed (especially from Greek) and it requires further investigation. The answer to the second question raised above – how the westernization of the 19th-century Turkish music is reflected in the Ottoman Turkish lexicography – is straightforward: the ongoing change in the musical culture and even the pace of the assimilation of the ideas and the changing popularity of newly brought instruments can easily be traced by investigating the dictionaries.

It is worth emphasizing that my research on musical terms of the Greek and Italian origin found in the Ottoman Turkish lexicography proves that dictionary makers can be interpreted as the attentive observers of the socio-cultural processes surrounding any linguistic (Siemieniec-Gołaś 2015) and extralinguistic (Doroszewski 1970, Ligara 2017) change. Because of that, a diligent reader can trace important extra-linguistic historical data on the pages of dictionaries. In the Ottoman Turkish context, it is particularly important to appreciate the value of information given in definitions written by Ottoman lexicographers (in my case: Şemseddin Sami, Anton B. Tinghir, Krikor Sinapian, İsmail Kâzım Uz), who were not only observers (like their contemporaries of European origin, such as James W. Redhouse, Jean D. Kieffer or Julius T. Zenker) but also, more importantly, the active participants of the Ottoman Turkish culture undergoing a significant change in the 19th century. Current research proves that the analysis of the extra-linguistic content of dictionary articles prepared by such authors may be useful information for research in other scientific disciplines within the framework of the Ottoman Studies – in my case, musicology. Lexicographical data perfectly reflects the socio-cultural change in this very specific element of the Turkish culture, and it can be used as a new source of information for future research on the history of Turkish music.

## References

- Aksan, Doğan. 1998. Türklerde Sözlükçülük, Bugün Türkiye’de Sözlük. *Kebikeç İnsan Bilimleri için Kaynak Araştırmaları Dergisi* 6. 115–118.

- Aksoy, Ömer Asım & Dehri Dilçin (eds.). 2009. *XIII. Yüzyıldan beri türkiye türkçesiyle yazılmış kitaplardan toplanan tanıklariyle tarama sözlüğü*. 3rd edn. Ankara: Türk Tarih Kurumu Basımevi.
- Aracı, Emre. 2002. Giuseppe Donizetti Pasha and the Polyphonic Court Music of the Ottoman Empire. *The Court Historian* 7(2). 135–143. DOI: 10.1179/cou.2002.7.2.004.
- Aracı, Emre. 2006. *Donizetti Paşa: Osmanlı Sarayının İtalyan Maestrosu*. İstanbul: Yapı Kredi Yayınları.
- Behar, Cem. 2006. The Ottoman Musical Tradition. In Suraiya Faroqhi (ed.), *The Cambridge History of Turkey*, vol. 3, 398–407. New York: Cambridge University Press.
- Bloomfield, Leonard. 1984. *Language*. Chicago: The University of Chicago Press.
- Bulut, Mustafa Hilmi. 2009. Türk Müzik Kültüründe Sözel Yapı. In 38. ICANAS (Uluslararası Asya ve Kuzey Afrika çalışmaları kongresi) 10-15.09.2007, ankara, bildirler: 13. Müzik Kültürü ve Eğitimi, 151–66. Ankara: Atatürk Kültür, Dil ve Tarih Yüksek Kurumu. <http://www.ayk.gov.tr/baglantilar/icanas-38/>.
- Çetin, Kamile. 2009. Musiki ve Musiki Terimlerinin İbrahim Raşid Divanı'ndaki Yansımaları. *Journal of Turkish Studies* 4(2). 199–225. DOI: 10 . 7827 / TurkishStudies.628.
- Dąbkowski, Grzegorz. 1991. Perspektywy badań porównawczych nad terminologią muzyczną w różnych językach europejskich. *Poradnik Językowy* 7. 283–286.
- Dąbkowski, Grzegorz. 1997. *Europejska terminologia muzyczna*. Kielce: Wyższa Szkoła Pedagogiczna im. Jana Kochanowskiego.
- Dąbkowski, Grzegorz. 2004. *Siedmiojęzyczny słownik terminów muzycznych: Włosko-francusko-angielsko-niemiecko-rosyjsko-czesko-polski* (Semiosis Lexicographica, vol. 22). Warszawa: Takt.
- Danielson, Virginia, Dwight Reynolds & Scott Marcus. 2002. 3.6. Turkey. In *The Garland Encyclopedia of World Music. Volume 6, The Middle East*, vol. 6, 755–820. New York, London: Routledge.
- Doroszewski, Witold. 1970. *Elementy leksykologii i semiotyki*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Farmer, Henry George. 1931. *The Organ of the Ancients from Eastern Sources*. London: William Reeves Bookseller Ltd. <https://archive.org/details/TheOrganOfTheAncientsFromEasternSources/mode/2up> (13 February, 2021).
- Farmer, Henry George. 1936. Turkish Instruments of Music in the Seventeenth Century. *Journal of the Royal Asiatic Society of Great Britain & Ireland* 68(1). 1–43. DOI: 10.1017/S0035869X00076322.

- Feldman, Walter. 1996. *Music of the Ottoman Court: Makam, Composition and the Early Ottoman Instrumental Repertoire*. Berlin: VWB-Verlag für Wissenschaft und Bildung.
- Galinski, Christian & Gerhard Budin. 1999. Deskriptive und präskriptive Terminologieerarbeitung. In Lothar Hoffmann, Hartwig Kalverkämper, Herbert Ernst Wiegand, Christian Galinski & Werner Hüllen (eds.), *Fachsprachen. Ein internationales Handbuch zur Fachsprachenforschung und Terminologiewissenschaft*. 2. Halbband, 2183–2207. Berlin, New York: Walter de Gruyter.
- Greve, Martin (ed.). 2015. *Writing the History of "Ottoman Music"*. Würzburg: Ergon Verlag in Kommission.
- Harras, Gisèle. 1989. Wörterbücher als Hilfsmittel der linguistischen Forschung. In Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta (eds.), *Wörterbücher: Ein internationales Handbuch zur Lexikographie*, vol. 1, 159–163. Berlin-New York: Walter de Gruyter.
- Hausman, Franz Josef. 1989. Die gesellschaftlichen Aufgaben der Lexikographie in Geschichte und Gegenwart. In Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta (eds.), *Wörterbücher: Ein internationales Handbuch zur Lexikographie*. Vol. 1, 1–19. Berlin-New York: Walter de Gruyter.
- Kadioğlu, İdris. 2012. Lebib Divanı'nda; Musiki Terimleri ve Bir "Gazel-i Ferah-Sâz". *Turkish Studies* 7(3). 1575–92.
- Kalender, Ruhi. 1981. XV. Yüzyılda Arapça Musiki Terimleri ve Türkçe Karşılıkları. *Ankara Üniversitesi İlahiyat Fakültesi Dergisi* 24(1). 485–489.
- Kieffer, Jean Daniel & Thomas Xavier Bianchi. 1835. *Dictionnaire turc-français*. Paris: Impr. royale.
- Komsuoğlu, Ayşegül & Namık Sinan Turan. 2007. From Empire to the Republic: The Western Music Tradition and the Perception of Opera. *International Journal of Turcologia* 2(3). 5–29.
- Korabinsky, Johann Matthias. 1788. *Versuch eines kleinen türkischen Wörterbuchs mit beygesetzten Deutsch-Ungarisch und böhmischen Bedeutungen, und einer kurzgefassten türkischen Sprachlehre*. Pressburg: Verlage des Verfassers. <https://www.digitalniknihovna.cz/mzk/view/uuid:2912c65a-ef72-4eea-99c2-abc06295dbd8?page=uuid:e0746b2c-c64e-43dc-a732-28afc612ba2a> (17 April, 2023).
- Kubbealti Lugati. 2023. <http://lugatim.com/> (15 March, 2023).
- Lewis, Geoffrey. 1991. *Turkish Grammar*. Oxford-New York: Oxford University Press. 303 s.
- Lewis, Geoffrey. 1999. *The Turkish Language Reform: A Catastrophic Success*. Oxford: Oxford University Press.

- Ligara, Bronisława. 2017. Terminologia specjalistyczna i kultura: Dychotomia czy punkty wspólne? In Renata Przybylska & Władysław Śliwiński (eds.), *Terminologia specjalistyczna w teorii i praktyce językoznawców słowiańskich*, 29–52. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego. <https://ruj.uj.edu.pl/xmlui/handle/item/47749> (30 January, 2021).
- Lukszyn, Jerzy & Wanda Zmarzer. 2006. *Teoretyczne podstawy terminologii*. 2nd edn. Warszawa: Katedra Języków Specjalistycznych Uniwersytetu Warszawskiego.
- Meniński, Franciscus à Mesgnien. 1680. *Thesaurus Linguarum Orientalium, Turcicae, Arabicae, Persicae...* Vienna: na.
- Miodunka, Władysław. 1989. *Podstawy leksykologii i leksykografii*. Warszawa: Polskie Wydawnictwo Naukowe.
- Ölmez, Mehmet. 1994. Türk Dillerinin Sözlükleri ve Türk Sözlükçülüğü. In *Uygulamalı Dilbilim Açısından Türkçenin Görünümü*, 88–100. Ankara: Dil Derneği Yayınları.
- Özalp, Mehmet Nazmi. 2000. *Türk Müsiki Tarihi*, vol. 2. İstanbul: Milli Eğitim Bakanlığı.
- Özcan, Nuri. 2020. Muzika-yi Hümâyün. In Tahsin Yazıcı (ed.), *TDV İslâm Ansiklopedisi*, vol. 31, 420–422. Ankara: Türkiye Diyanet Vakfı. <https://islamansiklopedisi.org.tr/muzika-yi-humayun>.
- Öztuna, Yılmaz. 1969. *Türk musikisi ansiklopedisi*, vol. 1. İstanbul: Milli Eğitim Basımevi.
- Parlatır, İsmail (ed.). 2017. *Osmanlı Türkçesi Sözlüğü*. 9th edn. Ankara: Yargı Yayınevi.
- Pawlina, Agata. 2014. Muzyka klasyczna europy w imperium osmańskim. *Przełqad Orientalistyczny* 1-2. 61–76.
- Pawlina, Agata. 2021. *Od sultańskiej mehterhane do filharmonii narodowej. Terminologia muzyczna w leksykografii osmańsko-tureckiej i tureckiej*. Kraków: Jagiellonian University. (Doctoral dissertation).
- Prokosch, Erich. 2011. Osmanische Lexikographie. Erster Teil. *Wiener Zeitschrift für die Kunde des Morgenlandes* 101. 367–406. [www.jstor.org/stable/23861926](http://www.jstor.org/stable/23861926) (22 May, 2020).
- Prokosch, Erich. 2013. Osmanische Lexikographie. Zweiter Teil. *Wiener Zeitschrift für die Kunde des Morgenlandes* 103. 299–335. [www.jstor.org/stable/23862204](http://www.jstor.org/stable/23862204) (22 May, 2020).
- Redhouse, James W. 1861. *A Lexicon, English and Turkish: Shewing in Turkish the Literal, Incidental, Figurative, Colloquial and Technical Significations of the English Terms*. London: Oriental Literature Society.

- Reinhard, Kurt & Ursula Reinhard. 1984. *Musik der Türkei. 1 - Die Kunstmusik*. Wilhelmshaven: Heinrichshofen.
- Reinhard, Kurt, Martin Stokes & Ursula Reinhard. 2001. *Turkey*. Stanley Sadie (ed.). 2nd edn. New York: Grove. 909–921.
- Rocchi, Luciano. 2011. Turkish Historical Lexicography. *Lexicographica* 27(2011). 195–220. DOI: 10.1515/9783110236484.195. (2 September, 2020).
- Sager, Juan C. 1989. The dictionary as an aid in terminology. In Franz Josef Hausmann, Oskar Reichmann, Herbert Ernst Wiegand & Ladislav Zgusta (eds.), *Wörterbücher: Ein internationales Handbuch zur Lexikographie*, vol. 1, 167–171. Berlin-New York: Walter de Gruyter.
- Sâmî, Şemseddîn (Ch. Samy-Bey Frascbery). 1885. *Dictionnaire turc-français*. Constantinople: Imprimerie Mihran.
- Sanal, Haydar. 1964. *Mehter Musikisi: Bestekâr Mehterler, Mehter havaları*. İstanbul: Millî Eğitim Basımevi.
- Schweickard, Wolfgang. 2014. Türkische Wortgeschichte im Spiegel europäischer Quellen. *Zeitschrift für romanische Philologie* 130(3). 815–832. DOI: 10.1515/zrp-2014-0063.
- Siemieniec-Gołaś, Ewa. 2005. *Turkish Lexical Content in Dittionario della lingua italiana, turchesca by Giovanni Molino (1641)*. Kraków: Księgarnia Akademicka.
- Siemieniec-Gołaś, Ewa. 2015. Some remarks on Turkish Dictionaries Published in Constantinople/Istanbul Before and Soon After Language Reform in Turkey (1928). *Rocznik Orientalistyczny* LXVIII(2). 134–142.
- Stachowski, Marek. 2019. *Kurzgefaßtes etymologisches Wörterbuch der türkischen Sprache*. Kraków: Księgarnia Akademicka.
- Stachowski, Stanisław. 2002. *Lexique turc dans le Vocabulaire de P. F. Viguier (1790)*. Kraków: Księgarnia Akademicka.
- Stachowski, Stanisław. 1995. Studien zur türkischen Wortgeschichte (1). In *Studia Turcologica Cracoviensia (STC)*, vol. 1, 185–209. Kraków: Jagiellonian University Press.
- Stein, Heidi. 1990. Lexikographie. In György Hazai (ed.), *Handbuch der türkischen Sprachwissenschaft*, vol. 1, 335–370. Wiesbaden: Otto Harrassowitz.
- Tietze, Andreas. 2016. *Tarihî ve Etimolojik Türkiye Türkçesi Lugati*. Semih Tezcan (ed.). Ankara: Türkiye Bilimler Akademisi.
- Tinghir, Anton B. & Krikor Sinapian. 1891. *Fransızcadan Türkçeye işlâhât lugati. Dictionnaire français-turc des termes techniques des sciences, des lettres et des arts*. Constantinople: Impr. de K. Bagdadlian.
- Tohumcu, Ahmed. 2009. Türk Müziği Terminolojisinde Yozlaşma/Örnek Olay Analizi: Şarkı Formu. In 38. ICANAS (Uluslararası Asya ve Kuzey Afrika Çalışmaları Kongresi) 10-15.09.2007, Ankara, Bildirler: 13. Müzik Kültürü ve Eğitimi,

- 711–18. Ankara: Atatürk Kültür, Dil ve Tarih Yüksek Kurumu. <http://www.ayk.gov.tr/baglantilar/icanas-38/>.
- Turan, Namık Sinan. 2004. Yüzyılda Osmanlı İmparatorluğu'nda Batılılaşma ve Müzik. *Bilgi ve Bellek* 1(1). 87–104.
- Turan, Namık Sinan. 2011a. Opera Tutkunu Bir Sultan ii. Abdülhamid. *Evrensel Kültür* (1). 27–31.
- Turan, Namık Sinan. 2011b. Osmanlı'nın Son Döneminde Opera'. *Evrensel Kültür* (240). 22–26.
- Redhouse Sözlüğü. 2017. *Türkçe/Osmanlıca-İngilizce Redhouse Sözlüğü: Turkish/Ottoman-English Redhouse Dictionary*. 21st edn. İstanbul: SEV Yayıncılık Eğitim ve Ticaret A.Ş.
- Uz, İsmail Kâzım. 1892. *Ta'lim-i Musiki, yahut, Musiki Istılahatı*. Constantinople: Matbaa-yı Ebüzzıya.
- Wright, Owen. 1992. *Words without Songs: A Musicological Study of an Early Ottoman Anthology and Its Precursors*. London: School of Oriental & African Studies, University of London.
- Yalçın, Gökhan. 2013. Türkiye'de 1900-2012 Yılları Arasında Yayımlanan Batı Müzik Teorisi Kitaplarındaki Terminolojinin İncelenmesi. *İdil Sanat ve Dil Dergisi* 8(2). 65–91. DOI: 10.7816/idil-02-08-05.
- Yöre, Seyit. 2011. Kültürleşmenin Bir Parçası Olarak Osmanlı'da Operanın Görünümü. *Zeitschrift für die Welt der Türken / Journal of World of Turks* 3(2). 53–69.



# Chapter 14

## Exploring the unique method for encoding sinograms in the first known Chinese–Polish dictionary

Andrzej Swoboda

Jagiellonian University in Krakow

The aim of this chapter is to examine Doman Wieluch's Chinese character encoding method, introduced in his *Słownik chińsko-polski do czytania tekstów chińskich bez przygotowania* (Wieluch 1936) – the oldest (known) Chinese-Polish dictionary in history. In the first part of the chapter, the reference work in question is briefly described and discussed in the context of selected earlier lexicographic and linguistic publications. The macrostructure and microstructure of Wieluch's dictionary are outlined and compared to those of selected other monolingual and bilingual dictionaries from various periods. The second part of the chapter consists in a critical analysis of the author's encoding method: its general premise, specific rules (for isolating character components, replacing them with code words and arranging code words in sequences), selected examples of encoded headwords, as well as the peculiarities and limitations of the system. In the final part, Wieluch's ideas are discussed in the context of selected formerly and currently utilised methods for ordering, decomposing, encoding and inputting Chinese characters. It is hoped that this chapter will shed new light on this relatively unknown episode in the history of Chinese-Polish lexicography, which may be of interest to future dictionary authors.

### 1 Wieluch's Chinese-Polish dictionary

#### 1.1 Title(s) and author(s)

A relatively small number of Polish-Chinese and Chinese-Polish dictionaries have been published so far (cf. Sroka-Grądział 2018: 32). The earliest (surviving)



reference work in the above-mentioned language pair is *Słownik chińsko-polski do czytania tekstów chińskich bez przygotowania* ('A Chinese-Polish Dictionary for Reading Chinese Texts Without Preparation'), also known by its Chinese title 切字中波辭典 (*Qiē Zì Zhōng-Bō Cídiǎn*<sup>1</sup> 'A Chinese-Polish Character-Splitting Dictionary'). It was compiled by Doman Wieluch, a Polish chemist, Sinologist and diplomat, who wrote the entries, the preface and the indexes to the dictionary. While Wieluch is credited as the sole author, additional contributions to his publication were made by three other individuals (Wieluch 1936: 15<sub>p</sub>)<sup>2</sup>. Yu Houo-Joei (虞和瑞 *Yú Héruì*), a Chinese diplomat and Polonist, helped to create the visual layout of the work and calligraphed the Chinese character components in the appendices (*Table I* and *Table II*). Witold Jabłoński, a renowned Polish Sinologist working at the University of Warsaw, provided Wieluch with source materials and advice. Konstanty Matuszek, Wieluch's colleague, was thanked by the author "for a great deal of painstaking and intelligent work put into the book"<sup>3</sup>.

## 1.2 Characteristics and reception

*Słownik chińsko-polski...*, which contains approximately 6,000 entries on its 338 pages, was published by Wieluch himself in 1936 in Katowice. The reference work does not include a particularly large number of articles, which seems understandable, as the author himself perceived it as "merely the first step" for Chinese-Polish lexicography (Wieluch 1936: 14<sub>p</sub>) and explained that he did not have the financial means to publish a more comprehensive edition (Wieluch 1936: 3<sub>p</sub>). *Słownik...* remained the only bilingual general dictionary of its kind for several decades, but it was hardly ever used by Sinologists and translators and never received widespread acclaim. Its unpopularity may be demonstrated by the fact that it is actually absent from some of the publications which aim to give an account of the history of Polish-Chinese lexicography (cf. Sroka-Grądziel 2018). To

---

<sup>1</sup>Unless otherwise stated, the pronunciation of Chinese words was indicated using the Hanyu Pinyin Romanisation system (漢語拼音 *Hànyǔ Pīnyīn*) throughout the chapter and all the spellings is given in traditional characters (繁體字 *fántǐzì* or 正體字 *zhèngtǐzì*). However, other transcriptions and simplified sinograms are preserved if they have been originally used in names or titles of publications.

<sup>2</sup>The three main sections of Wieluch's dictionary (excluding the *Tables I & II*) are numbered separately using Hindu-Arabic numerals, with no further distinctions. For the sake of clarity, the following letters were added to the relevant page numbers included in the references throughout this chapter: <sub>p</sub> (*Preface and Explanation*), <sub>d</sub> (*Dictionary*) and <sub>i</sub> (*Indexes*; see also section 1.4).

<sup>3</sup>Original quote: "za masę mrówczej i inteligentnej pracy, włożonej w niniejszą książkę" (Wieluch 1936: 15<sub>p</sub>).

some extent, the limited success of Wieluch’s dictionary might be attributed to the lack of interest in Chinese-themed books among the general public or the relatively small print run of the work. In fact, only 200 numbered copies were originally created and some of the surviving volumes are incomplete.<sup>4</sup> On the other hand, it may also be suspected that one of the reasons for the relative obscurity of this reference work is the author’s innovative and controversial approach to recording Chinese lemmata. Unlike most other lexicographers, Wieluch opted to encode the spelling of Chinese headwords in all of his dictionary entries instead of printing the relevant sinograms or writing them by hand, citing the complexity and inaccessibility of the Chinese script as the reason for this decision. The author explained his original character encoding system in the preface to the dictionary and promoted the system as a reliable and simple method of transcribing and printing Chinese texts, especially for foreign readers (Wieluch 1936: 1-3<sub>p</sub>). However, Wieluch’s encoding system was never commonly accepted and did not appear in later bilingual dictionaries (see section 3.1).

### 1.3 Sources and inspirations

According to Wieluch, *Słownik chińsko-polski...* was compiled using the most recent editions of six previously published reference works (although no specific dates of publication were given). Four of these books were monolingual Chinese dictionaries: 說文解字 (*Shuōwén Jiězì*) by Xu Shen (許慎), 康熙字典 (*Kāngxī Zìdiǎn* ‘[the] *Kangxi Dictionary*’) by Zhang Yushu (張玉書), Chen Tingjing 陳廷敬 et al., 攷正字彙 (*Kǎozhèng Zìhuì*) by Chen Haozi (陳淏子) and 通用新字彙 (*Tōngyòng Xīn Zìhuì*, author uncertain). In addition, Wieluch claimed to have consulted two bilingual dictionaries: one of them was the Chinese-English 漢英大辭典 (*Hàn-Yīng Dà Cídiǎn*) by O. Z. Tsang, also known as 張鵬雲 (*Zhāng Péngyún*), whereas the other was the Chinese-Russian 新中俄大字典 (*Xīn Zhōng-Ē Dà Zìdiǎn*) by 路大和 (*Lù Dàhé*); (Wieluch 1936: 14<sub>p</sub>).

Selected works of certain Western scholars, such as Angelo Zottoli, Séraphin Couvreur, Herbert Allen Giles and William Edward Soothill were reportedly also consulted by Wieluch while compiling his reference work, although he did not mention any specific titles. The author claimed to have relied on those publications in order to establish and describe the meaning of the Chinese lemmata

---

<sup>4</sup>For instance, during the preliminary research for this chapter it was revealed that the appendices (*Table I and Table II* – see section 1.4) were missing from the copy of *Słownik...* stored at the Jagiellonian Library in Kraków, which made that particular volume significantly more difficult to use.

included in his book, albeit to a limited extent (as his main sources were the aforementioned monolingual and bilingual Chinese dictionaries). He stressed that he had not limited himself to simply copying other Western authors and pointed out that the contents of *Słownik...* were in fact a result of his own deliberations (Wieluch 1936: 11<sub>p</sub>).

Wieluch's character encoding system was reportedly partially inspired by the earlier publications of John Chalmers and Léon Wieger (Wieluch 1936: 1<sub>p</sub>). Although, once again, no specific titles were mentioned by the author, it may be suspected that the works consulted by Wieluch were *An Account of the Structure of Chinese Characters Under 300 Primary Forms After the Shwoh-Wan, 100, A. D., and The Phonetic Shwoh-Wan, 1833* (Chalmers 1882), and *Chinese Characters: Their Origin, Etymology, History, Classification and Signification. A Thorough Study from Chinese Documents* (Wieger 1965, originally published in French in 1915).

#### 1.4 Contents

The first part of Wieluch's dictionary, *Wstęp i wyjaśnienie* ('*Preface & Explanation*') contains a general introduction to the reference work and to the author's original character encoding method. The latter was not only theoretically presented, but also practically demonstrated by including a sample Chinese text on page 12<sub>p</sub> and its encoded version on page 13<sub>p</sub>. The second part, called *Słownik* ('*Dictionary*'), is a single Chinese-Polish wordlist, which constitutes the main body of the publication. Its entries were arranged alphabetically based on the encoded spellings of the Chinese lemmata. The wordlist is followed by four indexes: the chronologically arranged index of Chinese dynasties and sovereigns, as well as three other indexes arranged alphabetically by character encodings: the index of Chinese era names, the index of Chinese provinces, prefectures, cities and towns, and the index of reversed character encodings. In addition, *Słownik polsko-chiński...* contains two appendices, printed on separate pieces of paper: *Table I* (an index of character components arranged alphabetically by code words) and *Table II* (an index of character components arranged by the number of brush strokes).

Most of the above-mentioned elements, such as prefaces, wordlists or indexes (including those containing typically extralinguistic information on history, geography or other related topics) appear quite frequently in various Chinese-Polish dictionaries (cf. Xu Lin (许琳), Yao Xishuang (姚喜双), et al. 2010, Kocyba-Grych & KołECKA 2012, etc.). On the other hand, indexes of reversed character encodings are exceedingly rare (if not completely unheard-of outside of Wieluch's

publication). According to the author, this special index was designed for deciphering unclear sinograms. Wieluch perceived it as a new, groundbreaking tool, which could enable future Sinologists (and other users) to make sense of partly erased or misprinted glyphs in the analysed Chinese texts (Wieluch 1936: 14<sub>p</sub>).

### 1.5 Macrostructure and microstructure

The entries in Wieluch's dictionary were arranged alphabetically by character encodings. This specific type of macrostructure is highly unusual for both monolingual and bilingual Chinese dictionaries. However, it does share certain characteristics with the traditional method of arranging Chinese-language lemmata used in *Shuowen Jiezi*, the *Kangxi Dictionary* and numerous other monolingual reference works (see 2.2, 3.1). In both of these types of macrostructure the order of entries is essentially determined by the structure of sinograms which appear in the spelling of the relevant headwords.

On the other hand, it might also be argued that the macrostructure of *Słownik...* is somewhat similar to the more contemporary and commonly used alphabetic order of entries based on the Hanyu Pinyin Romanisation system (cf. Xia Zhengnong (夏征农) & Chen Zhili (陈至立) 2009, Xu Lin (许琳), Yao Xishuang (姚喜双), et al. 2010, Kocyba-Grych & KołECKA 2012, Achinger & Radziwiłł 2015, etc.). In both of these types of macrostructure the order of headwords on the wordlist is determined by the conventional order of letters in the Latin alphabet, even if the methods of assigning sequences of letters to the relevant Chinese lemmata are anything but the same.

Either way, the macrostructure of *Słownik...* does appear radically dissimilar from the thematic order of entries which can be found in certain other Chinese-Polish dictionaries (cf. Wu Yuemei (吴月梅) 2010: Yin Xiangfeng (殷相峰) 2013: etc.).

Each entry in Wieluch's dictionary contains the encoding of one or more characters representing the relevant Chinese headword. In addition, every article includes one or more Polish equivalents of the lemma, one or more of its definitions in Polish or both of the aforementioned elements. Furthermore, each entry contains the information on the part (or parts) of speech to which the headword in question belongs. According to the author, this simple grammatical information (expressed by means of the commonly used Latin abbreviations, such as *n.*, *v.* or *adj.*) had been gleaned from Tsang's Chinese-English dictionary (汉英大辞典 *Hàn-Yīng Dà Cídiǎn* – cf. 1.3). As Wieluch himself pointed out, these part-of-speech symbols may occasionally seem to be at odds with the Polish equivalents or definitions provided in the entry, since – for the sake of conciseness – only

selected meanings of certain polysemic and multifunctional headwords had been considered (Wieluch 1936: 11<sub>p</sub>).

The pronunciation transcription of each headword, written in upper case, was indicated using a modified version of the Wade-Giles Romanisation. According to Wieluch, this particular transliteration system had been adapted to the spelling habits of Polish readers, mostly following prof. Jabłoński's suggestions (Wieluch 1936: 11<sub>p</sub>). The two most noticeable modifications were: the use of doubled letters instead of single letters with apostrophes<sup>5</sup> to denote aspirated consonants and the use of digraphs inspired by Polish orthography (with each letter doubled separately to represent an aspirated sound). For example, the pronunciation of the word 財 (*cái* 'wealth/property'), which typically would be rendered as *ts'ai*<sup>2</sup> in the Wade-Giles Romanisation system, was actually written as *TTSSAI* in *Słownik...* Moreover, as can be seen in this example, the syllable tones were not indicated in any way in the transcription used in Wieluch's dictionary, as a result of which the information on pronunciation is incomplete for any given headword. Although it may be surprising, the above-mentioned imperfections appear to have been intended (or at least ignored) by the author. In fact, Wieluch admitted to "have treated the phonetic aspect of [his] dictionary with reprehensible negligence, as [he] wanted to emphasise that understanding Chinese texts [did] not depend on [knowing] their Chinese pronunciation."<sup>6</sup>

The phonetic transcription included in the entries may also serve a different purpose, as it sometimes provides additional information on the usage of certain headwords: the Romanisation is underlined if the lemma in question may function as a Chinese surname.

Some of the articles include other types of data, such as: information on antonyms, synonyms or other related words, information on spelling variants, full forms of abbreviated headwords, etc. Many compound words, collocations and fixed expressions appear in subentries, rather than as headwords in separate entries on the wordlist. In each of these subentries, the headword character encoding was replaced by the repetition symbol (*./.* or *././.*, depending on the number of repeated characters). For selected examples of microstructural elements in the entries of Wieluch's dictionary, see Figure 1.<sup>7</sup>

<sup>5</sup>Although the diacritic in question is often referred to as an *apostrophe* (i.e. ' ; cf. (Wieluch 1936: 11<sub>p</sub>), it may actually be the rough breathing character (i.e. ' or ' , also known as the *spiritus asper*), as originally used by Thomas Wade.

<sup>6</sup>Original quote: "Stronę dźwiękową naszego słownika traktowałem z karygodnem [*sic*] niedbalstwem, chcąc uwydatnić niezależność rozumienia tekstów chińskich od ich chińskiego dźwięku." (Wieluch 1936: 11<sub>p</sub>).

<sup>7</sup>All figures included in this chapter were created using the photographs of the copy number 65 of *Słownik...*, which is currently owned by the Józef Lompa Voivodeship Pedagogical Library in Katowice (Pedagogiczna Biblioteka Wojewódzka im. Józefa Lompy w Katowicach).

In certain respects, the microstructure of *Słownik...* is quite similar to that of many later Chinese–Polish dictionaries: the entries vary in length, but are usually relatively short; they contain some basic information on the key aspects of the headwords in question, such as their spelling, meaning, pronunciation or grammatical category (cf. Wu Yuemei (吴月梅) 2010, Xu Lin (许琳), Yao Xishuang (姚喜双), et al. 2010, Kocyba-Grych & Kofecka 2012, Yin Xiangfeng (殷相峰) 2013, etc.). However, the entries in Wieluch’s dictionary are somewhat unique because of certain design choices made by the author. His novel Romanisation system is now completely obsolete and Hanyu Pinyin is almost invariably used instead; however, it should be noted that the latter system did not exist at the time of *Słownik...*’s creation and the idea of adopting a universal, global Chinese Romanisation system for speakers of all languages was less commonly accepted. The other unprecedented element of Wieluch’s dictionary articles – the character encodings – will be more thoroughly discussed in the next part of this chapter.

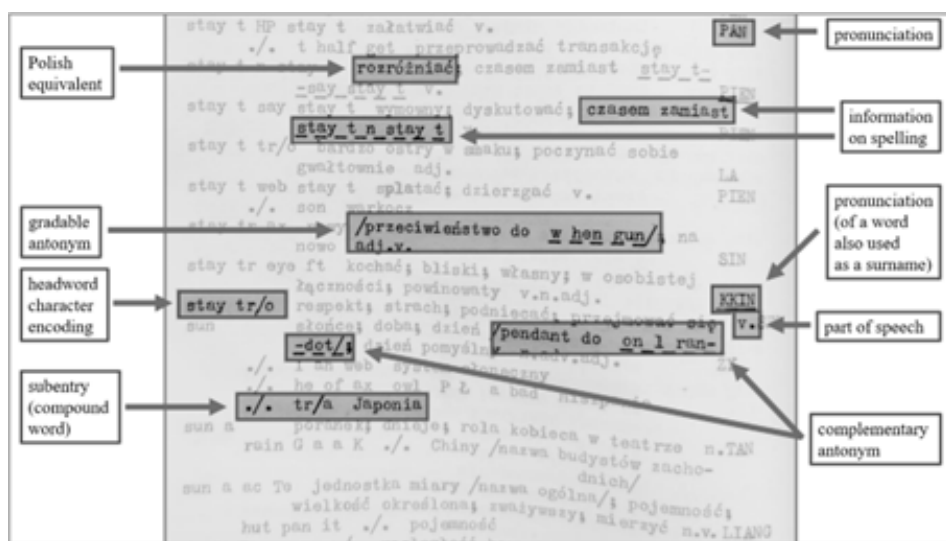


Figure 1: The microstructure of “*Słownik...*” (page 204<sub>d</sub>) – selected types of data (CC BY-NC-ND Doman Wieluch)

## 2 Wieluch’s character encoding system

### 2.1 Introduction

As mentioned earlier, sinograms were used quite sparingly in Wieluch’s dictionary: they only appear in the appendices, on the front cover and on page 12

(which contains a sample text in the regular form, next to its encoded version on page 13). All other Chinese characters were replaced throughout the dictionary by the author's original character encodings.

Every sinogram in Wieluch's system was essentially represented as a set of its parts, i.e. as a set of (*Chinese*) *character components*. Each of these character components was replaced by one of the 296 *code words*, which were supposed to be as simple and short as possible (Wieluch 1936: 1<sub>p</sub>). Every code word was either a specific letter of the Latin alphabet or a specific sequence thereof. While encoding entire Chinese characters, such code words were written in a particular order to represent the structure of the relevant sinograms. When two or more characters were being encoded, the corresponding sets of code words were written in rows, from left to right, top to bottom, so as to reflect the entire original text in a manner familiar to Western readers.

## 2.2 Selecting code words

The original criteria for selecting code words were explained by Wieluch in the preface to his dictionary (Wieluch 1936: 4-9<sub>p</sub>). A revised, simplified version of these criteria, based on the author's explanations and a new typological analysis of the character components used in his encoding system, will be presented in this part of the chapter. While the discrepancies between this new classification and the one originally included in the preface to Wieluch's dictionary are not particularly significant, they will be briefly discussed later on (see 2.3).

1. Some of the code words in Wieluch's system were chosen because of the similarity in shape between the relevant character components and certain letters of the Latin alphabet or groups thereof. For instance, the component 尸 was replaced by the code word *P*, the component 十 was substituted by the code word *t* and the component 匚 was paired with the code word *Ls* (Wieluch 1936: 4<sub>p</sub>, *Tables I & II*).
2. Some other code words were chosen because of the similarity in shape between the relevant character components and certain letters of the Greek alphabet. For instance, the component 厶 was assigned the code word *D*, as it resembles the Greek letter Δ; the component 冫 was replaced by the code word *G*, since it resembles the Greek letter Γ; the component 彡 was paired with the code word *s*, as it resembles the Greek letter Σ (Wieluch 1936: 4<sub>p</sub>, *Tables I & II*).

3. A number of other code words were chosen because of the similarity in shape between the relevant character components and the real-world objects denoted by the code words in question. For instance, the component 丶 was assigned the code word **dot**, the component 亠 was replaced by the code word **hill** and the component 凵 was paired with the code word **owl** (as it vaguely resembles the outline of a perched owl, (Wieluch 1936: 4<sub>p</sub>, *Tables I & II*).
4. Some other code words were chosen because of the similarity in shape between the relevant character components and some *elements* of certain real-world objects (or sets thereof). For instance, the component 冂 was assigned the code word **cot** [*sic* – see 2.3], as it vaguely resembles the outline of the main part of a coat (with missing sleeves); similarly, the component 卅 was replaced by the code word **pan**, as it is somewhat similar in shape to a pan hanging from a spit (with the bottom part of the pan missing; (Wieluch 1936: 4-5<sub>p</sub>, *Tables I & II*)).
5. Some of the code words were chosen because they denote the same objects which are referred to by the relevant character components (when these components are used as individual sinograms representing specific Chinese words). For instance, the component 身 (*shēn* ‘body’) was assigned the code word **body**; the component 犬 (*quǎn* ‘dog’) was replaced by the code word **dog**; the component 雨 (*yǔ* ‘rain’) was substituted by the code word **rain** and the component 羽 (*yǔ* ‘feather/wing’) was paired with the code word **wing** (Wieluch 1936: 5-7<sub>p</sub>, *Tables I & II*).
6. In certain cases a similar rule was followed, but the meaning of the code word in question was only close – and not identical – to that of the relevant Chinese character component. For instance, the component 虫 (*chóng* ‘insect/worm’) was assigned the code word **bee**; the component 鬼 (*guǐ* ‘ghost/demon’) was paired with the code word **elf**; the component 女 (*nǚ* ‘woman’) was replaced by the code word **her**, whereas the component 束 (*shù* ‘bunch, bouquet’) was substituted by the code word **rose** (Wieluch 1936: 7-8<sub>p</sub>, *Tables I & II*).
7. In some other cases a similar rule was applied, but the meaning of the code word was (almost) identical to that of a certain *more complex sinogram* containing the relevant Chinese character component. For instance, the component 卅 was assigned the code word **all**, as the component in question appears in the character 共 (*gòng* ‘common/total/altogether’); the

component 丁 was replaced by the code word **can**, as the component in question is part of the character 可 (*kě* ‘can, be able to’); the component 曲 was substituted by the code word **law**, as the component in question appears in the character 典 (*diǎn* ‘canon/law/ceremony’; Wieluch 1936: 8<sub>p</sub>, *Tables I & II*).

8. One of the code words, **hu**, was selected because of the similarity between its suggested pronunciation and the reading of the corresponding character component used as a separate sinogram: 互 *hù* (Wieluch 1936: 9<sub>p</sub>, *Tables I & II*).
9. The code word **a** was selected to represent the component 一 because of the perceived similarity between the traditional sequence of letters in the Latin alphabet and the conventional order of *Chinese character radicals* (i.e. principal character components, which often determine the order of articles in dictionaries – see section 3.1). While *a* is the first letter of the Latin alphabet, 一 (*yī*) is the first Chinese character radical (Wieluch 1936: 4<sub>p</sub>, *Tables I & II*).
10. Interestingly, Wieluch’s system is actually partially self-referential: some of the code words were selected because of the similarity in shape between the corresponding character components and certain other character components, to which certain other code words had already been assigned. For instance, as the component 一 had already been replaced by the code word **a**, and the component 尸 had already been replaced by the code word **P**, the component 尸 (which may be perceived as a combination of 一 and 尸) was substituted by the code word **a P**. Since the component 羊 had already been assigned the code word **ram**, the simpler component containing one fewer horizontal brush stroke at the bottom was paired with the code word **kid**. Similarly, as the component 車 had already been given the code word **car**, the component 𠂔 (whose bottom part is slightly different in shape) was replaced by the code word **van** (Wieluch 1936: 4<sub>p</sub>, 9<sub>p</sub>, *Tables I & II*).

Although typically only one of the above-mentioned rules was followed while choosing the code word for a specific character component, in some cases there may have actually been two applicable selection criteria. For example, the component 世 (*shì* ‘generation/world’) was paired with the code word **gen** because of the similarity in meaning between the words 世 and *gen[eration]*. Nevertheless, the same component was also presented by the author as a combination of three **t** (十) components put together. As he explained, such a description was

suggested by the above-mentioned similarity in shape, but also by the fact that a *generation* may be understood as a *period of thirty* (i.e. three times ten) *years* and 十 (*shí*) means ‘ten’ as a separate Chinese word (Wieluch 1936: 5<sub>p</sub>). Similarly, while the code word assigned to the previously mentioned component 一 (*a*) was selected because of the identical placement of the words *a* and 一 in monolingual dictionaries, Wieluch additionally cited the similarity in meaning between the English indefinite article *a* and the Chinese word 一 (*yī* ‘one’) as another reason for this pairing (Wieluch 1936: 4<sub>p</sub>).

While the above classification of code words within Wieluch’s system does give a general idea about the author’s intentions, it is by no means exhaustive. As Wieluch himself explained in the preface to his dictionary, some of the code words were in fact chosen “somewhat freely”, i.e. with hardly any regard for the above-mentioned criteria (Wieluch 1936: 9<sub>p</sub>). This issue, along with some other peculiar characteristics of the code word selection process within Wieluch’s system, will be elaborated on in the next part of the chapter.

### 2.3 Peculiarities of the code word selection

While considering the code word selection in Wieluch’s character encoding system, it might be useful to discuss some of the potentially controversial characteristics of this process. Although the task of dissecting and methodically recreating all Chinese characters is admittedly daunting and likely impossible to complete in a perfect way, it may be argued that some of the peculiarities of Wieluch’s code word selection process create certain practical limitations to his system and render it more challenging for the prospective users.

As previously indicated, some of the code words in Wieluch’s system were selected quite arbitrarily, without any explicitly presented reason, which the author himself admitted. For example, the component 又 was assigned the code word *it*, which seems in no way related to its shape, its reading (*yòu*) or its meaning as a separate sinogram (‘again/also’). Similarly, the component 甫 (*fǔ* ‘just, barely/[one’s courtesy name]’) was paired with the code word *my* (Wieluch 1936: 9<sub>p</sub>).

In some other cases there actually was a specific reason for choosing a certain code word, but it was not particularly obvious and did not clearly correspond to any of the previously established selection criteria. For instance, the component 齊 (*qí* ‘neat/simultaneous/[name]) was assigned the code word *Poland* (i.e. the name of Wieluch’s home country), which the author himself explained with this somewhat fatalistic remark: “[齊] represents grain on a field, it denotes a fertile plain, and also a certain ancient state – for this patriotic gesture I am going

to have difficulties with my system in international education[/academia] and Polish Sinology is not going to support me [on this].<sup>8</sup> In other words, in certain cases the previously delineated criteria for selecting code words were treated rather as general guidelines than as strict, unbreakable rules.

As may be understood from the quotation included in the previous paragraph, Wieluch hoped that his method for encoding sinograms could one day be used all over the world. To that end, he set out to create a truly international system for encoding Chinese characters, which would be useful not only to his compatriots, but also to speakers of various other languages. In order to accomplish that goal, he chose English as the source language for the majority of his code words, as evidenced by the examples included in the previous parts of this chapter; Wieluch 1936: 1-2<sub>p</sub>.

While Wieluch mostly abided by the above-mentioned rule, it should be noted that exceptions were made in certain cases. Some of the code words were actually derived from certain French or Latin lexemes and not their English counterparts. For instance, the French code words *bon* ('good') and *pas* ('step') were used instead of their English equivalents, as were the Latin code words *cor* ('heart'), *et* ('and') and *mus* ('mouse'; Wieluch 1936: 5-6<sub>p</sub>, 8<sub>p</sub>).<sup>9</sup> Although it might be argued that both French and Latin enjoyed the status of international languages at a certain point in history (or even that they continue to do so now), the use of such code words may have introduced additional (and possibly unnecessary) complexity to Wieluch's character encoding system.

Furthermore, it should be noted that some of the aforementioned code words derived from French or Latin are quite similar to certain other code words derived from English, even though they obviously do not correspond to the same character components. For instance, while the code word *et* ('and') was selected for the component 而 (*ér* 'and/and so/yet'), the code word *and* was chosen for the component 且 (*qiě* 'and/moreover/yet'). Selecting each of these code words appears justified at first glance, as each of them (rather accurately) reflects the meaning of the relevant component as a separate Chinese character. On the other hand, it might be argued that the inclusion of both *et* ('and') and *and* as code words could be confusing to the potential users of the system, given that these

---

<sup>8</sup>Original quote: "[齊] wyobraża łąn zboża, oznacza żyzną równinę, a także pewne starożytne państwo – za ten gest patryjotyczny [sic] będę miał trudności z moim systemem w nauce międzynarodowej, a polska sinologia [sic] mnie nie poprze" (Wieluch 1936: 8<sub>p</sub>).

<sup>9</sup>The code words *cor*, *et* and *mus* can be identified as Latin (and not French or English) based on the comments and translations provided by Wieluch; for the same reason, it is certain that the code word *pas* is undoubtedly of French origin (Wieluch 1936: 5-6<sub>p</sub>, 8<sub>p</sub>).

words are identical in meaning in their respective source languages (Wieluch 1936: *Tables I & II*).

Another possible issue for the prospective foreign users of Wieluch's encoding system is the inclusion of the Polish letter *Ł* (also known as *L with stroke*, lower-case form: *ł*; Wieluch 1936: *Tables I & II*). As the letter in question only exists in the spelling systems of about a dozen languages (most notably Polish, but also Kashubian or Navajo, among others) and was actually even less common at the time of the creation of Wieluch's dictionary, it could be argued that its presence in a supposedly universal encoding system may have rendered it slightly more difficult for many potential users.

Since the code words in Wieluch's system were meant to be short, it is perhaps not surprising that some of them were coined by abbreviating certain relatively long source-language words. For example, *HP* was used instead of *horsepower*, *Me* was chosen instead of *metal*, *rt* was selected instead of *right* and *sec* appeared instead of *second*. Similarly, the code word *M* was used instead of the French *Monsieur* ('Mister') and the code word *Te* appeared instead of the Latin *Tellus* ('Earth'); (Wieluch 1936: 4–8<sub>p</sub>).

For the sake of brevity, non-standard spellings were also used as code words in certain cases. *Coat* was respelled as *cot* (cf. section 2.2), *door* was rendered as *dor*, *hand* was replaced by *had* and *plough* was written as *log* (Wieluch 1936: 4–5 ;<sub>p</sub>). The analysis of these examples points to another possible problem with Wieluch's code word selection: while the above-mentioned words are intended to be understood as the unusual variant spellings of the words *coat*, *door*, *hand* and *plough*, they might be erroneously recognised as the regular spellings of the words *cot* ('a small or portable bed'), *dor* ('trick, joke/[a kind of insect]'), *had* ('[past simple or past participle form of *have*']') and *log* ('a thick piece of wood/a written record'), thus making the system more confusing.

While the majority of code words refer specifically to single character components (as the general rules would suggest), this is actually not always the case. To some extent, this multireferentiality of certain code words mirrors one of the characteristics of the Chinese script: the same character component may be written in a slightly different way (i.e. have a somewhat different shape) when it appears as a component in different complex sinograms. For instance, the character component 心 (*xīn* 'heart/mind/centre') is written as 心 in the character 想 (*xiǎng* 'to think/to want/to miss'), but as 忄 in the character 怕 (*pà* 'to fear, to dread') and as 小 in the character 恭 (*gōng* 'respectful, polite'); however, all of these variant forms are replaced by the same code word (*cor* 'heart') in Wieluch's encoding system (see Figure 2). Since both the origin and the meaning of all such variant forms are essentially identical, the author's decision to group all these forms

under the same code word appears justified, even if it might increase the complexity of his system. Nevertheless, in certain rare cases the decision to assign the same code word to more than one component seems to have been rather arbitrary, as no etymological reason may be cited to support it. For instance, while the code word *tow* corresponds to the mutually related components 申 and 电, it also refers to the similarly-shaped, but unrelated component 册 (Wieluch 1936: *Tables I & II*).

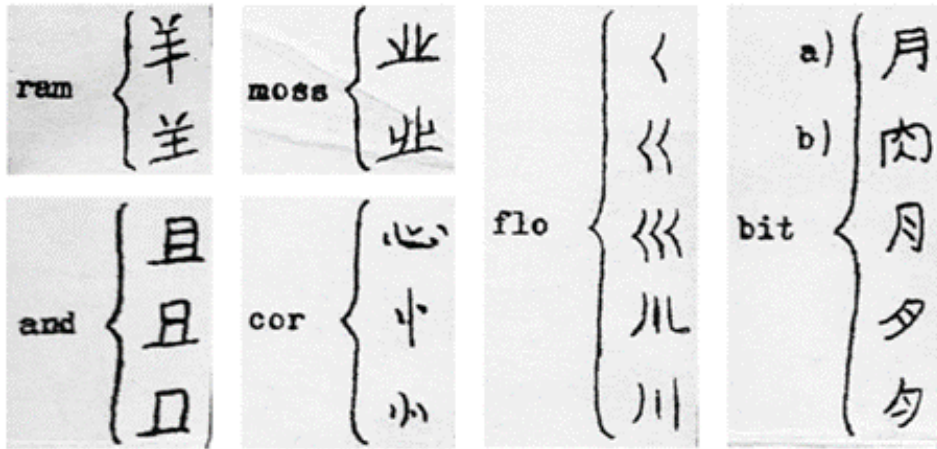


Figure 2: Multireferentiality of code words in Wieluch's system (CC BY-NC-ND Doman Wieluch)

Another possible issue with Wieluch's code word selection process is the fact that both upper-case and lower-case letters were used as single-letter code words. In certain cases the relevant corresponding elements are markedly different: for example, the code word *b* (derived from the English word *born*) corresponds to the component 生 (*shēng* 'to be born/to give birth') and the code word *B* (selected because of the similarity in shape between the letter and the relevant component) corresponds to the component 卩, which makes these two code words relatively easy to distinguish. On the other hand, some pairs of single-letter code words (containing the same letter in upper case and in lower case) were chosen solely because of the similarities in shape between the letters and the relevant components, which makes selecting and identifying such code words significantly more challenging. These difficulties are sometimes exacerbated by the multireferentiality of such single-letter code words and the similarities between the letters themselves. This issue is perhaps most noticeable when we consider the code words *I*, *l*, *L* and *l* (all of them being distinct and multireferential), which together correspond to a total of eleven different character components (see Figure 3; Wieluch

1936: *Tables I & II*). The code words **o** (口) and **O** (口) may be considered a similar, but somewhat particular example, as they are additionally differentiated by the combination properties of their corresponding components. While the **o** component never contains other components, the **O** component invariably does. These properties may be observed in the encoding of the character 固 (*gù* ‘solid, hard’), rendered as **O t o** (Wieluch 1936: 4<sub>p</sub>, 163<sub>d</sub>; see also 2.4).

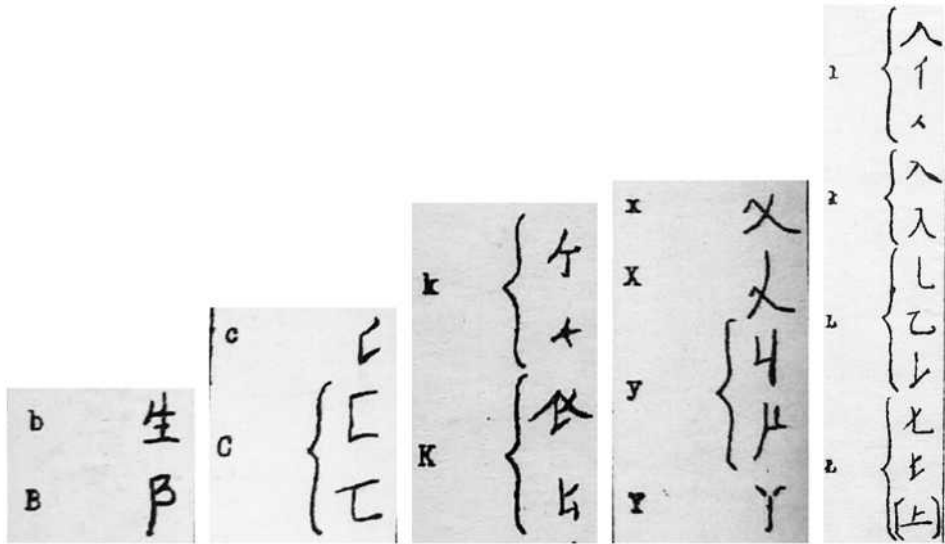


Figure 3: Upper-case and lower-case letters as code words in Wieluch's system (CC BY-NC-ND Doman Wieluch)

The above-mentioned multireferentiality of certain code words is additionally complicated by the fact that some character components may occur as variant forms which are identical to certain other components. As a result, some forms are associated with more than one code word (see Figure 4 – the code words associated with variant forms are given in square brackets). For example, while the character component 上 was primarily paired with the monoreferential code word **up**, it was also described as the third variant form associated with the code word **L** (cf. Figure 3. Wieluch 1936: *Tables I & II*).

Finally, as Wieluch himself admitted in the preface to his dictionary, his encoding system is actually not completely exhaustive. While the code words listed in his publication are perfectly sufficient to encode numerous sinograms, some Chinese character components were not assigned any specific code words of their own. Thus, if one of those components appears in a certain Chinese character (and cannot be broken down into two or more encodable components), it has

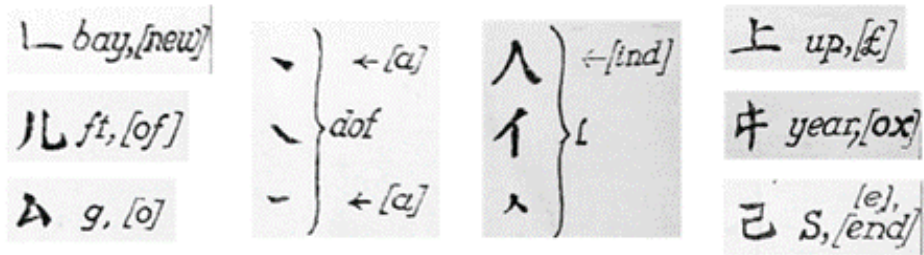


Figure 4: Multireferentiality of character components in Wieluch's system (CC BY-NC-ND Doman Wieluch)

to be encoded in the same way as we would normally encode *the most similar* component explicitly included in the system (Wieluch 1936: 9<sub>p</sub>).

As the various types of code words and the peculiarities of the process of code word selection in Wieluch's system have already been discussed, it might be useful to end this part of the chapter by briefly summarising the differences between the previously presented classification of code words and the one originally included in the preface to *Słownik...*:

1. In the original classification, the code word **ah**, which had been derived from the code word **a** because of the similarity in shape between the relevant corresponding components (see Figure 5), was not described as yet another example of a code word based on another, previously selected code word, but instead was listed in a separate category (Wieluch 1936: 4<sub>p</sub>).
2. In the original classification, the code words based on other, already selected code words, such as **a P** (戶), based on **a** (一) and **P** (尸) or **van** (庚), based on **car** (車), were divided into two different groups: combined components (such as **a P**) and modified components (such as **van**; Wieluch 1936: 4-9<sub>p</sub>).
3. In the original classification, there was a small separate group for code words based on mnemonics, even though these code words arguably could have been placed in other categories. One of these code words was **bud** (朱), selected because of the similarity in shape between its corresponding component and the component paired with the code word **tr** (木 'tree'; Wieluch 1936: 8-9<sub>p</sub>).
4. In the original classification, the arbitrarily assigned code words, such as **it** (又) or **my** (甫) were grouped in a separate category (Wieluch 1936: 9<sub>p</sub>).

5. In the original classification, some of the code words discussed in this chapter, such as *owl* (巴), belonged to different, but similar categories (Wieluch 1936: 4-9<sub>p</sub>).

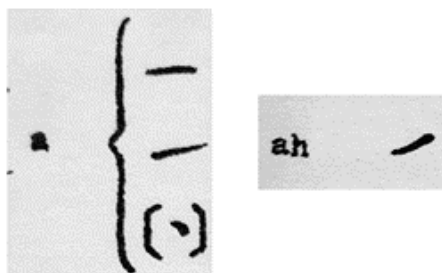


Figure 5: Components corresponding to the code words *a* and *ah* in Wieluch's system (CC BY-NC-ND Doman Wieluch)

## 2.4 Encoding entire sinograms

The process of replacing Chinese character components with code words, described in greater detail in the previous parts of the chapter, was a key element of Wieluch's encoding system. However, it should be noted that most sinograms are actually composed of two or more character components. For this reason, it was also necessary to devise a set of rules regarding the process of forming entire sequences of code words, which could represent more complex Chinese characters.

If a certain sinogram consisted of a single component, the relevant code word was simply written down instead of the character in question. For instance, the character 失 (*shī* 'to lose/to miss/to fail') was substituted by the code word *lose*, the character 山 (*shān* 'mountain') was replaced by the code word *Mt* and the character 言 (*yán* 'speech/to say') was substituted by the code word *say* (Wieluch 1936: 137<sub>d</sub>, 146<sub>d</sub>, 190<sub>d</sub>).

As indicated earlier, variant forms of the same character component were typically substituted by the same multireferential code word in Wieluch's system. While the author suggested the possibility of clearly identifying specific component forms occurring in complex sinograms by adding numerals to the relevant code words (Wieluch 1936: 9<sub>p</sub>), he decided against adopting this solution in his dictionary. However, if two variant forms of a certain component could exist as two different separate sinograms (representing two different words), the relevant

forms were additionally marked with the letters *a*) and *b*),<sup>10</sup> so as to avoid ambiguity. This solution was used for exactly two pairs of characters corresponding to two pairs of variant component forms. 月 (*yuè* ‘moon/month’) was assigned the sequence **bit a**), while 肉 (*ròu* ‘meat, flesh’) was replaced by the sequence **bit b**) (cf. Figure 2); 子 (*zǐ* ‘son/child’) was represented as **son a**), whereas 孑 (*jié* ‘lonely, alone’) was substituted by **son b**) (Wieluch 1936: 9<sub>p</sub>, 23<sub>d</sub>, 198<sub>d</sub>, *Tables I & II*).

If a certain sinogram included two or more character components, the consecutive code words corresponding to these components were written one after another with either single spaces or hyphens (-) between them. The specific order of these code words, reflecting the structure of the original Chinese glyph, was determined by the following rules:

1. If the character contained two or more vertical components placed next to one another, the relevant code words were written from the left-hand side to the right-hand side of the sinogram. For example, the character 袖 (*xiù* ‘sleeve’) was encoded as **robe by**, as it includes the component 衤 (corresponding to the code word **robe**) on the left-hand side and the component 由 (corresponding to the code word **by**) on the right-hand side. Similarly, the character 謝 (*xiè* ‘to thank/to wither/to decline’) was replaced by the sequence **say body cm**, as it includes the component 言 (corresponding to the code word **say**) on the left-hand side, the component 身 (corresponding to the code word **body**) in the centre and the component 寸 (corresponding to the code word **cm**) on the right-hand side (Wieluch 1936: 10<sub>p</sub>, 186<sub>d</sub>, 191<sub>d</sub>).
2. If the character contained two or more horizontal components placed next to one another, the relevant code words were written from the top part to the bottom part of the sinogram. For instance, the character 分 (*fēn* ‘to divide’) was encoded as **of n**, since it includes the component 八 (corresponding to the code word **of**) as its top part and the component 刀 (corresponding to the code word **n**) as its bottom part. Similarly, the character 夢 (*mèng* ‘dream/to dream’), which consists of the components 卜 (**w**), 𠃉

<sup>10</sup>In fact, these symbols were variously spelled as *a*) and *b*) (Wieluch 1936: *Tables I & II*), /*a*/ and /*b*/ (Wieluch 1936: 23<sub>d</sub>) or /*forma a*/ and /*forma b*/ (Wieluch 1936: 198<sub>d</sub>) in the dictionary. For the sake of clarity, only the first of these three pairs of symbols is consistently used in this chapter.

(*net*), 冫 (*m*) and 夕 (*eve*), was represented by the sequence **w net m eve** (Wieluch 1936: 10<sub>p</sub>, 165<sub>d</sub>, 238<sub>d</sub>).<sup>11</sup>

3. If the character contained two or more components arranged in a different manner, the relevant code words were written from the top left-hand corner to the bottom right-hand corner of the sinogram. For example, the character 但 (*dàn* ‘but/merely’) was encoded as **l sun a**, as it includes the component 亻 (*l*) on the left-hand side and the components 日 (*sun*) and 一 (*a*) on the right-hand side. The character 架 (*jià* ‘to support/frame’) was encoded as **HP o tr**, since it includes the components 力 (*HP*) and 口 (*o*) as its top part, and the component 木 (*tr*) as its bottom part. The character 能 (*néng* ‘can, to be able’) was encoded as **g bit L L**, as it includes the components 亠 (*g*) and 月 (*bit*) on the left-hand side and two 匕 (*L*) components on the right-hand side. The character 固 (*gù* ‘solid, hard’) was rendered as **O t o**, since it includes the components 十 (*t*) and 口 (*o*) placed inside the component 囗 (*O*; cf. 2.3; Wieluch 1936: 10<sub>p</sub>, 62<sub>d</sub>, 108<sub>d</sub>, 133<sub>d</sub>, 163<sub>d</sub>).

In some rare cases, applying the above-mentioned rules was insufficient to unequivocally represent a certain character, as the only difference between the glyph in question and a certain other sinogram was the placement of components, with identically encoded components placed on the left-hand and right-hand sides in one character, and in the top and bottom parts in the other. The encodings of such characters included additional symbols to indicate the vertical (||) or horizontal (=) relative placement of the relevant components. For instance, the character 叭 (*bā* ‘[sound of a sharp noise]’) was represented by the sequence **o of ||**, while the character 只 (*zhǐ* ‘only, merely’) was replaced by the sequence **o of =**. A similar solution was adopted to encode more complex sinograms, such as 暉 (*huī*, ‘sunshine/to shine upon’), paired with the sequence **sun m car |=** and 暈 (*yūn* ‘dizzy/to faint’ or *yùn* ‘halo/dizzy/to faint’), substituted by the sequence **sun m car ≡** (Wieluch 1936: 10<sub>p</sub>, 159<sub>d</sub>, 206<sub>d</sub>).

The forward slash symbol (/) was used between two code words in certain sequences to indicate that one character component was placed on top of another. For instance, the character 本 (*běn* ‘root/origin’) was replaced by the sequence **tr/a**, as it consists of the component 一 (*a*) placed on top of the bottom part of the component 木 (*tr*). The character 便 (*biàn* ‘suitable, convenient’ or *pián* ‘[bound form]’) was represented by the sequence **l a X/sun**, since it contains the

<sup>11</sup>The component 卩 in the character 夢 may be substituted by the component 卩+ in certain fonts; nonetheless, both 卩 and 卩+ are treated as variant forms of the same component in Wieluch’s system and are both replaced by the code word **w**.

component 亠 (*l*) on the left-hand side, the component 一 (*a*) in the top right-hand corner and the combination of the components 又 (*X*) and 日 (*sun*) in the bottom right-hand corner. Similarly, the character 費 (*fèi* ‘to cost/to spend/fee’) was replaced by the sequence **bow/fj i d**, as it consists of the components 丩 (*f*) and 丨 (*i*) placed over the component 弓 (*bow*) in the top part of the sinogram and the component 貝 (*d*) in its bottom part (Wieluch 1936: 10<sub>p</sub>, 28<sub>d</sub>, 124<sub>d</sub>, 217<sub>d</sub>).

Numerals were added in front of certain code words in some of the sequences in order to avoid repetition. For example, the sinogram 桑 (*sāng* ‘mulberry tree’) was replaced by the sequence **3 it tr**, as it contains three 又 (*it*) components in the top part and the component 木 (*tr*) in the bottom part. Similarly, the character 饒 (*ráo* ‘abundant’) was assigned the sequence **eat 3 Te a ft**, as it contains the component 饣 (*eat*) on the left-hand side, three 土 (*Te*) components in the top right-hand corner and the components 一 (*a*) and 儿 (*ft*) in the bottom right-hand corner (Wieluch 1936: 10<sub>p</sub>, 54<sub>d</sub>, 116<sub>d</sub>).

However, some other sequences did include repeated code words, since no numerals were used in the same way as in the above-mentioned examples. For instance, the character 鑰 (*yuè/yào* ‘key’), which contains the component 钅 (*Me*) on the left-hand side, the component 厶 (*D*) in the top right-hand corner, three 口 (*o*) components below it and the component encoded as *vol* in the bottom right-hand corner, was substituted by the sequence **Me D o o o vol**. Similarly, the sinogram 傘 (*sǎn* ‘umbrella’), including five 人 (*l*) components and one 十 (*t*) component, was replaced by the sequence **lll t ll** (Wieluch 1936: 10<sub>p</sub>, 129<sub>d</sub>, 141<sub>d</sub>).

In some cases it was possible to dissect the same character in different ways, thus generating more than one corresponding set of components and, by extension, more than one possible sequence of code words. To avoid any confusion on the part of the reader, only one of the possible sequences was consistently used to represent such sinograms. For example, the character 辛 (*xīn* ‘pungent/laborious/eighth in order’) was assigned the sequence **stay t** (立十), and not **\*on kid**. Such sequences were also consistently used in more complex sinograms: for instance, 辣 (*là* ‘spicy’) was represented as **stay t tr/o** (立十木/口), and not as **\*on kid tr/o** (Wieluch 1936: 10<sub>p</sub>, 204<sub>d</sub>).

*Słownik chińsko-polski...* is essentially a *character dictionary* (or 字典 *zìdiǎn*; cf. Huang Chu-Ren et al. (2016: 545-551), as the vast majority of its headwords are monosyllabic lexemes, originally represented by single sinograms. Compound words, just like collocations and fixed expressions, typically appear in subentries; they contain repetition symbols (./.) instead of the headword characters and they are written with double spaces between the encodings of separate glyphs. For example, the word 日本 (*Rìběn* ‘Japan’) was listed as ./ **tr/a** under the headword 日 (*rì* ‘sun/day’ *sun*), while the word 義大利 (*Yìdàlì* ‘Italy’) was listed as ./ **big gr**

**n** under the headword 義 (*yì* ‘justice, righteousness’ **ram I**; Wieluch 1936: 183<sub>d</sub>, 204<sub>d</sub>; cf. 1.5, Figure 1). However, if a certain sinogram is used exclusively in the spelling of bound forms, i.e. parts of disyllabic, two-character words, these disyllabic lexemes appear as headwords, and not as compound words in subentries. Examples of such headwords include 骷髏 (*kūlóu* ‘human skull/human skeleton’), rendered as **os bit t o os bit lot**, 霹靂 (*pīlì* ‘thunderclap/a bolt [from the blue]’), replaced by **rain P o stay t rain G gr gr leg** and 邂逅 (*xièhòu* ‘to meet by chance’), substituted by **way n use n ox way G a o** (Wieluch 1936: 172<sub>d</sub>, 182<sub>d</sub>, 245<sub>d</sub>).

### 3 Wieluch’s method and its applications in the context of other systems

#### 3.1 Ordering dictionary entries

Wieluch’s method of encoding sinograms may be compared to a number of other systems designed to categorise Chinese characters by their structure, i.e. by the shape and number of their components. Such systems have been used for centuries for the purpose of establishing the order of sinograms (and, by extension, Chinese headwords) on the wordlists of Chinese dictionaries (cf. 1.5, 2.2). Examples of such systems include the system of 540 *character radicals* introduced in 說文解字 (*Shuowen Jiezi*) by Xu Shen (許慎), the system of 214 radicals (or 部首 *bùshǒu* in Chinese), introduced in 字彙 (*Zihui*) by Mei Yingzuo (梅膺祚) and featured in the 康熙字典 (*Kangxi Dictionary*) by Zhang Yushu (張玉書), Chen Tingjing (陳廷敬) et al. (cf. 1.3), as well as the systems of approximately 200 (e.g. 188, 189, 200 or 201) components in later simplified-character dictionaries (cf. Cheung Kam-Siu 2016: 531-539, Huang Chu-Ren et al. 2016: 545-551).

Many character components are identical in the 540-radical system, the 214-radical system and Wieluch’s 296-code word system (disregarding the differences in font styles). For instance, 火 (*huǒ* ‘fire’) is known as the 382<sup>nd</sup> radical in *Shuowen Jiezi*, the 86<sup>th</sup> radical in the *Kangxi Dictionary* and a component paired with the code word **hot** in *Słownik...* Similarly, 言 (*yán* ‘speech/to say’) is the 56<sup>th</sup> radical in *Shuowen Jiezi*, the 149<sup>th</sup> radical in the *Kangxi Dictionary* and a component represented in *Słownik...* by the code word **say**. Such correspondences often extend to the variant forms of components: for example, 灬 is classified as a variant form of 火 (**hot**) in all of these systems. In certain cases, the code word proposed by Wieluch is analogous to a single *Kangxi* (*Zihui*) radical, but to two separate *Shuowen* radicals. For instance, although the components 乙 and 乚 are both denoted by Wieluch’s code word **L** and classified as the *Kangxi* radical number 5, in *Shuowen Jiezi* 乙 is the 514<sup>th</sup> radical and 乚 is the 456<sup>th</sup> radical.

Nevertheless, Wieluch's system is actually unique in some respects. For example, despite the fact that the component 曰 (*yuē* 'to speak, to say'; not to be confused with 日 *rì* 'sun/day') is described as either the 152<sup>nd</sup> *Shuowen* radical or the 73<sup>rd</sup> *Kangxi* radical, it is a complex sinogram represented as **O a** (口一) in *Słownik...* Similarly, while the component 竹/<sup>竹</sup> (*zhú* 'bamboo') is also a character radical (number 143 or 118, respectively), in Wieluch's system it is encoded as **k k**, since each of its two parts (cf. Figure 3) is classified as a separate component (cf. *HanDian* (漢典) 2023, Wieluch 1936: *Tables I & II*, 117–122<sub>d</sub>, 162<sub>d</sub>).

Although the system of 214 radicals and its slightly altered versions are still commonly used in the creation of wordlists, indexes and entries of both monolingual and bilingual Chinese dictionaries (cf. Xia Zhengnong (夏征农) & Chen Zhili (陈至立) 2009, Xu Lin (许琳), Yao Xishuang (姚喜双), et al. 2010, Kocyba-Grych & KołECKA 2012, Achinger et al. 2023, *HanDian* (漢典) 2023, *MDBG Chinese Dictionary* 2023 etc.), Xu Shen's system has long fallen out of use and Wieluch's system never became popular, as evidenced by the fact that it has not been featured in any of the later Chinese-Polish dictionaries (cf. Sroka-Grądział 2018). This is perhaps not particularly surprising, given that the advancements made since the publication of *Słownik...* in the fields of linguistics (e.g. the invention of the Hanyu Pinyin Romanisation system), computer science (e.g. the development of the Internet or the Unicode standard) and electronics (e.g. the introduction of touchscreen devices) have largely contributed to rendering Wieluch's method replaceable and obsolete.

Wieluch himself made the following remark: “[T]here does not – and cannot – exist a dictionary [...] which would *unequivocally* lead [the reader] from the sound [of a certain Chinese word] to the glyph which it denotes.”<sup>12</sup> While this statement technically still holds true today for the majority of monosyllabic Chinese words, modern users of electronic Chinese dictionaries (cf. Achinger et al. 2023, *HanDian* (漢典) 2023, *MDBG Chinese Dictionary* 2023 etc.) have mostly become accustomed to searching headwords by rewriting the unfamiliar characters on the screen or using pronunciation-based queries. While the latter method usually does not yield accurate results (with certain exceptions: for example, the queries *shei*, *shei2* etc. unequivocally lead to the word 誰 *shéi* 'who?'), it is relatively easy to browse a number of (nearly) homophonic lemmata, which can be generated by the dictionary in a matter of split seconds. Similarly, the modern paper-based monolingual and bilingual Chinese dictionaries usually also have a Romanisation-based (typically Hanyu Pinyin-based) alphabetic macrostructure

<sup>12</sup>Original quote: “[N]ie istnieje i istnieć nie może słownik [...], któryby [sic] od tego dźwięku jednoznacznie prowadził do wyrażonego przezeń charakteru [sic].” (Wieluch 1936: 3<sub>p</sub>).

(cf. Xia Zhengnong (夏征农) & Chen Zhili (陈至立) 2009, Xu Lin (许琳), Yao Xishuang (姚喜双), et al. 2010, Kocyba-Grych & Kolečka 2012, Achinger & Radziwiłł 2015, etc.; cf. 1.5). While this significant shift in the common lexicographic practice needs to be acknowledged from the modern perspective, it may be argued that predicting such a drastic change was rather difficult (if not completely impossible) in Wieluch’s times.

### 3.2 Encoding and inputting sinograms

Wieluch’s method, designed to uniquely represent every sinogram as a sequence of code words denoting character components, may also be compared to its modern counterparts, such as the Intelligent Chinese Character Encoding System. For example, the character 峰 (*fēng* ‘peak, summit’) is rendered as *Mt ran ff* (山夕丰) in Wieluch’s system and as 山夕丰 in ICCES (Wieluch 1936: 148<sub>d</sub>, Huang Chu-Ren et al. 2016: 546–547). Unlike the method presented in *Słownik...*, the modern encoding systems are not designed to transcribe entire Chinese texts (which nowadays are rather easy to type), but may be used in electronic dictionaries to facilitate the analysis and acquisition of unfamiliar characters or to introduce cross-references to variant forms of sinograms (cf. Achinger et al. 2023, *HanDian* (漢典) 2023, *MDBG Chinese Dictionary* 2023, Huang Chu-Ren et al. 2016: 546–547).

As a tool for writing in Chinese, Wieluch’s system is based on roughly the same principles as the modern shape-based Chinese character input methods, such as *Cangjie* (倉頡), *Zhengma* (鄭碼) or *Wubi* (五筆). For example, just like the word 這 (*zhè* ‘this’) could be rendered as *way say* (讠言) in Wieluch’s system, it may also be encoded as *YYMR* (卜卜一口) in *Cangjie* (cf. *HanDian* (漢典) 2023, Wieluch 1936: 246<sub>d</sub>). However, compared to the above-mentioned methods, Wieluch’s system contains significantly more basic components and (at least currently) cannot be used to input Chinese characters on electronic devices. While it may be hypothesised that Wieluch’s system could be adapted for this purpose, it probably would not be more efficient than the already existing methods.

### 3.3 Concluding remarks

Wieluch’s encoding system was certainly a unique and interesting proposition, which aimed at solving the actual problems of contemporary non-native speakers and learners of Chinese. Although it was inspired by some earlier works of Western scholarship and the traditional classifications of character radicals, it was the very first method of analysing and transcribing sinograms to be practically applied in the creation of a Chinese–Polish dictionary. The groundbreaking

character of Wieluch's ideas is clearly demonstrated by the fact that the reference work in question was written at a time when most likely no such publications existed.

Today, this original Chinese character encoding system and the unprecedented dictionary compiled to showcase its abilities are both largely forgotten, not unlike Doman Wieluch himself. It may be argued that the solutions which he proposed almost ninety years ago were too complicated or too arbitrary, at least in certain respects. We might also suspect that the system was not sufficiently popularised at the time or demonstrate that it was effectively rendered obsolete by the advancement of technology and science. Nevertheless, appreciating this ambitious, even if ultimately unsuccessful endeavour is an essential prerequisite to fully understanding the history of Chinese-Polish lexicography. Although *Słownik chińsko-polski do czytania tekstów chińskich bez przygotowania* is not likely to receive widespread acclaim in the modern era, studying Wieluch's magnum opus could be a source of inspiration for future dictionary authors.

## References

- Achinger, Krzysztof et al. 2023. *Słownik chińsko-polski i polsko-chiński – Bohan*. <http://www.bohan.pl> (23 April, 2023).
- Achinger, Krzysztof & Antoni Radziwiłł. 2015. *Polsko-chiński i chińsko-polski słownik biznesowy 波汉-汉波商务词典*. Kęty: Wydawnictwo Bohan.
- Chalmers, John. 1882. *An account of the structure of Chinese characters under 300 primary forms after the shwoh-Wan, 100, a. D., and the phonetic shwoh-Wan, 1833*. London: Trübner & Co.
- Cheung Kam-Siu. 2016. Lexicography. In Chan Sin-wai et al. (eds.), *The Routledge encyclopedia of the Chinese language*, 215–226. London/New York: Routledge.
- HanDian (漢典). 2023. HanDian. <https://www.zdic.net> (23 April, 2023).
- Huang Chu-Ren, Li Lan & Su Xinchun. 2016. Lexicography in the contemporary period. In Chan Sin-wai et al. (eds.), *The Routledge encyclopedia of the Chinese language*, 545–562. London/New York: Routledge.
- Kocyba-Grych, Katarzyna & Zuzanna KołECKA. 2012. *Słownik chińsko-polski*. Warsaw: Wydawnictwo Edgard.
- MDBG Chinese Dictionary. 2023. MDBG. <https://www.mdbg.net> (23 April, 2023).
- Sroka-Grądziel, Anna. 2018. Z zagadnień leksykografii polsko-chińskiej. *Gdańskie Studia Azji Wschodniej* 13. 30–42.

- Wieger, Léon. 1965. *Chinese characters: Their origin, etymology, history, classification and signification. A thorough study from Chinese documents*. Trans. by Leo Davrout. 2nd edition, enlarged and revised according to the 4th French edition. New York: Paragon Book Reprint Corp./Dover Publications, Inc.
- Wieluch, Doman. 1936. *Słownik chińsko-polski do czytania tekstów chińskich bez przygotowania* 切字中波辭典. Katowice: (self-published).
- Wu Yuemei (吴月梅). 2010. *Obrazkowy słownik języka chińskiego* 汉语图解词典 (波兰语版). Beijing: 商务印书馆 (The Commercial Press).
- Xia Zhengnong (夏征农) & Chen Zhili (陈至立). 2009. *辞海*. 6th edition. Shanghai: 上海辞书出版社 (Shanghai Lexicographical Publishing House).
- Xu Lin (许琳), Yao Xishuang (姚喜双), et al. 2010. *Podręczny słownik języka chińskiego* 汉语800字. Trans. by Justyna Szpakowska. Beijing: 外语教学与研究出版社 (Foreign Language Teaching & Research Press).
- Yin Xiangfeng (殷相峰) (ed.). 2013. *汉语波兰语分类词典* *chińsko-polski słownik tematyczny*. Beijing: 商务印书馆 (The Commercial Press).



## Chapitre 15

# Les travaux lexicographiques de Carlo da Castorano et ses tentatives pour faire imprimer un dictionnaire européen de chinois

Mariarosaria Gianninoto

ReSO, Université Paul Valéry-Montpellier 3, France

Michela Bussotti

Ecole française d'Extrême-Orient, UMR 8173 (CNRS, EHESS, Université de Paris Cité, France)

The Franciscan missionary Carlo Orazi da Castorano (1673–1755) compiled a monumental trilingual Latin-Italian-Chinese dictionary (completed in 1732). Castorano tried several times to have his dictionary published, but only a few pages were finally printed. This article describes the elaboration of this dictionary by situating it in the context of the development of Sino-European lexicography, at the crossroads of various linguistic traditions, and then focuses on the attempts made by its author to have a European Chinese dictionary printed, several decades before the first Chinese-Latin-French dictionary printed in France in 1813 and the Chinese-English dictionary published in Macao by Robert Morrison (1782–1834) from 1815. These attempts came at a time when Etienne Fourmont (1683–1745) was engaged in engraving Chinese characters in wood in Paris. Despite his failure, Carlo Orazi da Castorano was one of the first missionaries who understood the importance of print for the dissemination and standardization of knowledge about the Chinese language and script in a foreign context.



## 1 Introduction

Les premiers dictionnaires européens des langues chinoises ont été conçus à la fin du XVI<sup>e</sup> siècle. Cette première étape de la lexicographie sino-européenne est associée à l'établissement de missions occidentales en Chine, les missionnaires ayant été les premiers Européens à avoir eu besoin de maîtriser le chinois pour leur vie quotidienne dans le pays, ainsi que pour leurs activités de prosélytisme (RAINI 2010, PATERNICÒ 2013). Ces dictionnaires bilingues ou multilingues, tout comme les grammaires européennes des variétés de chinois,<sup>1</sup> étaient ainsi des outils pédagogiques indispensables pour apprendre les langues locales. Un nombre considérable de dictionnaires bilingues ou trilingues apparaissent dans les siècles suivants, essentiellement dans des éditions manuscrites, avant l'essor impressionnant de la lexicographie bilingue au XIX<sup>e</sup> siècle (YANG 2014, KLÖTER 2019).

Parmi les travaux réalisés au cours du XVIII<sup>e</sup> siècle, ceux du franciscain italien Carlo Orazi da Castorano (1673–1755) méritent notre attention. Auteur prolifique, Castorano rédige aussi des ouvrages de nature linguistique. Il convient de mentionner notamment ses grammaires du chinois en latin : une version composée de 38 folios datée de 1706 et une version étendue composée de 88 folios et achevée en 1732. Castorano a également élaboré un monumental dictionnaire trilingue, le *Dictionarium Latino-Italico-Sinicum* [Dictionnaire latin-italien-chinois], achevé en 1732, qu'il a cherché à plusieurs reprises à faire publier, mais dont seules quelques pages ont été finalement imprimées.

Dans cet article, après avoir brièvement présenté les dictionnaires des langues chinoises conçus par les missionnaires européens, nous décrivons la contribution de Carlo Orazi da Castorano à la lexicographie sino-européenne, en analysant en particulier l'un des exemplaires de son dictionnaire trilingue. Nous nous penchons ensuite sur ses tentatives pour imprimer le premier dictionnaire européen de chinois.

## 2 La lexicographie missionnaire sino-européenne

La Chine possède l'une des traditions lexicographiques les plus anciennes et les plus riches au monde. Les plus anciens ouvrages lexicographiques sont rédigés au cours du premier siècle avant J.-C. (BOTTÉRO 2019) et la lexicographie chinoise se développera tout d'abord sous la forme de dictionnaires de caractères et de recueils de gloses ; les dictionnaires organisés par rimes figurent aussi plus tard

---

<sup>1</sup>Le terme « chinois » désigne en français la langue standard contemporaine, mais il est aussi utilisé pour indiquer une macro-langue, caractérisée par une forte variation diatopique et diachronique. Sur les langues chinoises, cf. PEYRAUBE (2011) et ARCODIA & BASCIANO (2021 : 6–61).

parmi ses genres principaux (VEDAL 2019). Cette tradition lexicographique est essentiellement représentée par des dictionnaires monolingues.<sup>2</sup>

Les premiers dictionnaires des variétés de chinois en langues occidentales sont écrits par les missionnaires établis en Chine et en Asie sud-orientale. Pour répondre à leurs besoins d'apprentissage des langues locales, les premiers missionnaires adaptent tout d'abord des ouvrages chinois d'apprentissage des caractères (BROCKEY 2007 : 247–249), puis rédigent également des grammaires, des méthodes et des dictionnaires des variétés de chinois en langues européennes. Les premiers dictionnaires missionnaires d'une variété de chinois ont été très probablement compilés à Manille à la fin du XVI<sup>e</sup> siècle, où résidait une communauté d'origine chinoise (ZWARTJES 2019 : 615, KLÖTER 2011). Le premier dictionnaire européen du chinois mandarin<sup>3</sup> est un dictionnaire manuscrit trilingue ayant le portugais comme langue source, des entrées traduites en italien, et le chinois comme langue cible. Compilé probablement dans la dernière décennie du XVI<sup>e</sup> siècle (RAINI 2010 : 51), ce dictionnaire est attribué aux missionnaires jésuites italiens Michele Ruggieri (1543–1607) et Matteo Ricci (1552–1610).<sup>4</sup> L'élaboration de dictionnaires de chinois en langues européennes doit répondre au défi d'inclure un système d'écriture non alphabétique, les caractères chinois ou sinogrammes, dans des ouvrages en langues alphabétiques. Des systèmes de romanisation sont ainsi élaborés pour noter la prononciation et faciliter la mémorisation des mots. Nous trouvons des dictionnaires entièrement rédigés en lettres latines et n'indiquant donc les mots chinois qu'en romanisation. Le dictionnaire espagnol-chinois *Vocabulario de la lengua Mandarina* (1687) du missionnaire dominicain Francisco Varo (1627–1687) ne comporte par exemple aucun caractère chinois.<sup>5</sup> Toutefois, dans la plupart des dictionnaires, caractères chinois et roma-

<sup>2</sup>La lexicographie bilingue et multilingue était beaucoup moins présente dans la tradition linguistique chinoise. Ceci est probablement lié au statut des langues étrangères en Chine, qui n'étaient pas valorisées dans le système éducatif traditionnel (PEVERELLI 1986 : 28, PELLIN 2009 : 40) et dans les concours mandarinaux (avec quelques exceptions, ELMAN 2000 : 166–167, 223), ne faisant pas partie de la formation des élites (LACKNER 2001 : 358).

<sup>3</sup>Pour une analyse détaillée de la langue décrite dans ce dictionnaire, cf. YANG (2001).

<sup>4</sup>Ce manuscrit a été découvert par le jésuite Pasquale D'Elia dans les archives romaines de la Compagnie de Jésus en 1934. D'Elia l'a appelé *Dizionario Portoghese-Cinese* [Dictionnaire portugais-chinois] et en a attribué la paternité aux missionnaires jésuites Michele Ruggieri et Matteo Ricci, même si Ruggieri devrait probablement être considéré comme l'auteur principal. Sur ce dictionnaire cf. YANG (2001).

<sup>5</sup>Selon COBLIN (2006 : 12), ce dictionnaire a été élaboré par Varo « *by the final decade of his life* ». Pour ce dictionnaire, nous avons consulté la version numérisée du manuscrit « Libr. Sin. 29 » de la Bibliothèque nationale d'Allemagne (accessible à l'adresse suivante : [https://digital.staatsbibliothek-berlin.de/werkansicht?PPN=PPN3308102102&PHYSID=PHYS\\_0005&DMDID=DMDLOG\\_0001](https://digital.staatsbibliothek-berlin.de/werkansicht?PPN=PPN3308102102&PHYSID=PHYS_0005&DMDID=DMDLOG_0001)) ainsi que l'édition critique de COBLIN (2006).

nisation coexistent. Ces dictionnaires doivent ainsi prévoir « *a digraphic arrangement of entries involving Chinese characters and transcriptions using the Roman alphabet* » (KLÖTER 2019 : 325). C'est le cas par exemple du *Diccionario de la lengua Mandarina*, écrit par le missionnaire dominicain espagnol Francisco Díaz (1606–1646), dont les entrées sont représentées par les caractères, dans le cadre de la tradition lexicographique chinoise. Le dictionnaire de Díaz marque un développement important de la lexicographie bilingue, les entrées en chinois y étant disposées selon l'ordre alphabétique de prononciation de l'initiale, suivant la tradition lexicographique occidentale (MASINI 2006 : 236). Cette innovation dans l'organisation des entrées avait été rendue possible par l'introduction des systèmes de romanisation<sup>6</sup> (LUÓ CHÁNGPÉI 羅常培 1930). Ces ouvrages pionniers dus aux missionnaires jésuites et dominicains ouvrent la voie à une riche production de dictionnaires bilingues ou plurilingues, souvent rédigés par des missionnaires franciscains, et qui proposent le chinois comme langue source, plus rarement comme langue cible. Ces ouvrages nous sont parvenus dans des éditions manuscrites, datant du XVIIe et du XVIIIe siècles, tandis qu'il faut attendre le début du XIXe siècle pour les premiers dictionnaires imprimés. C'est dans la tradition de cette production que s'inscrit l'ouvrage trilingue de Carlo Orazi da Castorano et dans ce contexte historique que se situent ses tentatives de faire imprimer un dictionnaire européen du chinois.

### 3 Carlo Orazi da Castorano et son dictionnaire latin-italien-chinois

#### 3.1 Carlo Orazi da Castorano

Carlo Orazi (ou Carlo Horatii da Castorano, Carlo di Orazio da Castorano ; en chinois Kāng Hézi<sup>7</sup> 康和子) est un missionnaire franciscain, né en 1673 à Castorano (près d'Ascoli Piceno, dans la région italienne des Marches). Il entre dans les ordres mineurs en 1690. Arrivé en Chine vers 1700, il y vivra pendant plus de trois décennies. Carlo Orazi travaillera dans les deux provinces du diocèse de Pékin (Shandong et Zhili), cf. CATTO 2013). En 1724, il se retire dans la résidence

---

<sup>6</sup>Les dictionnaires monolingues chinois étaient traditionnellement classés par catégories sémantiques, par radicaux et traits, ou encore par tons et rimes.

<sup>7</sup>Dans le présent article, le système de transcription *pinyin* (système de romanisation officiel de la République Populaire de Chine), est utilisé. Les romanisations originales sont maintenues dans les citations. Les caractères traditionnels (non simplifiés) sont utilisés pour toutes les sources de l'époque impériale. Pour les sources contemporaines, le choix entre caractères simplifiés ou traditionnels dépend de la source elle-même.

des missionnaires de Propaganda Fide près de Pékin. Il est très impliqué dans la controverse sur les rites chinois (VON COLLANI 2016).<sup>8</sup> En 1733, il quitte la Chine pour se rendre à Rome, puis retourne à Castorano en 1741, où il meurt en 1755 (CATTO 2013).

Ce missionnaire était un auteur prolifique, ayant écrit de nombreux textes sur la controverse des rites chinois, sur la mission en Chine, sur les activités de prosélytisme (YU 2017). Parmi ses écrits, nous pouvons citer la *Parva elucubratio super quosdam libros sinenses* [Brève explication sur quelques livres chinois], catalogue manuscrit décrivant les livres chinois ayant appartenu au missionnaire franciscain Giovanni Francesco Nicolai da Leonessa et le *Brevis narratio itineris ex Italia usque ad Chinam* (1739) [Bref récit du voyage de l'Italie en Chine], qui, comme le précise Stafutti, est un « récit des activités des missionnaires catholiques en Chine, [décrivant] les critères suivis par les différents ordres religieux pour mener à bien leur travail d'évangélisation, et les controverses qui ont surgi entre les ordres et le Saint-Siège en raison de ces critères<sup>9</sup> » (STAFUTTI (1981 : 35), notre traduction).

Parmi ses ouvrages, figurent aussi deux grammaires. Tout d'abord, la *Manuductio ad linguam sinicam*, une « Introduction à la langue chinoise », rédigée pour la première fois en 1706, dans une version composée de 38 folios. La *Grammatica seu Manuductio ad linguam sinicam* [Grammaire ou Introduction à la langue chinoise, 1732] est une version étendue de la première grammaire.<sup>10</sup> La *Grammatica* comprend 88 folios et inclut les caractères chinois, alors que seules les transcriptions étaient utilisées pour écrire les mots chinois dans la grammaire précédente.

Enfin, Castorano a également élaboré un dictionnaire trilingue et a participé à d'autres projets lexicographiques, notamment en copiant des dictionnaires manuscrits et en cherchant à faire imprimer son propre dictionnaire.

---

<sup>8</sup>Comme le précise VON COLLANI (2016 : 479), « *The so-called Chinese Rites Controversy is the terminus technicus for the quite heated debates and arguments about the meaning of Chinese terms and Rites. It is a story which lasted more than 100 years and absorbed much energy in Europe and in China* ». Sur la question des rites cf. aussi MUNGELLO (1994).

<sup>9</sup>Texte original en italien : « *testimonianza sull'attività dei missionari cattolici in Cina, sui criteri seguiti dai vari ordini religiosi nel condurre la propria opera di evangelizzazione e sulle controversie che proprio a causa di questi criteri sorsero tra gli ordini stessi e la S. Sede* ».

<sup>10</sup>Sur la *Grammatica seu Manuductio ad linguam sinicam*, cf. LI (2017).

### 3.2 Le *Dictionarium Latino-Italico-Sinicum*

L'exemplaire consulté du *Dictionarium Latino-Italico-Sinicum* [Dictionnaire latin-italien-chinois], achevé en 1732 à Pékin, comprend 1 008 folios.<sup>11</sup> Le titre complet de cet ouvrage en explicite la nature de dictionnaire trilingue et donne des indications sur son lectorat cible : *Dictionarium Latino-Italico-Sinicum Tàm vocum, quàm Litterarum seu Characterum usualium Sinensium ad usum et commoditatem PP. Missionariorum in hanc Sinicam Missionem noviter adventantium* [Dictionnaire latin-italien-chinois des mots, des lettres ou caractères du chinois courant pour l'usage et le profit des missionnaires arrivant dans cette mission chinoise].

Ce manuscrit s'ouvre par des « Avertissements aux lecteurs » (*Ad Lectores Monita*<sup>12</sup>), où l'auteur décrit le contexte dans lequel son ouvrage a vu le jour et donne des précisions sur son processus de rédaction. Cette préface est suivie par la *Grammatica seu Manuductio ad linguam sinicam* [Grammaire ou introduction à la langue chinoise<sup>13</sup>], la deuxième version, étendue, de la grammaire chinoise de Carlo da Castorano. Puis, nous avons le dictionnaire, dont la langue cible est le chinois<sup>14</sup> et la langue source est le latin, bien que les entrées soient aussi traduites en italien. Ce dictionnaire devait en effet combler l'absence d'outils adaptés pour les missionnaires italo-phones, lacune dont l'auteur se rendra compte dans son processus d'apprentissage de la langue. Carlo da Castorano commence ainsi la rédaction de son ouvrage sur incitation du père Giovanni Battista da Illiceto, missionnaire de l'Ordre des Frères Mineurs (LI 2017 : 178). Toutefois, Carlo Orazi opte pour un dictionnaire latin et dans « Les avertissements aux lecteurs » explique ce choix par sa volonté d'élaborer un ouvrage accessible et utile à un public plus large :

Il vaut mieux que je fasse un dictionnaire latin-chinois, car il sera plus utile, y compris pour les missionnaires des autres pays, et pour beaucoup d'autres personnes.<sup>15</sup>

---

<sup>11</sup>L'édition consultée est la version numérisée du manuscrit « Vat.estr.or.4 » de la Bibliothèque Apostolique Vaticane, qui correspond à l'une des versions manuscrites autographes (écrite par Carlo da Castorano, à l'exception des caractères chinois du dictionnaire, dus à des lettrés chinois, cf. LI 2017 : 180). Manuscrit accessible à l'adresse suivante : [https://digi.vatlib.it/view/MSS\\_Vat.estr.or.4](https://digi.vatlib.it/view/MSS_Vat.estr.or.4).

<sup>12</sup>Pages 3 à 15 du manuscrit consulté.

<sup>13</sup>Pages 17 à 103.

<sup>14</sup>Mandarin septentrional, enrichi de mots et d'expressions venant de la région du Shandong, en raison du long séjour de Carlo da Castorano à Linqing 臨清 LI (2017 : 177).

<sup>15</sup>Texte original : « *Verum melius est si faciam dictionarium Latino-Sinicum, sic enim erit magis comunicabile bonum, etiam pro aliis Ppbus Misionariis aliarum Nationum, et pro multis Tertibus* » ORAZI (s. d. : f. 4).

Dans la préface, Castorano décrit le processus de rédaction de son dictionnaire, en commençant par reprendre les entrées du dictionnaire de Cesare Calderino Mariani,<sup>16</sup> ne pouvant pas disposer du dictionnaire d'Ambrogio Calepino<sup>17</sup> (« *statim misi manum ad opus, deferibendo scilicet Dictionarium Calderinum [carebam enim Calepino]* »).<sup>18</sup> Partant de cette base, l'auteur ajoute, lorsque cela était possible, les traductions chinoises des entrées, aussi bien en transcription qu'en caractères chinois (« *ut dum liceret verba sinica pomerem, et tandem Litteras seu characteres sinicos opponi facerem*<sup>19</sup> »). Il abandonna ensuite ce travail jusqu'en 1728, année où il en reprit la rédaction. Il disposera alors du dictionnaire de Calepino, qui sera une autre source importante pour les entrées en latin (cf. LI 2017 et CASTORINA 2019).

Les sources des équivalents chinois sont variées. Nous pouvons tout d'abord mentionner les dictionnaires chinois-latin de son confrère Basilio Brollo (1684-1704), nom chinois Yè Zūnxiào 葉尊孝.<sup>20</sup> Brollo est l'auteur de deux dictionnaires. Le premier dictionnaire, complété en 1694, était organisé selon le système chinois des radicaux et des traits et contenait plus de 7 000 entrées ; le second dictionnaire, achevé en 1699, était classé par ordre alphabétique, en fonction de l'initiale du mot en romanisation et contenait plus de 9 000 entrées. Il est important de mentionner que l'un des exemplaires du dictionnaire alphabétique de Brollo a été recopié par Castorano,<sup>21</sup> qui a inclus dans ce même manuscrit la première version de sa grammaire (*Manuductio ad linguam sinicam*, 1706). Raini souligne également les similitudes au niveau des systèmes de transcription, affirmant que « la romanisation utilisée dans le dictionnaire et la grammaire de Castorano, coïncide parfaitement avec la romanisation principale utilisée dans la copie du dic-

<sup>16</sup>Il s'agit du *Dictionarium tum latini, tum italici sermonis studiosis, apprime congruens*, dont la première édition date de 1587 et qui a fait l'objet de nombreuses éditions aux XVIIe et XVIIIe siècles LI (2017 : 178).

<sup>17</sup>La première édition du dictionnaire latin de l'augustinien Ambrogio Calepino (1435-1511), intitulée *Ambrosii calepini bergomatis eremitani dictionarium*, a paru en 1502. Ensuite « *at least two hundred editions were published before 1779, starting with three languages (Latin, Greek, and Italian) before being expanded to a total of eleven languages* » (ZWARTJES 2021).

<sup>18</sup>ORAZI (s. d. : f.4)

<sup>19</sup>*Ibidem*.

<sup>20</sup>Mattia Andrea Brollo (1648-1704) est également un missionnaire franciscain. Né à Gemona près de Udine (d'où son nom Père Basilio « da Gemona », parfois transformé en « da Glemona »), il est ordonné en 1674 et arrive en Chine en 1684. Il est nommé vicaire apostolique du Shaanxi en 1696 où il se rendra quelques années plus tard et où il restera jusqu'à sa mort en 1704. À son propos, voir POLMONARI (2008).

<sup>21</sup>Il s'agit du manuscrit « Vat. estr.-or. 8 » de la Bibliothèque apostolique vaticane, intitulé *Dictionarium Sinico-Latinum Reverendissimi Patris Basilij à Glemona* [Dictionnaire chinois-latin du très révérend père Basile da Gemona].

tionnaire de Brollo datée de 1699 » et qu'il s'agit « d'une transcription fortement marquée par la prononciation italienne » (RAINI 2010 : 236 ; notre traduction).<sup>22</sup>

Dans la préface, Castorano explique combien il est difficile de trouver les équivalents de nombreuses entrées latines. Plusieurs équivalents ou explications sont fournis par des informateurs chinois. Ces informateurs étaient des domestiques et paysans qu'il avait pu rencontrer dans son entourage, puis, dans un deuxième temps, un enseignant (cf. LI 2017 : 179). Les caractères chinois manquaient pour de nombreux mots de la langue parlée (« *multa verba esse in locutione vulgari sinico, quibus Litterae seu characteres Sinici vulgares non occurrant* », ORAZI s. d. : 6). Des équivalents écrits sont repris de la lexicographie sino-mandchoue.<sup>23</sup> Le dictionnaire de Castorano, enrichi de tous ces apports, constitue une source importante pour la langue de l'époque. L'intérêt pour la langue parlée, le registre familier, les mots de la vie quotidienne, tout ce qui était motivé par les besoins de la vie de tous les jours mais aussi pour les activités d'évangélisation, mérite d'être souligné.<sup>24</sup> Il s'agit d'une caractéristique importante de la lexicographie missionnaire, par rapport à la tradition lexicographique autochtone chinoise, centrée sur la norme écrite.

En ce qui concerne la macrostructure, les entrées sont classées en fonction de la première lettre du mot latin ; la première entrée est constituée par les prépositions latines *A, ab, ex*. « de, depuis » ; la dernière entrée est *zhytyum* [en latin] /*acqua d'orzo tostato* [en italien], « *zythum* » ou « *zython* ». Cette organisation par ordre alphabétique selon l'initiale s'inscrit dans le cadre de la tradition lexicographique occidentale et elle est liée au choix de la langue source, mais aussi aux modèles suivis et aux sources adoptées par Castorano. En effet, la première entrée correspond, y compris en ce qui concerne la définition, à une entrée du dictionnaire de CALDERINO (1586 : 12) ; la dernière entrée, par contre, correspond à la dernière entrée du dictionnaire de Calepino.<sup>25</sup>

---

<sup>22</sup>Texte original italien : « *la romanizzazione usata nel vocabolario e nella grammatica di Castorano, coincide perfettamente con la romanizzazione principale utilizzata nella copia del dizionario di Brollo datata 1699* » ; « *una trascrizione a forte base italoфона* ».

<sup>23</sup>ORAZI (s. d. : 6) parle d'un « dictionnaire mandchou-chinois » (« *Tartaro=Sinico Dictionario* »). La Chine de l'époque Qing (1644-1911) est un empire multilingue où le mandchou (langue de la dynastie régnante) est une langue officielle. Une riche production de manuels et de dictionnaires bilingues chinois-mandchou ou trilingues chinois-mandchou-mongol voit le jour (cf. SÖDERBLOM SAARELA 2019). Ces manuels et dictionnaires constituent une source importante pour l'étude du mandarin de l'époque.

<sup>24</sup>LI (2017 : 183) écrit à cet égard que « l'auteur a transcrit pratiquement toutes les entrées qui concernaient chaque aspect de la vie quotidienne » (« *l'autore ha trascritto quasi tutti i lemmi che hanno a che fare con qualsiasi ambito della vita quotidiana* »).

<sup>25</sup>Dans l'édition que nous avons pu consulter, datée de 1513.

Chaque page compte dix-huit entrées. Les entrées se situent à gauche et les équivalents chinois à droite de la page. Comme c'est le cas dans la plupart de ces dictionnaires, bilingues ou multilingues, nous retrouvons une mise en page à l'occidentale : la page est organisée par lignes et non par colonnes (ce qui est habituellement le cas dans la lexicographie traditionnelle chinoise), et écrite de gauche à droite (tandis que les colonnes sont écrites de droite à gauche dans les ouvrages chinois de l'époque).

En ce qui concerne la microstructure, les entrées se composent de mots ou de locutions d'abord en latin (éventuellement accompagnés de synonymes), puis en italien (les traductions en italien sont placées en dessous de l'entrée en latin) ; les équivalents chinois sont indiqués en caractères chinois mais aussi en romanisation (qui apparaît sous les caractères correspondants) ; certaines entrées incluent également des exemples illustratifs en caractères chinois et en romanisation. Cela correspond à un objectif pédagogique : la romanisation était utile pour les missionnaires et les apprenants visant à étudier uniquement la langue parlée, les caractères et la romanisation pour ceux souhaitant étudier à la fois la langue parlée et la langue écrite, comme l'explique Castorano dans sa préface :

Ainsi, les missionnaires [ou bien les autres (lecteurs)] qui souhaitent uniquement apprendre la langue chinoise (parlée), ne doivent lire que les mots placés en bas à droite ; mais ceux qui souhaitent apprendre aussi les caractères chinois, doivent étudier les caractères disposés sur la même ligne ou au-dessus.<sup>26</sup>

Nous citons des exemples constitués par des entrées successives, pour donner une idée de l'organisation des entrées dans ce dictionnaire (notre traduction française est ajoutée entre parenthèse) (cf. Tableau 1).

Dans cette séquence d'entrées,<sup>27</sup> la première entrée est constituée par le lemme latin « *Abloco, as* », correspondant à la première et à la deuxième personne du présent du verbe *ablocare* « louer ». À côté de ce lemme, nous trouvons un renvoi au synonyme « *Loco* » (première personne du verbe *locare* « louer »). Ce renvoi est présent dans le dictionnaire de Calepino, où à l'entrée *Abloco, as*, le lecteur est invité à « se rapporter au mot '*Loco*' » (« *vide in dictione loco* », CALEPINO

<sup>26</sup>Texte original en latin : « *Ergo PP Missionarij [aut etiam Alii] qui volerint addiscere solumodò Linguam sinicam, legant solum nomina seu verba in linea recta inferiorius posita; Qui autem addiscere voluerint etiam characteres Sinicos, inspiaciant eos intra eadem lineam seu areolam superius dispositos* » ORAZI (s. d. : 6)

<sup>27</sup>F. 2v du dictionnaire. Il est important de préciser que la première page du dictionnaire se trouve à la page 109 du manuscrit (les pages précédentes étant numérotées jusqu'à 108, en haut de la page).

TABLE 1 : L'organisation des entrées

|                                                                                                            |                                                                                                                                                                                                                                |
|------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Abloco, cas<br>(Je loue, tu loues)<br>Vide Loco<br>(cf. [entrée] « Loco »)<br>Locare, affittare<br>(Louer) | 賃。租。賃房子。租房屋。<br>jín lín zū lín fàngzù zū fàngwǔ<br>(Donner ou prendre à loyer. Louer. Donner à loyer une maison. Louer une maison)                                                                                             |
| Abluo, is<br>(Je lave, tu laves)<br>Lavare<br>(laver)                                                      | 洗。洗一洗。  洗手。洗臉。  洗滌。<br>盥。澣。盥手<br>sì sì ĭ sì v.g. sì scèu sì lièn sì tǐ kuòn<br>han kuòn scèu<br>(Laver. Laver un peu. p. ex. Se laver les mains. Se laver le visage. Nettoyer. Se débarbouiller. Se laver. Se laver les mains) |
| Abluere seu lavare corpus<br>(laver ou se laver le corps)                                                  | 洗澡。洗身子。沐浴<br>sì zào sì scīnzù mǔiǔ mǔiǔ<br>(Prendre un bain. Laver son corps. Se baigner)                                                                                                                                      |

1513 : 3). Dans son dictionnaire, Carlo da Castorano fournit également l'entrée en italien, sous forme des verbes à l'infinitif *locare* et *affittare* « louer » (placé en dessous de l'entrée en latin). Comme le montre cet exemple, c'est le latin qui est pris en compte pour la disposition alphabétique. Deux équivalents chinois figurent pour cette entrée. Pour le premier équivalent, deux prononciations sont indiquées. Ces équivalents sont suivis par les exemples de collocations (ayant le sens de « louer une maison »).

Dans le dictionnaire de Carlo da Castorano, l'entrée « *Abloco, as* » est suivie par l'entrée « *Abluo, is* ». L'ordre ne correspond pas à celui du dictionnaire de Calepino (qui place *Abluo* devant *Abloco*; cf. CALEPINO 1513 : 3), mais correspond à l'ordre du dictionnaire de Calderino Mariani, lequel avait constitué sa principale source pour établir la structure des entrées latines de son dictionnaire (CALDERINO 1586 : 13).<sup>28</sup> L'entrée « *Abluo, is* », est accompagnée de sa traduction italienne « *Lavare* » (« laver »). L'entrée suivante est « *Abluere seu lavare corpus* » (« laver ou se laver le corps »), qui ne figure pas à la suite de l'entrée « *Abluo* »

<sup>28</sup>Toutefois, ce ne sont pas toutes les entrées latines qui sont reprises, du moins dans l'édition que nous avons pu consulter, le *Dictionarium D. Caesaris Calderini Mirani Veronensis, tum latini, tum italici sermonis studiosis, apprimè congruens*, 1586, accessible à l'adresse suivante : <http://bdh.bne.es/bnearch/detalle/bdh0000209967>.

dans le Calderino et qui est ajoutée par Castorano. Dans ces deux entrées, nous avons des synonymes ou quasi-synonymes, appartenant souvent à des registres différents : par exemple des mots de registre soutenu, comme *guàn* 盥 et *huàn* 澣, figurent à côté d'expressions et tournures relevant davantage de la langue parlée, comme *xǐ yī xǐ* 洗一洗 « se laver (un peu) ». Nous trouvons aussi des collocations et des exemples illustratifs, comme *xǐ shǒu* 洗手 « se laver les mains » et *xǐ liǎn* 洗臉 « se laver le visage ». Les divers exemples sont introduits par l'abréviation latine « v.g. » (« p. ex. ») mais uniquement dans la ligne écrite en romanisation, tandis qu'ils sont séparés par le symbole | dans la ligne écrite en caractères.

Un autre exemple d'entrée<sup>29</sup> :

TABLE 2 : L'entrée de 'Canus, Canutus'

|                                                           |                                                                                                                                                                                                                                           |
|-----------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Canus, ni et Canutus, tis<br>Cano Canuto<br>(chenu vieux) | 頭髮白了。白頭髮的人。年老了。年老的人<br>tèufǎ pě lào pě tèufǎ tǐ jìn nièn lào<br>lèào nièn lào tiè jìn<br>îd est Senex<br>(Les cheveux sont devenus blancs. Une personne qui a les cheveux blancs. Il a vieilli. Une personne âgée. C'est un vieil homme). |
|-----------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

Dans cette entrée, Castorano propose des périphrases pour rendre les mots latins (entrées au nominatif et génitif). La traduction italienne est fournie (*cano*, *canuto*, « aux cheveux blancs »). Pour les traductions en chinois, nous avons d'abord une phrase (p. ex. « Les cheveux sont devenus blancs »), qui est reprise pour former un syntagme (« Une personne qui a les cheveux blancs » ou « Une personne aux cheveux blancs »). Le même procédé est utilisé pour le deuxième équivalent (« Il a vieilli » ; « Une personne qui a vieilli » et donc « Une personne âgée ») afin de rendre avec une périphrase le sens des mots latins. Il convient de préciser que les deux équivalents sont séparés graphiquement par un symbole vertical à l'encre rouge.

L'entrée suivante, concernant la conjonction *tum*,<sup>30</sup> est intéressante du fait des exemples illustratifs en langue parlée qu'on y retrouve (cf. Tableau 3).

Dans cette entrée, une série de synonymes sont présentés comme équivalents du mot latin, accompagnés de phrases de la vie quotidienne. Aucune définition ou explication de nature grammaticale n'est fournie.

<sup>29</sup>F. 61r du dictionnaire.

<sup>30</sup>F. 477v.

TABLE 3 : l'entrée 'Tum'

|                                           |                                                                                                                                                                                               |
|-------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Tum copula<br>(alors, copule)             | 也。亦。又。這個也用著那個亦要。又要這個又要那個                                                                                                                                                                      |
| Si. cosi. ancora.<br>(Oui, ainsi, encore) | iè ĩ ieú v.g. cé kó iè iúng ciò, nákó ĩ iáo ieú iáo cé kó ieú iáo ná kó<br>(Aussi. Ainsi que. Encore. p. ex. J'utilise celui-ci, je veux aussi celui-là. Je veux celui-ci et aussi celui-là.) |

La première entrée du dictionnaire constitue un autre exemple d'entrée de nature grammaticale, où une série de mots fonctionnels (des prépositions) sont définis puis traduits par des équivalents (cf. Tableau 4).<sup>31</sup>

TABLE 4 : L'entrée 'A, ab, abs, ex'

|                                                                                                                                         |                                                                                                                |
|-----------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------|
| A, ab, abs, ex                                                                                                                          | 從。自。由。從小慣的                                                                                                     |
| Praeposit. <sup>nes</sup> servie <sup>tes</sup> ablat.o casui<br>(de, dès, depuis, par<br>Prépositions utilisées avec les cas ablatif.) | zûng. zú. ièu v.g. zûng siào kuón dĩ<br>a puero usus &c.<br>(De, Depuis. Par. P. ex. Utilisé par les enfants.) |

Il convient de souligner que cette première entrée, ainsi que les deux suivantes, sont écrites à l'encre rouge (à la différence des autres entrées, rédigées à l'encre noire). Par ailleurs, un caractère est raturé dans l'entrée et corrigé (慣) sur le côté gauche de la page. Cette première entrée introduit une série de prépositions latines, dont les premières commencent par la lettre « a », et qui sont employées avec le cas ablatif pour exprimer l'origine, le point de départ spatial et temporel, la cause et l'agent. Trois prépositions chinoises sont proposées comme équivalents : *cóng* 從 et *zì* 自 « de, depuis » et *yóu* 由 « de, par », dont la deuxième relève d'un registre plus soutenu et la troisième peut introduire l'agent. Dans cette entrée, nous trouvons par ailleurs une définition de ces mots qui en explique l'emploi : « prépositions utilisées avec les cas ablatif », définition qui est reprise du dictionnaire de Calepino,<sup>32</sup> et abrégée dans l'entrée de Carlo da Castorano.<sup>33</sup>

<sup>31</sup>F. 1r du dictionnaire.

<sup>32</sup>Où nous lisons « *A & Ab praepositiones ablativo casui servientes significant* » CALDERINO (1586 : 12).

<sup>33</sup>Dans l'entrée : « *praeposit.<sup>nes</sup> servie<sup>tes</sup> ablat.o casui* », c'est-à-dire *praepositiones servientes ablativo casui*.

Cette entrée témoigne ainsi de l'influence des modèles occidentaux sur ce dictionnaire. L'auteur y introduit une définition de nature grammaticale reprise des sources latines. Il convient toutefois de souligner que ce dictionnaire ne propose généralement que les équivalents en chinois et des exemples, sans indications concernant leur emploi LI (2017 : 185). Cet ouvrage était probablement conçu comme un outil pédagogique à utiliser en complément de la grammaire de l'auteur, contenue dans le même manuscrit.

## 4 Les tentatives avortées de publier les dictionnaires

Le trait le plus marquant que l'on retient toujours à propos du missionnaire franciscain est son opposition aux rites chinois, avec une telle véhémence que, une fois à Rome, il fut invité à limiter ses actions à ce propos.<sup>34</sup> Mais dans le cadre de notre présentation, le point marquant est que, contrairement à ses contemporains et ses confrères qui semblent se satisfaire et s'accommoder de l'usage des manuscrits, Castorano avait compris le potentiel de l'imprimerie pour diffuser ses travaux et standardiser l'information linguistique. Il chercha donc à plusieurs reprises à mettre en œuvre ce projet, car – comme c'est expliqué dans certains de ses écrits – il aspirait à un cercle de lecteurs plus large que les seuls religieux. Cependant, il semble avoir manqué une occasion unique de publier ses dictionnaires en passant par Paris, lors de son voyage de retour de la Chine en Europe, justement parce qu'il était trop préoccupé par les questions religieuses et donc pressé de se rendre à Rome.

### 4.1 Le passage à Paris

Castorano arrive en France le 23 août 1734 ; pendant le voyage, il s'est occupé de la copie de l'un de ses dictionnaires (LI 2017 : 182). Selon un petit livre à son nom publié de manière posthume à Livourne en 1759 (*Brevissima notizia, o relazione di varj viaggi, fatiche, patimenti, opere ec. nell'Imperio della Cina, ec. Del Reverendo padre F. Carlo Horatii da Castorano*), son navire arrive en Bretagne, dans le port Luigi o Ludovico (la citadelle de Port-Louis à Lorient ?). De ce lieu, Castorano se rend à Rehen (Rouen ?), puis à Versailles, où il aurait été reçu par le roi, la reine et le cardinal de Fleury, précepteur de Louis XV. Il va s'arrêter ensuite à

---

<sup>34</sup> Auteur de 37 mémoriaux sur la question, même après la publication de la bulle papale de condamnation définitive des rites chinois (1742), Castorano continua à solliciter les autorités pontificales sur ces questions, jusqu'en 1750 quand on lui ordonna de s'abstenir de ces démarches. Cf. CATTO (2013), voir la version en ligne [https://www.treccani.it/enciclopedia/antonio-orazi\\_%28Dizionario-Biografico%29/](https://www.treccani.it/enciclopedia/antonio-orazi_%28Dizionario-Biografico%29/) (dernier accès, le 09/04/2023).

Paris, avant de repartir pour le Sud, en s'embarquant à Marseille pour Livourne, puis en traversant la Toscane : il arrive à Rome en novembre 1934 (ORAZI 1759 : 60–62).

Pendant son séjour parisien, il loge chez Monseigneur Raniero Delci (1670-1761), nonce apostolique nommé auprès du roi de France entre 1730 et 1738. À cette occasion Castorano rend visite au Cardinal de Polignac (Melchior de Polignac, 1661-1741) qui lui propose de rester à Paris pour imprimer aux frais du roi le dictionnaire latin-italien-chinois. C'est probablement à ce moment que le missionnaire italien a été introduit auprès d'Étienne Fourmont (1683-1745). Nous savons que Polignac communique avec Fourmont à propos des « choses chinoises » : par exemple, lors d'une rencontre à l'Académie Royale des Inscriptions et Médailles portant sur la « Bussole chinoise » en date du 26 janvier 1734, Polignac, Fourmont et Nicolas Fréret (1688-1749) échangent sur le figurisme (LEUNG 2002 : 56–58, 275–276). Polignac avait donc la position pour mettre Castorano en contact avec Fourmont qui, à ce moment-là, était déjà engagé depuis une vingtaine d'années dans le projet de gravure de caractères mobiles en bois, afin de réaliser une « fonte » la plus complète possible de types chinois, et ainsi imprimer à Paris cette langue d'une façon adaptée aux pratiques typographiques d'impression européennes. Le projet de Fourmont est mieux connu par le nom des caractères chinois, appelés les « buis du régent »<sup>35</sup>. Il avait débuté en 1615, lorsque l'abbé Jean-Paul Bignon (1662–1743), directeur des Académies royales et de la Bibliothèque du Roi, en confia la direction à Étienne Fourmont, assisté de son frère cadet, l'abbé Michel (1690-1746) et d'un chrétien chinois présent à Paris, Arcade Huang (Huáng Jiālùe 黃嘉略).<sup>36</sup> Mais Huang décède en 1716. On comprend l'apport potentiel pour le projet français d'une collaboration avec Castorano, vu les connaissances linguistiques acquises sur le terrain par le missionnaire italien, à un moment où les personnes pouvant prétendre connaître la langue sont extrêmement rares.

Fourmont écrit d'ailleurs dans ses *Meditationes sinicae* (1737) au sujet d'une rencontre avec Castorano – sous le nom de « Casaranus »<sup>37</sup> –, il cite sa grammaire et son *Dictionarium italico-sinicum*, et il fait aussi mention d'une admiration réciproque. Plus tard, cette information est reprise par Jean-Pierre Abel-Rémusat (1788–1832) dans son *Plan d'un dictionnaire chinois*. Le premier sinologue du Collège de France, non moins attaqué par ses détracteurs à propos de ses limites

<sup>35</sup>Ces caractères mobiles en bois sont ainsi désignés parce que leur gravure a commencé sous la régence (1715–1723) de Philippe d'Orléans (1674–1723). Ils existent encore et sont conservés à l'Imprimerie nationale.

<sup>36</sup>À propos d'Arcade Huang et sa collaboration avec Fourmont, voir XU MINGLONG 許明龍 (2004), SIEBER (2013).

<sup>37</sup>Il donne comme nom complet Carulus Oratius de Casarano, et il se trompe d'année, car il parle de 1735, donc un an plus tard que le passage de l'Italien à Paris, « Praefatio » : XXIV.

linguistiques que le fut Fourmont (les deux hommes ayant manqué d'un apprentissage direct de la langue chinoise, notamment de la langue orale), nommé élogieusement Castorano, tout en se référant aux manuscrits de Brollo présents à Paris (ABEL-RÉMUSAT 1814 : 9-12). Nous ne pouvons pas établir avec certitude si, dans ce passage, Abel-Rémusat commente le manuscrit « Chinois 9278 » – un dictionnaire chinois-latin qui se trouve aujourd'hui à la Bibliothèque nationale de France et qui inclut une version de la grammaire de Castorano –, ou d'autres dictionnaires liés au nom de Brollo, par exemple le manuscrit « Chinois 9279 », conservé dans la même bibliothèque.

#### 4.2 Les difficultés romaines

Dès son retour en Italie, Castorano souhaite imprimer un volume de son dictionnaire trilingue. Deux versions de ce dictionnaire sont aujourd'hui conservées à la bibliothèque Vaticane, ainsi que quelques pages et des fragments d'un volume inabouti, témoignant de malheureuses tentatives de publier l'ouvrage.<sup>38</sup> Dans ce volume, le titre en haut de la première page est *Dictionarium // Latino-Italico-Sinicum // Tàm vocum, quàm Characterum // usualium Sinensium* [Dictionnaire Latin, Italien, Chinois, avec les caractères usuels et leur prononciation] et le contenu est organisé dans la page imprimée d'une façon assez proche de celle utilisée dans les manuscrits décrits plus haut. Le latin, avec l'italien, est placé dans une colonne à gauche de la page, occupant moins d'un quart de la feuille; les entrées et les équivalents sont écrits sur une, deux ou trois lignes superposées; on y fait recours à des abréviations si le texte est trop long. Le chinois à droite occupe le reste de la page, avec la transcription phonétique et tonale au-dessous de chaque caractère; points et marqueurs de division interviennent entre les exemples. Un examen ultérieur de l'original serait nécessaire pour le confirmer, mais ce qu'on observe semble être une association de types occidentaux et de matrices de bois, correspondant à la ligne ou à un segment de la ligne, sur lesquelles ont été gravés les caractères chinois. Le tout a été composé dans une grille qui divise chaque entrée et reconstitue la disposition des contenus d'une façon proche à celle du manuscrit. Au verso de la deuxième page est écrit « Premières épreuves du dictionnaire malchanceux Latin, Italien, Chinois, à cause de l'iniquité et de la trahison du graveur Filippo Zenoby » (*Prime prove dell'//infausto dizionario // Latino=Italiano+Cinese // Per l'iniquità e tradimento // di Filippo Zenoby incisore//*).

<sup>38</sup>Un volume est « Vat. Est Or 4 », décrit plus haut. Mais l'exemplaire que Castorano souhaitait imprimer est un autre : « Est Or 5 A »; ce volume, avant d'entrer à la Bibliothèque Vaticane, aurait appartenu à la Propaganda. C'est probablement la même chose pour les quelques pages imprimées d'un volume incomplet, répertorié sous la côte Est Or 5 B. Castorano aurait réalisé quatre copies manuscrites du dictionnaire, et une autre copie aurait été faite par un tiers (CIVEZZA 1879 : 95).

Filippo Zenobi est le descendant d'une famille d'imprimeurs et de graveurs, originaire de Macerata (dans les Marches) : il est légitime de se demander si la région d'origine, la même que Castorano, peut avoir influencé ce dernier dans le choix d'un collaborateur de cette famille spécialisée dans la profession depuis plusieurs générations. Filippo est le fils d'Antonio (graveur du bois, mort avant 1729), qui était un de trois fils de Gaetano Zenobi (1663 ?–1726). Gaetano, le grand-père de Filippo, est le petit-fils de Carlo, déjà dans la profession à Macerata ; il a commencé son activité en tant que *stampatore pubblico*, dans la ville de Foligono, près de Pérouse (Ombrie) et il s'est ensuite déplacé à Rome, où il semble être présent au moins dès 1691 ou 1693 (MARINELLI s. d. : 3). Gaetano travaille pour le pape Clément X (1700-1721) comme imprimeur et graveur, et il sera aussi au service de deux papes successifs, Innocent XIII (1721-1724) et Benoît XIII (1724-1730). Les générations suivantes de la famille Zenobi restent dans la profession, mais les descendants semblent avoir moins de succès. En plus des publications pour l'Église, ils travaillent pour des éditions théâtrales romaines : Filippo Zenobi apparaît comme éditeur installé à la place Capranica, près du théâtre Capranica (SARTORI 2002 : 139–140), dans les mêmes années où Castorano lui aurait consigné son ouvrage plurilingue. Il est proche de Girolamo Mainardi (1679 ?-1763) qui, depuis 1726, a obtenu la gestion de la « Stamperia Camerale » à côté du théâtre.<sup>39</sup> Filippo Zenobi travaille en tant que graveur sur bois pour certaines publications de Mainardi, et il est encore en activité dans les années 1750.<sup>40</sup> Il est tout à fait vraisemblable que Castorano l'ait contacté pour graver les caractères chinois sur bois. Un autre franciscain, Joseph Maria da Evora (1690-1752), aurait initialement aidé pour réunir les fonds nécessaires à cette entreprise (FRANCESCHI 2017 : 9, CIVEZZA 1879 : 96), mais sa nomination à Porto l'éloigna finalement de son confrère et de ces projets éditoriaux (LI 2017 : 179). Dans les Archives de la Propagande se trouve un mémorial adressé au Pape par Castorano, en date du 10 janvier 1735, juste après son retour à Rome.<sup>41</sup> Le missionnaire explique que, pendant son séjour en Chine, il s'est donné beaucoup de mal afin d'apprendre les caractères

<sup>39</sup> À propos de Mainardi, éditeur bien connu qui aurait même relevé la gestion du théâtre pendant quelques mois entre 1728 et 1729, voir la notice de FRANCHI (2006) en ligne dans le *Dizionario biografico italiano* (dernier accès le 2023-04-08) [https://www.treccani.it/enciclopedia/girolamo-mainardi\\_%28Dizionario-Biografico%29](https://www.treccani.it/enciclopedia/girolamo-mainardi_%28Dizionario-Biografico%29).

<sup>40</sup> Dans le catalogue OPAC SBN ('Catalogue collectif des bibliothèques du Service national des bibliothèques' d'Italie), on compte une dizaine de volumes au nom de Filippo Zenobi, publiés pour la plupart vers 1734, puis dans les années 1750 ; signalons aussi un *Processionario* de 1754, publié par Mainardi, présentant un frontispice illustré avec les noms de F. Zenobi et C. Grandi (en ligne : "Online ecclesiastical heritage", <https://www.beweb.chiesacattolica.it/subeweb/>, dernier accès le 08/04/2023).

<sup>41</sup> Congregazione "de Propagande Fide", Archivio Storico, Sc. Stamperia, vol. 2 (1721–1770), ff. 109–110.

chinois qui sont extrêmement difficiles (*...affatigato ad imparare li caratteri di quella Nazione, quali sono molti, e difficili*). Il présente ensuite son dictionnaire trilingue : il ne contient pas seulement les mots chinois en romanisation, mais aussi des caractères chinois, et ce qui est utile dans la langue courante ; il recense toutes les locutions et les phrases nécessaires pour rendre la langue plus accessible ; il est accompagné d'une grammaire. Castorano explique avoir compilé ce dictionnaire pour les missionnaires destinés à se rendre en Chine, mais aussi afin qu'il soit mis au service des marchands actifs à l'international et des personnes qui étudient les langues étrangères et les 'choses curieuses' (*...per comodità e utilità non solo dei nuovi Padri Missionari che andaranno a imphiegarsi nelle Sacre Missioni della Cina, ma anche dei Mercanti Europei che trafficano in quel Regno, e di chiunque si diletta di lingue e studii forastieri e curiosi*). Cette compilation lui a pris beaucoup de temps et de travail, et Castorano supplie Sa Sainteté de donner l'ordre d'imprimer immédiatement son dictionnaire, car il veut s'occuper de ce travail pendant qu'il se trouve encore à Rome (*...per il tempo che restarà in Roma, potria dirigere la stampa al modo cinese...*). Il explique qu'il sera trop difficile de publier le dictionnaire pour les personnes qui ne connaissent pas la langue chinoise (*...per essere molto difficile tale ditionario à caratteri cinesi stampati da chi ne sa poco, e molto più da chi non ne sa niente...*), point sur lequel nous ne pouvons que lui donner raison : Castorano connaît le chinois et il est conscient de la difficulté de cette entreprise éditoriale. Effectivement, vu le nombre extrêmement élevé des caractères chinois, ce n'est pas seulement l'impression qui pose des difficultés à des non-sinophones, mais la gestion et le rangement de l'ensemble des pièces, comme on le sait pour le cas français.<sup>42</sup> Dans le même texte il est encore question des mérites que le représentant de l'église tirera d'une telle entreprise nécessaire aux activités de prosélytisme. Enfin, la conclusion nous apprend que l'impression du dictionnaire original de Brollo a été interrompue parce que ce dictionnaire contenait trop d'erreurs ; Castorano suggère donc d'utiliser sa propre copie (corrigée) du volume de Brollo à cette fin.<sup>43</sup>

<sup>42</sup>Avec des milliers de caractères, chaque pièce mal rangée correspond à un caractère perdu définitivement ; c'est pour cette raison qu'en France, pour les « buis de régent », est théorisée au XVIIIe siècle une gestion rigoureuse. Il faut éviter que les caractères soient divisés ou déplacés inutilement par des personnes qui, ne connaissant pas le chinois, ne pourraient pas les remettre à leur place (DE GUIGNES 1790 : 77–78)

<sup>43</sup>Il s'agit probablement, pour l'original, du manuscrit « BAV EO 2 » du dictionnaire de Brollo ; quant à la copie du dictionnaire de Brollo par Castorano, il devrait s'agir de « BAV EO 8 ». À propos des tentatives inachevées d'imprimer les dictionnaires de Brollo, pour lesquelles fut vraisemblablement sollicité Matteo Ripa (1682-1746) dans les années immédiatement précédentes au retour de Castorano en Italie et sans que cela aboutisse, voir BUSSOTTI (2015 : 376–378).

Un autre texte conservé à la Propaganda est un mémorial de l'année suivante (1736) au pape Clément XII.<sup>44</sup> On comprend qu'aucune version de dictionnaire chinois n'a été finalement retenue pour l'impression.<sup>45</sup> Castorano présente des excuses pour les modifications qu'il a apportées au dictionnaire de Brollo et qui ont été mal accueillies. Mais il se justifie par le fait qu'il n'existait pas de copie originale en Europe et que toutes les versions du dictionnaire étaient discordantes. Pour ses corrections, il s'est basé donc sur sa propre copie du dictionnaire de Brollo en faisant appel à ses connaissances,<sup>46</sup> tout en signalant ses éventuels ajouts ; ces amendements lui semblaient d'autant plus nécessaires que l'impression serait coûteuse, et il était donc inopportun de proposer un texte erroné. Dans cet écrit, le franciscain demande finalement à être libéré de cette tâche ingrate et il mentionne Mainardi comme l'imprimeur chargé par le pontife de publier le dictionnaire.<sup>47</sup> On peut imaginer que, dans ce contexte, Castorano soit entré facilement en contact avec les Zenobi : non seulement Filippo est voisin de Mainardi vers 1730, mais il travaille parfois pour lui dans les années qui suivent.

Enfin, un fragment du journal de Castorano nous apprend qu'en 1742, il présenta au Pape deux livres « à imprimer pour le bien public », à savoir son dictionnaire trilingue et la *Parva elocubratio* comprenant le catalogue de la bibliothèque de Giovanni Francesco Nicolai, décrite plus haut (§ 3.1). Castorano explique qu'il était déjà en contact avec la *Propaganda fide* pour ces impressions, mais qu'il aurait aimé recevoir des fonds du pontife pour ce projet. Mais ce dernier ne donna pas une réponse positive, en ordonnant même à Castorano de reprendre ces livres que, par ailleurs, il connaissait déjà.<sup>48</sup>

---

<sup>44</sup>Cf. Congregazione de Propagande Fide, Archivio Storico, Sc. Stamperia, vol. 2 (1721–1770), ff. 111–113.

<sup>45</sup>En plus du dictionnaire trilingue (BAV EO 5) et de la version de Castorano du dictionnaire Chinois-Latin de Brollo (BAV EO 8), dans la lettre est mentionné le dictionnaire de Brollo transcrit en Chine par Telli (probablement BAV EO 2) auquel serait allée la préférence de Mainardi « imprimeur caméral », c'est-à-dire gérant la Stamperia Camerale pour le Pape : « *il Signor Mainardi stampatore camerale diede memoriale a vostra Santità di voler stampare non essi, ma quello addotto in mezzo dal signor Filippo Telli.* »

<sup>46</sup>Il écrit : « *onde non non essendovi in Roma, né in Europa, l'originale del dizionario del R. P. Gemona e gli altri esemplari trascritti da mano in mano, essendo con molti errori, nè concordando l'uno con l'altro, l'oratore stimo bene dovere correggerlo e secondo quello trascritto da sé in Cina (nel 1706) e secondo la scienza è notizia che esso ha della lingua e caratteri cinesi.* »

<sup>47</sup>Voir les notes n. 39 et 45.

<sup>48</sup>DI FIORI (1989 : 450), qui cite le « Document 26 », Fragment du journal de Carlo Orazi da Castorano 1735 -1742, Bibliothèque nationale de Naples, *Manoscritti Castorano seng.* XI B 73.

## 5 Conclusion

Le *Dictionarium Latino Italico Sinicum* est un dictionnaire monumental, d'un grand intérêt non seulement en raison du nombre de ses entrées, mais aussi des exemples illustratifs proposés, ainsi que de la présence de mots, expressions et collocations relevant de divers registres. Dans le processus de rédaction, l'auteur s'est appuyé sur deux des dictionnaires européens les plus répandus (les ouvrages de Calderino et Calepino), mais aussi sur les dictionnaires de chinois écrits en langues européennes par ses confrères (notamment par Basilio Brollo), ainsi que sur la lexicographie sino-mandchoue. Le dictionnaire trilingue de Castorano se situe au carrefour des traditions lexicographiques occidentales et chinoises, tout en ayant également recours à des informateurs.

Comme le souligne LI (2017 : 176), la richesse des entrées, les nombreux exemples et la présence de mots de la langue parlée contribuent à la valeur de ce dictionnaire pour la reconstitution du chinois mandarin de l'époque, ainsi que dans l'histoire de l'apprentissage du chinois.

L'ouvrage de Carlo da Castorano est représentatif de l'apogée de la lexicographie missionnaire italophone en Chine, mais aussi de sa dernière étape. Comme le rappelle Castorina : « avec Brollo et Castorano, la période dorée des études italiennes sur la Chine se clôt<sup>49</sup> » (CASTORINA 2019 : 79 ; notre traduction). Mais les tentatives infructueuses de Castorano pour faire financer l'impression témoignent aussi des difficultés liées au contexte social, économique et politique (RAINI 2017, CASTORINA 2019), en particulier de la difficulté à imprimer des ouvrages en chinois en Europe, sans un soutien extérieur important, comme ce fut le cas à Paris pour Fourmont et ses « buis du régent ». Utilisés en petite partie pour les œuvres de Fourmont, ces types de bois devront attendre Napoléon et l'Empire pour servir encore une fois à l'édition du dictionnaire Chinois-Latin-Français de 1813, un projet politique et de prestige (cf. BUSSOTTI & LANDRY-DERON 2020). La papauté, soucieuse des questions doctrinales et rituelles, tournée vers d'autres terres de mission que la Chine, ne développe pas un vrai programme éditorial concernant le chinois et ceci malgré le fait que les missionnaires, hommes de terrain, auraient été les premiers occidentaux à détenir toutes les compétences pour envisager en Europe un projet d'édition et d'impression du chinois viable dans le temps.

---

<sup>49</sup>Texte original italien : « Con Brollo e Castorano si chiude così il periodo d'oro degli studi italiani sulla Cina ».

## Références

- ABEL-RÉMUSAT, Jean-Pierre. 1814. *Plan d'un dictionnaire chinois, avec des notices de plusieurs dictionnaires chinois manuscrits...* Paris : Pillet.
- ARCODIA, Giorgio & Bianca BASCIANO. 2021. *Chinese Linguistics : An Introduction*. Oxford : Oxford University Press.
- BOTTÉRO, Françoise. 2019. Ancient China. In *The Cambridge World History of Lexicography*, 51-6. Cambridge University Press.
- BROCKEY, Liam M. 2007. *Journey to the East : The Jesuit Mission to China, 1579-1724*. Cambridge, MA : Harvard University Press.
- BUSSOTTI, Michela. 2015. Du dictionnaire chinois-latin de Basilio Brollo aux lexiques pour le marché : deux siècles d'édition du chinois en Italie et en France. *T'oung-Pao* 101(4). 363-406.
- BUSSOTTI, Michela & Isabelle LANDRY-DERON. 2020. Printing Chinese Characters, Engraving Chinese Types : Wooden Chinese Movable Type at the Imprimerie Nationale (1715-1819). *East Asian Publishing and Society* 10. 1-72.
- CALDERINO, Cesare. 1586. *Dictionarium D. Caesaris Calderini Mirani Veronensis, tum latini, tum italici sermonis studiosis, apprimè congruens*. Venezia : apud Felicem Valgrisiium.
- CALEPINO, Ambrogio. 1513. *Ambrosius Calepinus bergomates professor deuotissimus ordinis Eremitarum sancti Augustini*. Venezia : Alessandro Paganini per Leonhard Alantsee.
- CASTORINA, Miriam. 2019. Parole, parole, parole... Il contributo della sinologia italiana alla compilazione di dizionari di lingua cinese » [Des mots, des mots, des mots... La contribution de la sinologie italienne à l'élaboration des dictionnaires du chinois]. *LEA* 8. 73-86.
- CATTO, Michela. 2013. *Orazi, Antonio*. [https://www.treccani.it/enciclopedia/antonio-orazi\\_%28Dizionario-Biografico%29/](https://www.treccani.it/enciclopedia/antonio-orazi_%28Dizionario-Biografico%29/).
- LUÓ CHÁNGPÉI 羅常培. 1930. Yēsūhuìshì zài yīnyùnxué shàng de gòngxiàn 耶穌會士在音韻學上的貢獻 [La contribution des jésuites à la phonologie]. *Zhōngyāng yánjiùyuàn lìshǐ yǔyán yánjiūsuo jíkān* 中央研究院历史语言研究所集刊 [Bulletin de l'Institut d'histoire et de philologie de l'Academia Sinica] 1(3). 267-388.
- CIVEZZA, Marcellino. 1879. *Saggio di bibliographia geografica, storica, etnografica sanfrancescana*. Prato : Ranieri Guasti.
- COBLIN, Weldon South. 2006. *Francisco Varo's Glossary of the Mandarin Language*, t. 1. Sankt Augustin : Monumenta Serica Institute.
- DE GUIGNES, Joseph. 1790. *Principes de composition typographique pour diriger un compositeur dans l'usage des caractères orientaux de l'imprimerie royale*. Paris : Imprimerie Royale.

- DI FIORI, Giacomo. 1989. *La legazione Mezzabarba in Cina (1720-1721)*. Milano : Istituto Universitario Orientale.
- ELMAN, Benjamin. 2000. *A Cultural History of Civil Examination in Late Imperial China*. Berkeley : University of California.
- FRANCESCHI, Maurizio. 2017. Dopo tanti viaggi, fatiche, patimenti, pericoli ecc., anco vivo, o utcumque sano mi rivedo in casa Paterna. In Isabella Doniselli ERAMO (éd.), *Carlo da Castorano. Un sinologo francescano tra Roma e Pechino [Carlo da Castorano. Un sinologue franciscain entre Rome et Pékin]*, 1-13. Milan : Luni.
- FRANCHI, Saverio. 2006. *Mainardi. Girolamo*. [https://www.treccani.it/enciclopedia/girolamo-mainardi\\_%28Dizionario-Biografico%29/](https://www.treccani.it/enciclopedia/girolamo-mainardi_%28Dizionario-Biografico%29/).
- KLÖTER, Henning. 2011. *The Language of the Sangleys : A Chinese Vernacular in Missionary Sources of the Seventeenth Century*. Leyde : Brill.
- KLÖTER, Henning. 2019. China from c. 1700. In John CONSIDINE (éd.), *The Cambridge World History of Lexicography*, 317-339. Cambridge : Cambridge University Press.
- LACKNER, Michael. 2001. Circumnavigating the Unfamiliar : Dao'an (314-385) and Yan Fu (1852-1921) on Western Grammar. In Michael LACKNER, Iwo AMELUNG & Joachim KURTZ (éd.), *New Terms for New Ideas. Western Knowledge and Lexical Change in Late Imperial China*, 357-372. Leyde : Brill.
- LEUNG, Cecile. 2002. *Étienne Fourmont (1683-1745). Oriental and Chinese Languages in Eighteenth-Century France*. Leuven : Leuven University Press.
- LI, Hui. 2017. Il Dictionarium Latino-Italico-Sinicum di Carlo Orazi Da Castorano [Le Dictionarium Latino-Italico-Sinicum de Carlo Orazi Da Castorano]. In Isabella Doniselli ERAMO (éd.), *Carlo da Castorano. Un sinologo francescano tra Roma e Pechino [Carlo da Castorano. Un sinologue franciscain entre Rome et Pékin]*, 175-190. Milan : Luni.
- MARINELLI, Bruno. s. d. *Gaetano Zenobi, stampatore e intagliatore tra Foligno e Roma (sec. XVII-XVIII)*. [https://www.centrostudifrezzi.it/magazzino/docs/Gaetano\\_Zenobi.pdf](https://www.centrostudifrezzi.it/magazzino/docs/Gaetano_Zenobi.pdf).
- MASINI, Federico. 2006. Manoscritti sulla lingua cinese presenti negli archivi e nelle biblioteche italiani [Manuscrits sur la langue chinoise présents dans les archives et bibliothèques italiennes]. In G BOCCALI & M SCARPARI (éd.), *Scrittura e codici nelle culture dell'Asia : Giappone, Cina, Tibet, India. Prospettive di studio [Ecritures et codes dans les cultures d'Asie : Japon, Chine, Tibet, Inde. Perspectives d'études]*, 233-244. Venice : Caforscarina.
- XU MINGLONG 許明龍. 2004. *Huang Jialüe yu zaoqi Faguo hanxue 黃嘉略與早期法國漢學*. reprint, Shanghai : Shangwu yinshuguan, 2014. Beijing : Zhonghua shuju.

- MUNGELLO, David. 1994. *The Chinese Rites Controversy : Its History and Meaning*. Nettetal : Steyler Verlag.
- ORAZI, Carlo. 1759. *Brevissima notizia o relazione di varj viaggi, fatiche, patimenti, opere ec. Nell'Imperio della Cina, ec. Del reverendo padre f. Carlo Horatii da Castorano*. Livorno : per gli eredi Santini.
- ORAZI, Carlo. s. d. *Dictionarium Latino Italicum Sinicum Tam vocum, quam Litterarum seu Characterum usualium Sinensium ad usum et commoditatem PP. Missionariorum in hanc Sinicam Missionem noviter adventantium*. Ms. « Vat.estr.or.4 » (1732). Rome : Vatican Apostolic Library. [https://digi.vatlib.it/view/MSS\\_Vat.estr.or.4](https://digi.vatlib.it/view/MSS_Vat.estr.or.4).
- PATERNICÒ, Luisa. 2013. *When the Europeans Began to Study Chinese. Martino Martini's Grammatica Linguae Sinensis*. Leuven : F. Verbiest Institute.
- PELLIN, Tommaso. 2009. *Lessico grammaticale in Cina (1859-1924)*. Milan : Franco Angeli.
- PEVERELLI, Peter. 1986. *The History of Modern Chinese Grammar Studies*. University of Leiden. (thèse de doct.).
- PEYRAUBE, Alain. 2011. Les langues sinitiques. In A PEYRAUBE (éd.), *Dictionnaire des langues*, 979-983. Paris : PUF.
- POLMONARI, Simonetta. 2008. *Padre Basilio Brollo da Gemona in dialogo con la cultura cinese*. Vicenza : LIEF.
- RAINI, Emanuele. 2010. *Sistemi di romanizzazione del cinese mandarino nei secoli xvi-xviii [Systèmes de romanisation du chinois mandarin aux XVIe et XVIIe siècles]*. Rome : Sapienza Università di Roma. (thèse de doct.).
- RAINI, Emanuele. 2017. A preliminary analysis of the romanization in the 'Vocabulario da letra china'. In Pieter ACKERMAN (éd.), *Chinese missionary linguistics*, 109-136. Leuven : Ferdinand Verbiest Institute.
- SARTORI, Orietta. 2002. *Le impressioni sceniche. Dizionario bio-bibliografico degli editori e stampatori romani e laziali di testi drammatici e libretti per musica dal 1579 al 1800, volume II : Integrazioni, aggiunte, tavole, indici*. Rome : Edizioni di storia e letteratura.
- SIEBER, Patricia. 2013. The Imprint of imprints : Sojourners, xiaoshuo translations, and the transcultural canon of early Chinese fiction in Europe, 1697-1826. *East Asian Publishing and Society* 3. 31-70.
- SÖDERBLOM SAARELA, Mårten. 2019. The Chinese Periphery to c. 1800. In *The Cambridge World History of Lexicography*, 202-222. Cambridge : Cambridge University Press.
- STAFUTTI, Stefania. 1981. Note Sul Manoscritto inedito Brevis Narratio Itineris Ex Italia Usque Ad Chinam ... Di Carlo O. Da Castorano (1673-1755). *Cina* 17. 35-79.

- VEDAL, Nathan. 2019. China, c. 600–c. 1700. In J CONSIDINE (éd.), *The Cambridge World History of Lexicography*, 109-129. Cambridge : Cambridge UNiversity Press.
- VON COLLANI, Claudia. 2016. Carlo Orazi da Castorano and the Jesuits in the Chinese Rites Controversy. *Antonianum* XCI(2). 479-510.
- YANG, Huiling. 2014. The Making of the First Chinese-English Dictionary. *Historiographia Linguistica* 41(2). 299-322.
- YANG, Paul Fu-mien. 2001. The Portuguese-Chinese Dictionary of Michele Ruggieri and Matteo Ricci : An Historical and Linguistic Introduction. In J CONSIDINE (éd.), *Dicionário Português-Chinês*, 169-217. Lisbon : Instituto Português do Oriente, Bibliothèque nationale du Portugal.
- YU, Dong Clara. 2017. Gli scritti di Carlo da Castorano nella Biblioteca Apostolica Vaticana [Les écrits de Carlo da Castorano dans la Bibliothèque Apostolique Vaticane]. In Isabella Doniselli ERAMO (éd.), *Carlo da Castorano Un sinologo francescano tra Roma e Pechino [Carlo da Castorano. Un sinologue franciscain entre Rome et Pékin]*. Milan : Luni.
- ZWARTJES, Otto. 2019. Missionary Traditions in East Asia. In J. CONSIDINE (éd.), *The Cambridge World History of Lexicography*, 614-633. Cambridge : Cambridge University Press.
- ZWARTJES, Otto. 2021. The Jesuits' contribution to the study, documentation, and teaching of "exotic" languages. In Robert A. MARYKS (éd.), *Jesuit historiography online*. Leiden : Brill. DOI : 10.1163/2468-7723\_jho\_COM\_225017.



## Chapter 16

# The bilingual dictionary as a mediator between West and East: The beginnings of English-Polish lexicography

Mirosława Podhajecka

University of Opole

This article looks at the beginnings of the English-Polish lexicographical tradition by focusing on the earliest dictionary for both languages compiled in 1799 and extant in manuscript form. It is something of a curiosity, inasmuch as English enjoyed very little recognition among speakers of Polish at the end of the eighteenth century. French and German, along with classical languages, would at that time have been the only foreign languages taught in the school system. We may assume that the dictionary was made by a bilingual author, but, since borrowing from predecessors was a widespread phenomenon in lexicography, other options also need to be considered. Indeed, a comparative analysis of the manuscript indicates that the compiler took into account an edition of Abel Boyer's English-French dictionary published originally in 1699 (Boyer 1780); Nathan Bailey's 1721 monolingual dictionary expanded into an English-French-Latin-German work in 1771; and two English-French grammars. The area of overlap is traced through the spelling of the headwords, the choice of phraseology and illustrative citations, and French and German items added to the entry when Polish equivalents were apparently difficult to provide. The findings are illustrated on the basis of examples.

## 1 Introduction

This paper is based on the premise that there is no better example of a linguistic mediator between West and East than the bilingual dictionary. The reason is simple: by describing the vocabulary of two different languages, the bilingual



dictionary juxtaposes two different worlds (Podhajecka 2016a: 19). The divide between West and East has been pervasive in a variety of discourses. While it may be drawn roughly along geographical lines, it has admittedly been more of an ideological dichotomy between the West, with its presumed superiority, and the East, with its presumed inferiority (see, e.g., Spencer & Wollman 2002: 98). This ideological facet has been ignored here, allowing me to focus on linguistic and cultural perspectives.

In what follows, I look at the beginnings of the English-Polish lexicographical tradition by focusing on the earliest dictionary extant in manuscript form, *Do Słownika Angielsko-Polskiego Dodatek ktorego układanie przedsięwzięte zostało Roku 1799 Dnia 17 Czerwca* [An addendum to an English-Polish dictionary whose compilation was undertaken on 17 June 1799] (Anonymous n.d.).<sup>1</sup> From the European perspective, English represents a Western and Polish an Eastern language. Printed reference works pairing both tongues appeared a few decades later: the first monoscopal dictionary, Krystyn Lach-Szyrma's *Słownik angielsko-polski ...* [An English-Polish dictionary ...], was published in 1828 (Lach-Szyrma 1828) and the first biscopal one, Erazm Rykaczewski's *A complete dictionary English and Polish and Polish and English ...* (Rykaczewski 1849–1851), in the mid-nineteenth century.

## 2 The historical context

Let me begin with brief historical explanations to put the analysis into the right context. Between the sixteenth and eighteenth centuries, Polish was the main language of the Great Polish-Lithuanian Commonwealth,<sup>2</sup> one of Europe's largest and most populous countries. The turbulent period that followed, however, resulted in its loss of independence. Annexed by Russia, Prussia, and Austria-Hungary in the wake of three subsequent partitions, Poland was erased from the world's map for 123 years and, as a consequence, Polish became a language without a nation state. English, by contrast, was the official language of the vast British Empire, whose effective administration, robust parliamentary system, and economic power, among other things, were held in high esteem. It is,

---

<sup>1</sup>The manuscript was discovered by Siekierska (1985: 134–136) in the Vernadsky National Library of Ukraine in Kiev. Two other unpublished dictionaries were compiled in the mid-nineteenth century: Paweł Sobolewski's in 1840 (Podhajecka 2016b) and Michał Wiszniewski's around 1854 (Jajdelski 2000).

<sup>2</sup>More exactly, Polish was the superstratum, whilst Ruthenian and Lithuanian were the substrata (Walczak 2018: 280–281).

therefore, unsurprising that there were attempts to achieve close political cooperation with Britain, a relationship that would, in the fullness of time, lead to the restoration of Poland, but the plans eventually fell through.<sup>3</sup>

Both before and after the partitions, foreign languages were regarded as an integral part of the education of young Poles, so French, German, and, to a lesser extent, Italian were taught alongside classical languages. French, in particular, became the common language of the Polish nobility; what France represented was perceived as the height of civilisation. As Brandes (1903: 31) observed,

French is the auxiliary language of the Poles, the language which among the higher classes all know perfectly (...) the language which is spoken as fluently as the mother-tongue and even better than the Russians speak it. In aristocratic circles Poles frequently converse with one another in French, a state of things which from the beginning of the century was promoted not only by the continual intellectual intercourse with France, and the emigration thereto.

Interest in English, formerly treated as a “barbarous tongue lacking in refinement” (Cieśla 1974: 121), emerged falteringly at the end of the eighteenth century. It may be worthwhile noting that, in the Age of Reason, English was not only the language of Shakespeare and Milton, but also that of British inventors and men of science. As Julian Antonowicz (1788), the author of the first grammar of English for Polish students, so eloquently expressed it,

The glory of being the native Country of Creative Genius’s is granted to Great Britain by all Europe. It was this happy Kingdom that has brought forth the famous Bacons and Newtons, to whom, it seems, nature has opened its closet (...). Then it is pity, that this Language is not still so universally known, as it deserves to be, above all, in the Kingdom to which the blessing of liberty is common with Great Britain.

Native speakers of English in Poland, a country soon to fall under foreign rule, were a rarity. John Lind (1737–1781), the governor of the Cadet School in Warsaw established by King Stanisław August, and Joseph Saunders (1778–1836), an engraver affiliated with the University of Wilno, are among the few acknowledged

---

<sup>3</sup>The aristocratic Czartoryski family attempted to introduce a set of reforms modelled on Britain and advocated a Polish-British alliance. Prince Adam Kazimierz Czartoryski’s British experiences earned him the nickname “the perfect Englishman”, but it was his son Adam Jerzy who became an ardent anglophile. He studied in London, travelled throughout Britain, and was fluent in English (Lipoński 1978: 18–19, 28; cf. Konopczyński 1947).

in the literature. No wonder the early history of English-Polish language contact relied, out of necessity, on autodidacts such as Antonowicz (c. 1750--1824), a Basilian monk and teacher in Włodzimierz (in the Wołyń province, now in Ukraine).<sup>4</sup> Benjamin Haustein (1778–1836), a professor of German and English at the University of Wilno (now in Lithuania),<sup>5</sup> as well as Gaspar Friese and a Hennig, who taught English to senior students of Toruń's academic gymnasium (Łukaszewicz 1849: 408), might have been eligible had they not been Germans.

The schools in Warsaw, Włodzimierz, and Toruń were not the only ones to promote the teaching of English. The Wołyń Gymnasium in Krzemieniec, initiated and headed by Tadeusz Czacki, was another. Staffed with distinguished teachers, it was the best Polish educational facility with ambitious aims: to educate students in line with the highest teaching standards in order to give them “a solid knowledge and comprehension of problems as well as [to] stimulate independent thinking” (Buczek 2016a: 3). Teaching languages was a distinct priority; the school's curriculum featured Polish and five foreign languages, i.e., French, German, Russian, Latin, and Greek. Lessons in English, recommended to Czacki by Antonowicz, were introduced in 1807 (Cieśla 1974: 125, Buczek 2016b: 154). Originally an optional subject, English was taught by the Englishmen George Forster and William Rose, both of whom were brought from abroad specifically for that purpose,<sup>6</sup> and Józef Mikulski, a teacher of Latin and Polish (Słowikowski 1863: 344).<sup>7</sup>

There seems to be a link between the Wołyń Gymnasium and the draft dictionary. After the collapse of the November Uprising (1830–1831), an armed rebellion against the Russian Empire, the school was closed and its entire library of approximately 34,000 volumes was transported to Kiev (Danilewicz-Zielińska 1992: 82).<sup>8</sup> Buczek (2016b: 18) affirms that part of the school's book collection rescued from the wars has now been deposited in the National Library of Ukraine. Since

---

<sup>4</sup>Butterwick (1998: 175) states that, at the request of King Stanisław August, Antonowicz offered his English grammar to Charles Whitworth, the British Envoy Extraordinary and Minister Plenipotentiary at the Polish court.

<sup>5</sup>The Romantic poet Juliusz Słowacki, a graduate of the university, noted that Haustein's English was so awkward that “even native speakers would be unable to recognize what language he spoke” [M.P.] (Piechota 2016: 96).

<sup>6</sup>According to Buczek (2016b: 154), Forster spoke neither French nor Polish. It is thus unclear how he taught English with the so-called Grammar-Translation Method, as has been suggested by Piotrowski (2005: 214).

<sup>7</sup>Jack William MacDonald, a Scot (Nicoll 1923: 38), was a private teacher in Krzemieniec, but there is no evidence that he ever worked at the school.

<sup>8</sup>The school's library included 15,580 volumes from the royal book collection purchased by Czacki in 1803 after the death of King Stanisław August (Buczek 2016a: 8–9).

the manuscript comes from the holdings of that library, one might assume that it belonged to the Wołyń Gymnasium, having either been compiled in Krzemieniec or, perhaps more likely, donated to this centre of Polishness as a cultural artefact.

### 3 The manuscript dictionary

#### 3.1 The compiler

The identity of the compiler remains undisclosed. The same handwriting used throughout, clearly not Antonowicz's, suggests a single author, no doubt a man. This hypothesis stems from two factors: firstly, schooling was restricted to boys and, secondly, lexicography was a predominantly male business at that time.<sup>9</sup> The problem is that none of the Poles familiar with English either by dint of travel to Britain, including Krystyn Lach-Szyrma (1790–1866), Karol Sienkiewicz (1793–1860), and Michał Wiszniewski (1794–1865), or extensive self-study, including Hipolit Błotnicki (1792–1886), Józef Korzeniowski (1797–1863), and Konstanty Piotrowski (1790–1863), was born sufficiently early to undertake the project.<sup>10</sup> Jacek Idzi Przybylski (1756–1819), a professor, writer, and translator, was an exception, but his linguistic knowledge was far too sophisticated to be able to consider him the author. The manuscript thus appears to have been a spontaneous act of creativity on the part of an enthusiast of English who never achieved distinction.<sup>11</sup>

#### 3.2 The form

The manuscript consists of white sheets of paper ruled by hand and bound into a book with the word *Słownik* [*Dictionary*] printed on the spine.<sup>12</sup> Of the book's 404 pages, around 280 are filled, rather haphazardly, with text. Although the

<sup>9</sup>Hester Lynch Piozzi, the author of *British Synonyms ...* (1794) and “hostess, muse, amanuensis, intimate friend” of Samuel Johnson (Berglund 2010: 69), was one of very few recognised female lexicographers in the eighteenth century.

<sup>10</sup>The latter category included authors who translated English literature (e.g., Shakespeare, Milton, Byron, Goldsmith, and Stern) into Polish.

<sup>11</sup>At the same time, increasing numbers of Poles of high social rank such as Michał Mniszech (1748–1806) and Franciszek Bohusz (1746–1820) embarked on journeys abroad that included visits to Britain. While their peregrinations had intellectual purposes, other Poles such as Tadeusz Burzyński (1730–1773) and Franciszek Bukaty (1747–1797) remained in Britain for political and diplomatic reasons (see, e.g., Libiszowska 1972).

<sup>12</sup>The binding is of half-leather, which means that the sides are covered with board paper, but the spine and corners are in leather.

order is alphabetical, ascertaining the chronological sequence in which the dictionary was compiled is a sheer impossibility. There is every indication that, after writing a batch of entries, the compiler would return to his sources only to find more headwords that deserved inclusion, and he would add them between other letters of the alphabet. Headwords from B, for instance, are given at the beginning, in the middle, and even at the very end of the book. Here is a small sample of how the headwords were placed in the subsequent letters of the alphabet:

A 4, 5-7, 8-29, 32-43, 288, 317

B 4, 7, 43-61, 220-224, 289, 318-322, 397-404

C 7, 8-12, 15-16, 62-66, 69-89, 290, 323, 404

D 89-101, 291, 326-328

E 101-115, 117-119, 292, 329

As the title evinces, the material was an addition to another volume (see Figure 1). If we look at the title page more carefully, however, we will discern that the draft dictionary was originally titled *Słownik Angielsko-Polski ... [An English-Polish dictionary ...]*, but the title was later modified by using a pen with black ink. The same ink was applied for the ornament attached below, which brings to mind the engravings on the title pages of Renaissance dictionaries such as Calepino's.<sup>13</sup> The manuscript has neither the front matter, explaining the purposes of the project and its envisaged readership, nor the back matter.

Each page is divided into two columns with extensive margins left for the notes. Sometimes they include lexicographical information such as a missing headword (e.g., *Laden* 'ołowiany'; *Swan* 'łabędź'; *Ungodliness* 'niezbożność'), at other times providing what seem to be references to the sources consulted (e.g., I, 120, 134). From page 317, the prospective headwords were to be written in three columns.

The last pages (381-404) are titled "Przydatek do dykcjonarza" [An addendum to the dictionary]. This part, including single words, phrases, and citations, is arranged in no alphabetical order. Whereas the first 460 lines (entries?) are numbered, the other 64 numbers are provided in a way that cannot be explained logically. Generally speaking, a great deal of internal diversity, partly illegible handwriting, and numerous inconsistencies make the draft dictionary a challenge to research.

---

<sup>13</sup>With over 200 successive editions printed across Europe, Ambrogio Calepino's dictionary was one of the best-selling endeavours in the history of lexicography.

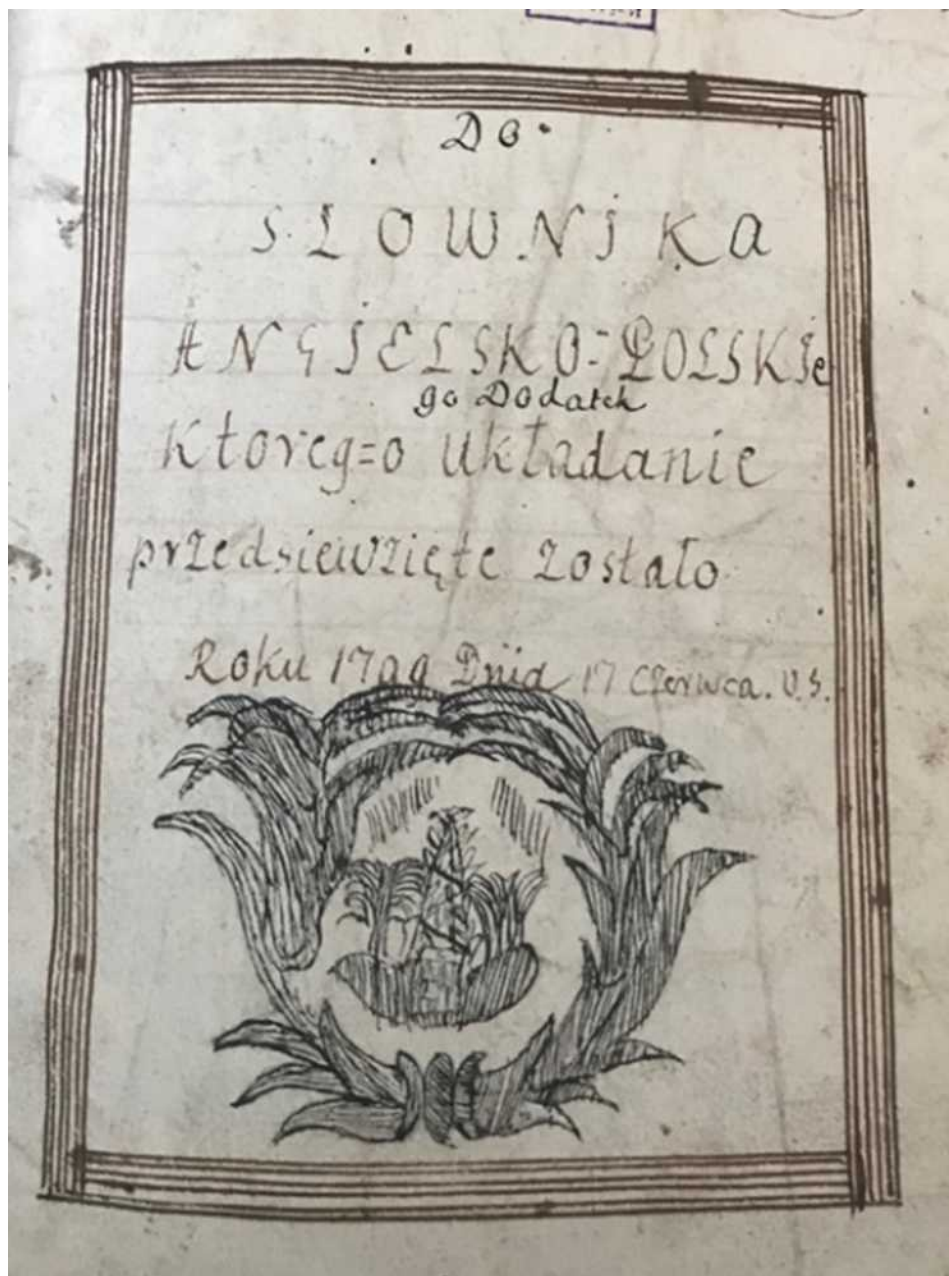


Figure 1: The title page of the manuscript dictionary

The material quoted below has been kept in the original orthography. If French or Polish equivalents were recorded without diacritics, for example, this is how they appear here. I also stick to the compiler's inconsistent spelling with upper and lower case.

### 3.3 The contents

What did it take to compile a bilingual dictionary of English and Polish at the end of the eighteenth century? One might venture to claim that fluency in both languages was a prerequisite, but it was not an overriding requirement in the past (Podhajecka 2016a: 554). Many English-Polish dictionaries, including Lach-Szyrma's and Rykaczewski's, were made by authors whose knowledge of English was imperfect, but who spoke other foreign languages, mainly French and German. It is little wonder, then, that the compiler treated both languages as a conduit between English and Polish.

My analysis was carried out on Google Books, a multilingual diachronic archive.<sup>14</sup> The research methodology was twofold: firstly, I searched for single words, sequences of words, and longer stretches of text recorded in the manuscript dictionary in order to identify the compiler's sources and, secondly, the rest of the searches were conducted specifically in the sources. The second part of the study was exceedingly labour-intensive.

The results suggest that at least two reference works were drawn on: an edition of Abel Boyer's *The royal dictionary ...* (1792?), considered "one of the most important and innovative works of all time" (Cormier & Fernandez 2005: 479–480);<sup>15</sup> and Nathan Bailey's great folio dictionary, originally monolingual, expanded into an English-German-French-Latin dictionary by Anton E. Klausing and retitled *A compleat English dictionary: oder vollständiges englisch-deutsches Wörterbuch ...* (Bailey & Klausing 1771).<sup>16</sup> Boyer's *The complete French master ...* (Boyer 1762), a grammar-cum-phrasebook-cum-dictionary,<sup>17</sup> and Guy Miede and Boyer's *Grammaire angloise-francoise ...* (Miede & Boyer 1763), also came into

---

<sup>14</sup>In doing so, I adjusted Google Books functionalities to my research needs. I restricted the time frame to specific periods (for example, 1700–1800) and searched for whole constructions (e.g., "she is a little out of order"), also with wildcards (e.g., "she \* a little out of order"), or various Boolean expressions (e.g., "she" "out of order"). Google Books is not a diachronic corpus (see, e.g., Podhajecka 2011), but, with 40 million scanned books, it is currently the largest resource available for historical scrutiny.

<sup>15</sup>The original edition of Boyer's dictionary was published in 1699.

<sup>16</sup>The dictionary must have been convenient to the compiler, who was undoubtedly familiar with German, French, and Latin, though possibly in different degrees.

<sup>17</sup>It is worthy of mention that *The complete French master ...* was one of Antonowicz's sources.

play. There is a clear rationale behind such a choice of materials: two different dictionaries allowed the compiler to decide, firstly, what headwords should be selected and, secondly, how to explain the source language (SL) meanings in the target language (TL).<sup>18</sup> The grammars were, among other things, a handy source of citations extracted from the sections titled “Familiar phrases”.<sup>19</sup>

A few words should be said about the compiler’s working practice. The headwords in upper case are recorded alphabetically. Every now and then, we encounter ostensible errors, but they reflect the type of arrangement employed in the sources. *Alehouse*, for instance, is followed by *Alecost*, but this is because *Alehouse* was clustered in the entry for *Ale* in both dictionary sources, whilst *Alecost* was accorded main headword status. Under the letter C, by contrast, *Camber*, *Càmbrick*, and *Càmber* were copied verbatim from Bailey and Klausning’s work together with their German equivalents.

Different parts of speech are treated as different headwords. A proportion of nouns are followed by a bracketed article and verbs by the bracketed particle *to*, e.g., *Artichoke* [: the] ‘karczoch’; *Pour* [: to:] ‘nalewać, wylewać’; and *Shepherdess* [a] ‘owczarka’. Irregular word forms are also recorded (e.g., *Begun* ‘part. od słowa *to Begin*’; *Dead* ‘imiesłow od tego słowa *Die*’; *Kept* ‘praet y part: od sł: *keep*’; *Lighted* ‘part y praet: od sł: *light*’; *Swolln* ‘Imiesł. od *Swell*’; and *Wrung* ‘praet y part: od sł: *wring* wykręcać’). Apart from that, grammatical information is practically non-existent.

In the first letters of the alphabet, the compiler provides variant spellings (e.g., *Antelope* ob. *Antilope*; *Arsevercy* or *Arsiversy*; *Bandalier* ob. *Bandaleer*; and *Esq* zamiast *Esquire* ‘szlachcic’), all of which come from Boyer’s or Bailey and Klausning’s works.<sup>20</sup> In the remaining part of the manuscript, when it dawned on him how time-consuming the compilation would be, fewer variant forms were incorporated.<sup>21</sup>

For the word list, the author drew from two bulky dictionaries, from which he copied some headwords and ignored others; many of his choices seem debatable.

<sup>18</sup>In the light of the above, the SL refers to English, French, and German, so it would be more accurate to speak of three SLs.

<sup>19</sup>Siekierska (1985) sensibly, albeit wrongly, assumed that the compiler had drawn on Kunstmann’s *Kurzes Wörterbuch in deutscher, französischer, italienischer, englischer und polnischer Sprache ...* (1794).

<sup>20</sup>One also encounters dead references. In the case of *Basin*, the user was told to look up *Bason*, but it is not the word list.

<sup>21</sup>See, however, abbreviated forms such as *Dr.* ‘zamiast *Doctor* nauczyciel’; *Excy* ‘zamiast *Excellency*’; *Knt* ‘zamiast *knight* kawaler’; *Lap* ‘zamiast *Ladiship*’; *O’er* ‘over nad’; *T’other* ‘zamiast *the other*’; and *Won’t* ‘zamiast *will not*’. Where they come from remains undetermined due to formal inconsistencies.

Under A, for instance, he admitted *Albacor* ‘see Dolphin’,<sup>22</sup> *Alburin*, *Alburn* ‘dark colour’, in its old-fashioned spellings (cf. OED3’s *auburn*); and *Alet* ‘an Indian partridge’ [M.P.],<sup>23</sup> at the same time leaving out more common words such as *Alarm-ing* ‘causing terror’, *Alchymy* ‘a secret part of chymistry’, and *Alcove* ‘a recess in a room’.<sup>24</sup> His motives may only be speculated upon, but the manuscript dictionary was probably intended as an aid in reading English literature, so, from this angle, the status of common and rare words was comparable. Here are a handful of other peripheral items or senses: *Agramatist* ‘nieumiejętny, nieuczony’; *Ban-stickle* ‘nazwisko pewney małej ryby’; *Bargh-master* ‘dozorca kopalni’; *Belagged* ‘zostawiony w tyle, zmoczony’; *Fillamot* ‘feuille morte, naz.[wa] koloru’; *Loll* ‘zgarbić się’; *Scullion* ‘kuchta’; *Squint-ey’d* ‘kosooki’; *Tight* ‘tęgi’; and *Unchew’d* ‘niepożuty’.

The manuscript includes some proper names, primarily place names (e.g., *Constantinopol* ‘Konstantynopol’, *Edinburgh* ‘Edinburg’, *Muscow* ‘Moskwa’, *Prague* ‘Praga’, *Stock-holm* ‘Sztokholm’, *Warsaw* ‘Warszawa’); names of regions (e.g., *Brandenburg* ‘Brandenburgia’, *Piedmont* ‘Pedemontana’, *Tuscany* ‘Toskania’); and names of countries (e.g., *Denmark* ‘Dania’, *Hungary* ‘Węgry’, *Norway* ‘Norwegia’, *Persia* ‘Persya’, *Scotland* ‘Szkocya’, *Turkey* ‘Turcyja’). Proper names were recorded in Bailey and Klausing’s trilingual “Catalogue of the most usual proper, Christian and nicknames of men and women ...”, located in the back matter, as well as in both grammars.

A number of headwords in the dictionary such as *Man* ‘męszczyzna’, *Manhood* ‘męski wiek’, and *Mankind* ‘Rodzaj ludzki’ are formally and semantically related. Once the compiler established the meaning of the main headword, meanings of derivatives or compounds should be easy to guess or find, but this is not always the case. This refers to, among others, *Despair* ‘rospacz’ vs. *To desper* ‘desesperer’; *Foot* ‘noga, stopa, miara’ vs. *Foot-stool* ‘un marche pied’; *Herb* ‘ziele’ vs. *Herb-woman* ‘herbiere’; *Juice* ‘sok’ vs. *Juicy* ‘succulente’; *Kitchin* ‘kuchnia’ vs. *Kitchin Tackling* ‘batterie de cuisine’; and *Wine* ‘wino’ vs. *Wine-press* ‘un pressoir’. Curiously, despite some variation in form, all the missing Polish equivalents may be found in the then contemporary reference works: ‘przywieść kogo do rospaczy’ (*Desesperer* → *To desper*); ‘podnożek’ (*Marche-pied* → *Foot-stool*); ‘zielenniczka,

<sup>22</sup>OED3 (Simpson & Proffitt 2000) defines *albacore* as ‘any of several small and medium-sized tunas ...’.

<sup>23</sup>In his *English dictionary ...* (1676), the British lexicographer Elisha Coles (1677) defines *Alet* as ‘The true Peru-falcon’. This word has not been incorporated in OED3 (Simpson & Proffitt 2000).

<sup>24</sup>The headwords are defined with English glosses added to Boyer’s dictionary, except for *Alet*, which stands for Bailey and Klausing’s German equivalent in my translation.

co różne ziele przedaie, ziołniczka, zieleniarka' (*Kräuterfrau* → *Herb-woman*);<sup>25</sup> 'soczysty, soku wiele w sobie maiący' (*Succulent* → *Juicy*); 'statki kuchenne miedziane i mosiężne ...' (*Batterie* → *Kitchin Tackling*); 'and 'prasa winna do tłoczenia wina y tym podobnych rzeczy' (*Pressoir* → *Wine-press*).<sup>26</sup>

The compiler's decisions are occasionally baffling. *Translate* 'przetłumaczyć' and *Translation* 'przetłumaczenie', for instance, are followed by *Translator* 'un savetier szwiec który naprawia stare obuwie' [a cobbler who repairs old shoes].<sup>27</sup> The reason behind neglecting the primary sense of the word, 'one who renders anything into a different language',<sup>28</sup> alluding to the compiler's temporary status, is unknown. There are also headwords with no Polish counterparts at all, of which the entry for *Break* is a good example (e.g., *To break a horse* 'ein Pferd bändigen'; *To break company* 'die Gesellschaft verlassen'; *To break silence* 'anfängen zu reden'; *To break a jest* 'scherzen'; and *To break out into tears* 'häufige Thränen vergiessen').

Clues as to which dictionary was perused should be sought in the selection of the phrases and/or their foreign equivalents. In the letter B, the compiler included *Back of the hand* 'Dłoń wierzchnia', *Back Bone* 'Pacierzowa kość', *Back of the Chimney* 'contre-coeur de cheminée, plague de fen', *Back-stairs* 'un Page, ou un Garçon de la chambre du lit',<sup>29</sup> and *Back Gammon* 'toutes tables'. As has turned out, only Boyer has all the English phrases and identical French explanations, with the exception of *Back-stairs*, the description of which comes from one of the grammars. Sometimes the wording is harder to trace to a source, which may be exemplified by *Balkers* → *Conders* 'Ci ktorzy z wysokiego mieysca ukazują Rybakom, gdzie są śledzie'. In this case, the syntactic patterns of French ('Ceux qui étant postés sur une hauteur montrent aux bateaux de péchceurs par où les harengs passent') and German ('Die von einem hohen Ort den Fischern zeigen, um welche Gegend die Heringe sind') could be translated into Polish in largely the same ways.

For most dictionaries, it would be challenging to give a precise count of the headwords, inasmuch as no single method of calculating them has gained universal acceptance among practical lexicographers and metalexicographers (see, e.g., Landau 2001: 109–112). It is even more puzzling in the case of a manuscript

<sup>25</sup>None of these have survived in Polish to date, being ousted by 'zielarka'.

<sup>26</sup>The equivalents quoted above come from a dictionary by Danet & Kola (1745) and two different dictionaries by Trotz (1764, 1772).

<sup>27</sup>The same definition is paired with the headword *Cobler*.

<sup>28</sup>This English gloss comes from Boyer's 1792 edition.

<sup>29</sup>In another part of the manuscript, *Back-stairs* is paired with Bailey and Klausning's German equivalent ('eine heimliche Treppe').

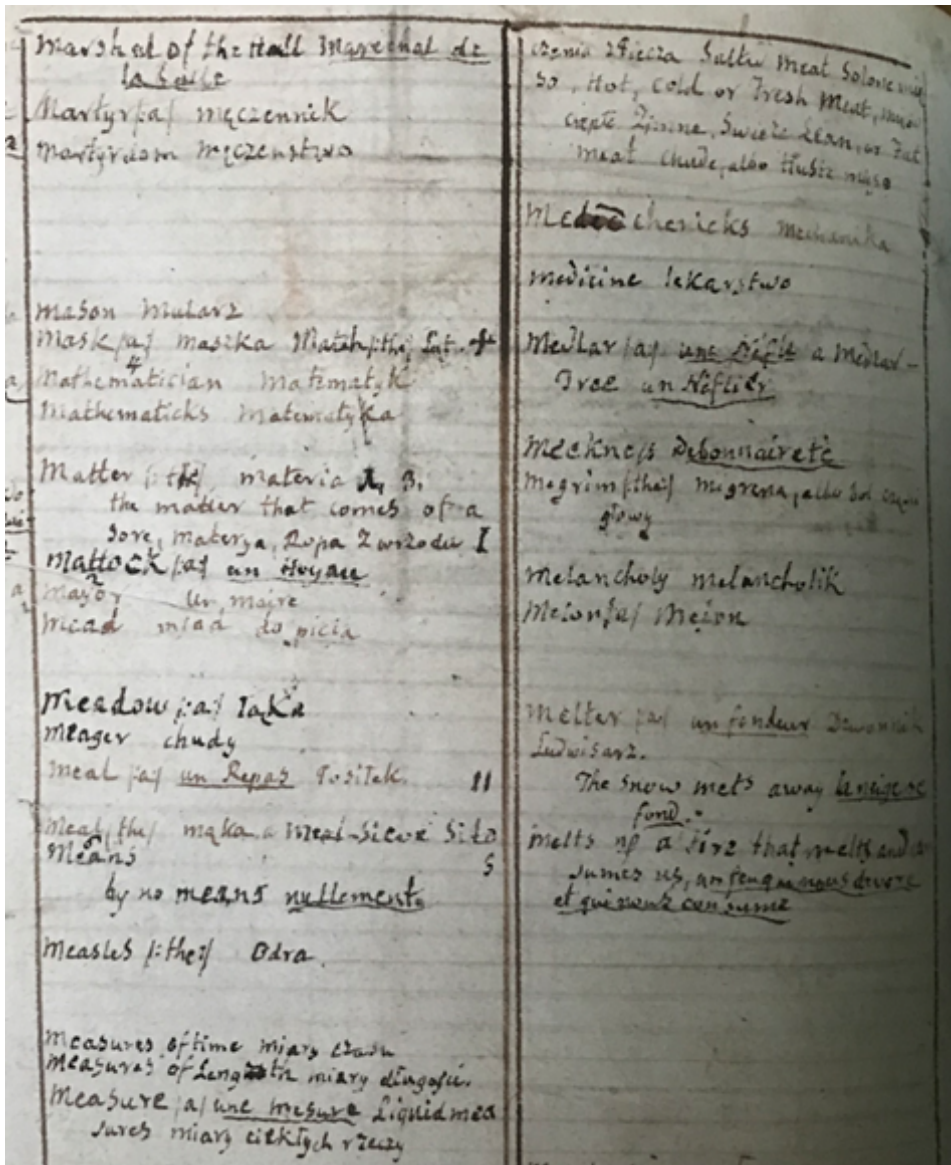


Figure 2: Part of a sample page of the manuscript dictionary

wordbook like this, where each line could be perceived as one entry or a component of a broader entry. For this reason alone, I made no attempt to count the number of potential headwords and, likewise, decided to eschew any reference to *macrostructure* and *microstructure*,<sup>30</sup> two standard lexicographical terms, because the borderline between these two is blurred (see Figure 2).

### 3.4 Seeking equivalence

Although the nature of equivalence between English and Polish remains unaddressed, the compiler must have taken a somewhat naïve approach to it by assuming that the lexical systems of different languages describe the world in exactly the same ways.<sup>31</sup> After all, a plethora of dictionaries which paired the vocabulary of two or more languages might attest that translatability is an undisputed fact. In any case, working with French and German was all the easier that German-Polish and French-Polish dictionaries had a long tradition in Poland. Frączek (1999: 358) lists ten German-Polish or Polish-German dictionaries issued in the eighteenth century, as well as those produced between the sixteenth and seventeenth centuries, whereas Jakubczyk (2016) analyses five lesser known French-Polish or Polish-French volumes, published alongside Pierre Danet and Dymitr Franciszek Kola's *Nouveau grand dictionnaire ... francois, latin & polonois* (1743–45) and other multilingual works of the eighteenth century.

The manuscript is lexicographically unsophisticated; the English headword is usually paired with one single Polish equivalent (e.g., *Law* 'Prawo'; *Perjury* 'krzywoprzysiędztwo'; *Spectacles* 'okulary'; *Triennial* 'trzyletni'). Two or more alternatives appear less frequently, as in the case of *Agree* 'zezwoić, godzić się'; *Barefacedly* 'iawnie, bez wstydu'; *Dragoon* 'dragon, żołnierz'; *Endeavour* 'starać się, usiłować'; *Pudding* 'kiszka, jątrznica'; *Sorrow* 'boleść, nieukontentowanie, smutek'; *Tutor* 'Guwerner, Nauczyciel'; and *Verily* 'w samey rzeczy, prawdziwie'.

A great many English headwords are polysemous, but the compiler rarely differentiates between senses. One suitable example is the entry for *Baggage* 'Bagaż, podróżne rzeczy, Nierządnicza wojskowa', which recreates in Polish two different senses of Bailey and Klausinger's headword ('das Reisegeräth, Wanderung; eine Soldatenhure'); another is the entry for *Baste* 'Polewać piecyste,

<sup>30</sup>Hartmann & James (2001: xii, 94) treat *macrostructure* as "the overall access format of a dictionary", usually alphabetical organisation, whereas *microstructure* refers to the internal design of the entry.

<sup>31</sup>This, as we know today, is entirely wrong. Different languages describe extralinguistic reality in quite diverse ways, which Zgusta (1971: 295) terms *anisomorphism of languages*. No wonder it has been argued that "perfect interlingual equivalence is an exception rather than the rule" (Adamska-Sałaciak 2020: 201).

co zle szyc, bić' ('gebratenes am Spiess begiessen; etwas Schlecht nähren; abschmieren, abprügeln'); and still another is the entry for *Deny* 'zapierać, przeczyć, odmówić' ('läugnen, verneinen; abschlagen, verweigern; absagen, verlügen').<sup>32</sup>

The need to negotiate meaning by making assumptions about what the lexicographers meant and correlating them with factual knowledge was an essential part of the compilation process.<sup>33</sup> This was intellectually demanding because the French and German glosses were often dissimilar. *He looks as big as Bull-Beef*, for example, was rendered into French as 'Il a le regard extremement fier' [He looks extremely proud] and into German as 'Er siehet aus, als ob er mich fressen wollte' [He looks as if he wanted to eat me].<sup>34</sup> It was only to be anticipated that mediating between the SLs would result in misconceptions (e.g., *Bed-rid* 'łożkiem bawiący się'; *Cellar* 'sklep na trunki'; *Charity* 'miłość'; *Dimple* 'Dolek pod nozdrzami nad wargami';<sup>35</sup> *Foot-men* 'stołeczek do ściągania butów'; *Humoursome* 'dziwak, chymeryk'; *Poem* 'poema rym'; *Speech* 'Parole słowo'; *Tempest* 'Niepogoda'; *Toast* 'pieczyste'; *Womb* 'żywot niewieści'; and *Ye* 'zamiast the').

The Polish part covers single-word equivalents (e.g., *Experience* 'doświadczenie'), descriptive equivalents (*Outworks* 'Dzieła zewnętrzne fortyfikacyi'), and translated citations (*Is there any of you that will undertake it?* 'Czy iest kto z was który by chciał to przedsięwziąć?'). Upon closer analysis, three categories of lexicographical treatment may be distinguished:

- Polish constructions documented in historical texts, including bilingual dictionaries
- Polish word-for-word translations from French or German unattested elsewhere
- missing Polish counterparts for citations, culture-bound terms, and other types of vocabulary

---

<sup>32</sup>Importantly, different senses in Bailey and Klausning's dictionary are separated with semicolons.

<sup>33</sup>In foreign-language teaching, meaning negotiation has been perceived as an interactive process, particularly in a language classroom (see, e.g., Macaro et al. 2015: 17). In the historical perspective, however, a student working with a foreign text was often bound to establish meaning solely on the basis of his or her dictionary.

<sup>34</sup>OED3 (Simpson & Proffitt 2000) explains that, in such examples of usage as *to bluster like bull-beef* and *as big as bull-beef*, *bull-beef* is a term of abuse.

<sup>35</sup>Cf. Bailey and Klausning's 'ein Grübchen im Kinn oder Backen' and Trotz's 'Dolek iako w brodzie, śmiechow y policzkach' (1744).

These categories will be discussed one by one and illustrated with examples.

The fact that many Polish constructions are documented in historical sources may be indicative of the compiler's native speaker competence, but they are more likely to derive from French-Polish or German-Polish dictionaries. As already remarked, French and German were taught extensively to Polish students, so a number of bilingual and multilingual resources were available on the market. Danet and Kola's French-Latin-Polish dictionary (Danet & Kola 1743, 1745) and Trotz's French-German-Polish dictionary (Trotz 1747) were the largest and the best (see, e.g., Kuryłowicz 2021), but there were also less comprehensive works such as Eberlein's German-Polish (Eberlein 1763), Trotz's German-Polish (Trotz 1772), and Kaliszewski's French-Polish dictionaries (Kaliszewski 1795). It may be safely assumed that the compiler used them in his foreign language education.

Let us look at a few examples, beginning with *Ankle* 'kostka u nogi'. Bailey and Klausning's definitions 'la cheville de pied'/'der Knöchel am Fuss' have been traced to Danet and Kola's 'kostka u nogi ...' (Danet & Kola 1743) and Trotz's 'kostka, nogi część ...' (Trotz 1744), so we could infer that the former served as the compiler's direct source. The verb phrase 'mieć płynienie krwi z nosa' (*Bleed at the nose*) was probably borrowed from Trotz's 1772 dictionary, whereby it is the first of five stylistic variants ('krwi płynienie z nosa, istic krwi z nosa, ciorkanie krwi z nosa, gdy kroplemi cieczy kroplami,<sup>36</sup> cieczenie krwi z nosa'). *Falling-Sickness* is paired with 'kaduk, wielka choroba'. In his *Royal dictionary*, Boyer offers two equivalents: 'le haut mal, le mal caduc', both of which, unsurprisingly, lead us to Danet and Kola's 'kaduk, wielka choroba'. The same procedure was followed for *Fob* 'kieszonka w spodniach', first looked up in Boyer ('gousset') and then in a French-Polish dictionary, this time Kaliszewski's. More examples are exhibited in Table 1. The symbol ≠ signifies a marked difference between the structure of the SL and TL items.

Whenever literal translation would produce constructions atypical of, or entirely alien to, the compiler's mother tongue, the Polish equivalents diverged from the SL patterns. 'Pieczenia nerkowa cielęca' for *Loyn of Veal* ('la langue de veau') and 'przegub ręki między dłonią y łokciem' for *Wrist* ('das Gelenke an der Hand') are suitable examples. The linguistic tradition was also a factor. Taking *Cousin* into consideration, it has been customary for Poles to refer to a male cousin in a twofold way as 'brat stryjeczny', i.e., the son of one's father's sibling, and 'brat cioteczny', i.e., the son of one's mother's sibling.<sup>37</sup> As for *Sixpence*

<sup>36</sup>This wording is erroneous; it should have been replaced by the gloss 'gdy krew cieczy kroplami'.

<sup>37</sup>Female cousins are called 'siostra stryjeczna' and 'siostra cioteczna'.

Table 1: Polish constructions documented in historical texts.

| The headword        | Background sources                                                                      | The manuscript                                      |
|---------------------|-----------------------------------------------------------------------------------------|-----------------------------------------------------|
| Ablepsy             | eine Blindheit der Augen oder des Gemüths                                               | Ślepotą oczów albo umysłu                           |
| Bantling            | (kleines Kind) ein Bankert, ein Kind, welches vor der Hochzeit erzenget worden          | Małe dziecię przed ślubem urodzone                  |
| To Blister Campeche | appliquer des vesicatoires lignum Indicum sic dictum                                    | wezykatoria przykładać drzewo Indyjskie tak nazwane |
| Cousin              | cousin, cousine                                                                         | ≠ Stryeczny cioteczny brat, krewny                  |
| Dizziness Embryo    | vertige, tournement de tête eine unzeitige Geburth, unvollkommene Frucht in Mutterleibe | Zawrót głowy<br>Płód w żywocie macierzyńskim        |
| Epiphany            | le jour <i>ou</i> la fête des Rois                                                      | Święto trzech królów                                |
| Fish-day            | jour maigre                                                                             | Postny dzień                                        |
| Loyn of Veal        | la langue de veau                                                                       | ≠ Pieczenia nerkowa cielęca                         |
| Make mouths         | faire la mouë                                                                           | Wykrzywiać się wargami komu                         |
| Parlour             | der Sprachsaal                                                                          | izba gościnna do rozmowy                            |
| Peace, peace there  | paix, paix-là                                                                           | ≠ nakazując milczenie                               |
| Sixpence            | ein Sechs Pence, oder halber Schilling                                                  | ≠ pieniędzy Gatunek                                 |
| Unbloody victory    | une Victoire obtenue sans effusion de sang                                              | Zwycięstwo bez wylania krwi                         |
| Wrist               | das Gelenke an der Hand                                                                 | ≠ Przegub ręki między dłonią y łokciem              |

‘Pieniądzy Gatunek’, culture-bound terms tend to be ambiguous to the target user; hence, a descriptive equivalent.<sup>38</sup> Although the idiomatic expression *Make mouths* was rendered word for word (‘Wykrzywiać się wargami komu’), the compiler should not be reproached for this, as Danet and Kola’s definition of *Moue* was as follows: ‘Wystawienie wargi na kogo wydrzeźniając mu się wykrzywianie się na kogo’ (1745). ‘Nakazując milczenie’ (*Peace, peace there*) looks more like an

<sup>38</sup>Cf. Penny ‘un sou gatunek pieniędzy’. In his *Anglia i Szkocja: przypomnienia z podróży roku 1823–24 odbytey* (Lach-Szyrma 1828), Lach-Szyrma preferred to borrow *penny* (*pense*) and *shilling* (*szyling*) rather than describe them.

explanatory gloss than a functional equivalent. The descriptive equivalent for *Campeche* suggests that Latin might have been another SL.

The second category sometimes sees the compiler go astray by wording the Polish equivalents in a way unattested in eighteenth-century materials. For example, *Alectromancy* denotes 'divination using a cock'. Bailey and Klausing's 'das Wahrsagen durch einen Hahn' was translated word for word into 'wrózenie przez koguta', but the preposition was incorrect; *wrózenie z koguta* (Osiński 1806: 117) sounded more natural.<sup>39</sup> 'Pasztet z węgorzy' (*An eel pie*), probably a rendition of Boyer's 'un pâté d'anguilles', is the earliest occurrence in Polish, but it is Lach-Szyrma's 'pasztet z węgorza' that became the standard form. *Empirick* 'lekarz z doświadczenia' denotes (positively) a doctor by experience,<sup>40</sup> which ought to be treated as a mistake. The traditional Polish equivalent found in eighteenth-century bilingual dictionaries, 'lekarz jarmarkowy', had pejorative connotations and there are reasons for thinking that 'une empirique' was a term of contempt in French, just like 'Quacksalber' and 'Marktschreier' in German. 'Mieysce do zdybania się' (*Rendevous*) must have been a translation of Bailey and Klausing's 'ein bestimmter Ort der Zusammenkunft, Sammelplatz'. The Polish deverbal noun 'zdybanie się' was less outdated than it is today, but 'szadzka' would, nonetheless, be a better option (cf. Trotz's *Schadzka, Zchadzka, Schacka* 'Ort der Zusammenkunft', 1764).

Table 2 includes additional Polish constructions, translated word for word on the basis of the background sources, which are undocumented in historical materials.

In the case of *Ferret*, the original equivalents did not match the Polish gloss, so I used a back-translation, which points to the entry for *Fureter* in a handful of works, including *Table analytique et raisonnée des matieres ...* (Anon 1782). The Polish equivalent in the entry for *Pedlar* was less erudite than Rykaczewski's 'kramarz wędrujący' (1849), whilst that for *Scull* was imprudent (cf. Trotz's 'czaszka głowy', 1764). It may be interesting to note that single-word equivalents such as 'jedwabnik' (*Silk-weaver*) and 'dachownik' (*Tyler*) were attested in a range of eighteenth-century dictionaries, including Danet & Kola (1743), Trotz (1744), and Eberlein (1763). 'Podnauczyciel', recorded in Antonowicz's Polish-English glossary included in his grammar (Antonowicz 1788), was a far more natural equivalent for *Usher*.

<sup>39</sup>Cf. 'wrózenie z ręki' [divination from the palm], 'wrózenie z rzeczy' [divination from things], and 'wrózenie z roślin' [divination from plants].

<sup>40</sup>*Empirick* must have been borrowed from Bailey and Klausing, as Boyer has it in the spelling *Empiric*.

Table 2: Word-for-word Polish translations unattested in texts.

| The headword      | Background sources                                                                         | The manuscript                                      |
|-------------------|--------------------------------------------------------------------------------------------|-----------------------------------------------------|
| An Alder-bed      | ein Ort, wo viele Erlenbäume wachsen                                                       | Miejsce gdzie wiele Drzew olchowych rośnie          |
| Anticor           | eine Geschwulst an der Brust der Pferde                                                    | puchlina na piersiach konia                         |
| Bestead           | se montrer ami de quelqu'un, le servir                                                     | Pokazać się komu przyjacielem, iemu usłużyć         |
| Bilboes           | eine Art der Strafe unter den Seeleuten                                                    | pewny Gatunek kary między Żeglarzami                |
| Calambo           | in wohlriechendes Holz, so aus der Insel Java wächst                                       | Gatunek pachnącego drzewa na wyspie Jawa            |
| Ferret            | ein Iltis, geringe Seide → faire sortir les lapins de leur terrier par le moyen des furets | łasica zwierząt do wypędzania krolikow z iam zdatna |
| Instrument-keeper | garde des instrumens                                                                       | Instrumentów muzycznych dozorca                     |
| Fall of the leaf  | la chute des feuilles, l'automne                                                           | opadanie liści albo Jesień                          |
| Megrim            | migraine, douleur de tête, mal de tête                                                     | Migrena, albo bol części głowy                      |
| Pedlar            | petit mercier, qui porte sa boutique sur soi/einer der sein Krämlin am Halse herum trägt   | torbiarz ten co towary swoje nosząc przedaie        |
| Reassume          | wieder annehmen                                                                            | znowu wziąć co przed się                            |
| Scull             | die Hirnschale, der Schedel                                                                | czaszka na głowie                                   |
| Silk-weaver       | ouvrier en soie                                                                            | Ten co w jedwabiu wyrabia                           |
| Tyler             | un couvreur en tuile                                                                       | ten co nakrywa dachy dachówką                       |
| Usher             | Unterschulmeister                                                                          | Namiestnik mistrza szkoły                           |
| A Venison Pasty   | pâté de venaison                                                                           | pasztet ze zwierzyny                                |

It needs to be explained that the constructions in Table 2 have not been found in their entirety, but components thereof were obviously known and used. 'Kara' [punishment], for instance, has been retrieved in contexts wholly unrelated to 'żeglarze' [sailors]; none of the uses of 'instrumenty muzyczne' [musical instruments] pertains to 'dozorca' [keeper, guard]; and nowhere are 'opadanie liści'

[the falling of leaves] and 'jesień' [autumn] put side by side.<sup>41</sup> Burdened with the task of expressing foreign concepts, especially when precise and compact Polish equivalents escaped him, the compiler was torn between the SL and TL patterns. In fact, this would be in line with modern research revealing that translators tend to introduce peculiarities of the SL into the TL. As Teich (2003: 222) puts it, the SL "shines through" the TL translation.

Let us consider the third category. Every now and then, the draft dictionary includes citations borrowed from one of the grammars; the author was evidently aware that language primarily serves communication. Regrettably, only a handful of contextual uses are translated into Polish (e.g., *Let him be ever so rich*<sup>42</sup> 'niech sobie będzie jakożkolwiek bogatym'), which undermines the quality of the English-Polish lexicographical project.

While some of the French citations in Table 3 might have been a little problematic to render into Polish, others posed no special difficulty, but they, too, were left untranslated. This raises questions of the compiler's proficiency in the SLs and of his method of work. One cannot exclude that, once the French or German counterparts had been put down, a search for the Polish items was to ensue. For some reason, however, this procedure failed to take place.

Other headwords lacking Polish equivalents are culture-bound terms, also taken from the two grammars. What the compiler must soon have realised is that they named realia typical of Britain (e.g., *the Keeper of the Privy Purse*) and France (e.g., 'le Trésorier des Menus'), whilst here equivalence was sought between English and Polish. As literal translation was not always the best strategy, more time and resources were needed to identify fitting Polish terms, so only French items were copied into the manuscript.<sup>43</sup>

Even today, few of the terms in Table 4 could be paired with Polish equivalents, of which *Izba Gmin* ('The House of Commons') and *sekretarz stanu* ('Secretary of State') are the easiest. Whilst *the Keeper of the King's Pictures* might potentially be expressed as 'strażnik obrazów królewskich',<sup>44</sup> other renditions (e.g., 'urzęd-

<sup>41</sup>To arrive at credible results, I also consulted Samuel Bogumił Linde's *Słownik języka polskiego* (Linde 1807–1814), the first fully-fledged dictionary in the history of Polish lexicography.

<sup>42</sup>This citation comes from another source, conceivably Peyton's *La nouvelle grammaire angloise ...* (Peyton 1767) or Haussner's *Phraseologia anglo-germanica ...* (Haussner 1798). In the former, the SL version was 'laissez-le être qu'il soit jamais si riche: quelque riche qu'il soit' and, in the latter, 'wenn er noch so reich wäre'.

<sup>43</sup>*Dean of his Majesty's Chappel* 'Dziekani Kaplicy Królewskiej' and *The Master of the King's Household* 'Marszałek Nadw.[orny] Król.[ewski]' may be viewed as exceptions to the rule.

<sup>44</sup>Google Books has attestations for, among others, 'strażnik lasów królewskich' (*Keeper of the Royal Forests*) and 'strażnik grobów królewskich' (*Keeper of the Royal Tombs*).

Table 3: English citations recorded only in French.

| The headword | English citations                                             | French versions                                                         |
|--------------|---------------------------------------------------------------|-------------------------------------------------------------------------|
| Able         | I have eat so much that I shan't be able to eat my Dinner.    | J'ai tant mangé, que je ne pourrai pas diner.                           |
| Conspicuous  | The English Valour has been Conspicuous on several occasions. | La Valeur angloise s'est fait voir avec éclat dans plusieurs Occasions. |
| Could        | I could not write it.                                         | Je n'ai pu l'écrire.                                                    |
| Favours      | The wedding ring and Favours are already bought.              | On a déjà acheté l'anneau nuptial, et les livrées.                      |
| Hop          | Shall we hop with one leg.                                    | Voulez-vous que nous fautions à cloche pied.                            |
| Order        | She is a little out of order.                                 | Elle est un peu indisposée.                                             |
| Parted       | Methinks 'tis not long since we parted.                       | Il me semble qu'il n'y a pas long-tems que nous nous sommes séparez.    |
| Races        | Let us run Races.                                             | Exerçons-nous à la course.                                              |
| Tattle       | She does nothing but prattle or tattle.                       | Elle ne fait que causer, ou caqueter.                                   |
| View         | Let's view those Fields and Meadows.                          | Promenons notre vûë sur ces champs & sur ces prairies.                  |
| Way          | We are a great way from the River.                            | Nous sommes éloignez de la riviere.                                     |
| Yonder       | I suppose they are those that walk yonder.                    | Je m'imagine que ce sont ceux qui se promonent là-bas.                  |

Table 4: English culture-bound terms recorded only in French.

| The headword | English terms                                                 | French equivalents                                         |
|--------------|---------------------------------------------------------------|------------------------------------------------------------|
| Almoner      | the Lord Almoner                                              | Grand-Aumônier                                             |
| Buttery      | the Gentleman of the great Buttery                            | le chef de la grande depense                               |
| Groom        | the Groom of the stole and first Gentleman of the Bed-Chamber | le premier gentilhomme de la chambre du lit du roi         |
| House        | the House of Commons                                          | la chambre des communes, <i>ou</i> la chambre basse        |
| Master       | Pay Master of his Majesty's Forces                            | Trésorier de l'ordinaire & de l'extraordinaire des Guerres |
| Picture      | the Keeper of the King's Pictures                             | Huissier du Cabinet des Peintures                          |
| Poultry      | the Clerk of the Poultry                                      | Le comptrolleur de l'office de la volaille                 |
| Privy        | a Gentleman of the King's Privy – chamber in ordinary         | Gentilhomme ordinaire de la Chambre Privée du Roi          |
| Seal         | Lord Privy Seal                                               | Garde du Petit-Sceau                                       |
| State        | Secretary of State                                            | Un Secretaire d'Etat                                       |
| Usher        | A Gentleman Usher of the Presence Chamber                     | un Ecuyer de la chambre de présence                        |
| Work         | the Surveyor-General of the works                             | Intendant des Bâtiments                                    |
| Yeoman       | Yeoman of the Pantry                                          | l'Aide de la Paneterie                                     |

nik do spraw drobiu' for *the Clerk of the Poultry*) would look idiosyncratic.<sup>45</sup> The fact that Rykaczewski succeeded in pairing *Groom of the Stole* with 'dozorca garderoby królewskiej', *Gentleman of the Bed-Chamber* with 'kamerjunker', *the House of Commons* with 'Izba niższa Parlamentu', *Pay Master General* with 'jeneralny płatnik', *Keeper of the Privy Seal* with 'podkanclerz',<sup>46</sup> and *Yeoman of the Pantry* with 'gwardzista piechotny' (1849) attests to his being a highly skilled lexicographer.

There are further headwords without Polish counterparts. Explaining the verb *to Bastard* ('überführen, das einer ein Hurenkind ist, ein Hurkind zeugen') must have been a pitfall because even Rykaczewski omitted it. 'Ospa wietrzna' is the modern equivalent of *Chicken-pox*, which was previously termed 'ospice owcze', i.e., a disease affecting sheep.<sup>47</sup> As for the noun *Cabbage*, the prototypical equivalent is self-evident ('kapusta'), but we also come across another nominal meaning, 'das Dicke und harte zu unterst am Hirschgeweihen', and two verbal meanings, 'pommer' and 'voler', which turned out to be more troublesome.<sup>48</sup> This notwithstanding, circumventing lexicographical problems by leaving foreign material untranslated into Polish is hardly surprising if we take into account the scope of the project, which sought to pair two languages that had never been paired before. The fact that the compilation was undertaken by a language learner justifies, at least to some extent, the shortcomings of the manuscript.

The compiler drew on bilingual dictionaries for different language pairs in order to ensure the feasibility of his endeavour, but the vocabulary of the SLs was not made up of single words that could be substituted with their TL equivalents in a neat and tidy way. Countless TL multi-word expressions such as collocations and idioms, moreover, were unlike their SL counterparts in terms of lexical and syntactic structures, and some had no equivalents at all. There is, therefore, no doubt that translatability had then, and still has today, inevitable limits. As Baker (1992: 17) explains,

the nature of language is such that, in the majority of cases, words have 'blurred edges'; their meanings are, to a large extent, negotiable and are only realized in specific contexts. The very notion of 'types of meaning' is theoretically suspect.

<sup>45</sup>Even *James* was juxtaposed with its French version 'Jacques' (cf. *Ralph* 'Rudolf').

<sup>46</sup>Diki (<https://www.diki.pl/>), an online dictionary, proposes the equivalent 'Lord Strażnik Tajnej Pieczęci' for *Keeper of the Privy Seal*, which is richly documented in Google Books (<https://books.google.com>).

<sup>47</sup>Cf. Trotz's *verole volante* 'Windpocken. Ospice wietrzne, luskawe' (1747).

<sup>48</sup>Only the latter was admitted into Rykaczewski's English-Polish volume ('kraść okrawki sukna (o krawcach)').

The bilingual lexicographer's aptitude for expressing the concepts of one culture in the language of another was often taken for granted, but, in reality, few, if any, reference works were compiled from scratch. This is the reason borrowing was notorious in lexicography, even though not always "performed on dictionaries of the same direct lineage" (Francoeur 2010: 137).<sup>49</sup> Suffice it to say that, due to the dominant tradition of teaching foreign languages via the Grammar-Translation Method, literal translation in both the West and the East prevailed in the eighteenth century as much as in previous centuries.

## 4 Conclusions

As demonstrated above, the compiler of the manuscript dictionary faced a number of quandaries, all the more because he was presumably a tyro. His lack of expertise has been exhibited, on the one hand, by leaving a great many words, phrases, and citations untranslated into Polish and, on the other, by failing to employ the full linguistic repertoire typical of an educated native speaker of Polish. Unable to match the English headwords with their Polish counterparts, he used bilingual dictionaries as a bridge in the hope of arriving at helpful patterns in French or in German. The comparative analysis indicates that he followed a straightforward method of gathering data. In other words, he rendered the SL material, often resorting to word-for-word translation, unless his native speaker intuition or bilingual dictionaries with Polish guided him otherwise. Whether he undertook his task with "mixed enthusiasm and apprehension" (Cormier 2010: 133), we are unlikely ever to establish.

Time and money have been crucial factors in any lexicographical project, and this one is no exception. The pressure of time, in particular, may explain why the manuscript was left unfinished: looking up thousands of headwords in the various sources to hand would have been painfully slow and laborious. All in all, relying on equivalents culled from English-French and English-German works, some of which were translated with the help of French-Polish and German-Polish dictionaries, the anonymous compiler showed a great deal of ingenuity. This should be appreciated, but, unlike Lach-Szyrma and Rykaczewski who spent several years in Britain, he was insufficiently acquainted with the English language and culture. Consequently, even if the manuscript had been published, it would have stood no chance of becoming a genuine cross-linguistic and cross-cultural intermediary.

---

<sup>49</sup>Instead of speaking of blatant plagiarism, Dolezal (2023: 97) takes a less critical stance by considering borrowing [a] "time-honored lexicographic practice".

One might wonder at the motives that prompted the author to take on such a taxing project. Browsing through the material, I discovered three citations from Antonowicz's bilingual preface to his grammar, of which the following is worth quoting here: *had I contributed to render English more common to my Countrymen* 'gdybym się przyłożył do uczynienia powszechniejszym Języka angielskiego'. Perhaps one of Antonowicz's "few gentle boys" (1788) was so impressed with English that he decided to contribute, as did his mentor, to familiarising his compatriots with the English vocabulary. The names of three of Antonowicz's students are found in the literature: Alojzy Feliński (1771–1820), a poet, translator, and literary critic; Count Jan Feliks Tarnowski (1777–1842), Czacki's nephew, a historian, bibliophile, and senator of the Kingdom of Poland; and Rev. Antoni Bęczkowski (?–1858), a polyglot, teacher, and rector of the Piarist gymnasium in Międzyrzecze Koreckie (in the Wołyń province). The last of these is known to have encouraged his students to read English literature in the original (Kowalski 1859: 236), which makes him a most suitable candidate.

## References

- Adamska-Sałaciak, Arleta. 2020. Issues in compiling bilingual dictionaries. In *The Bloomsbury handbook of lexicography*, 2nd edn., 193–208. London: Bloomsbury Academic.
- Anon. 1782. *Table analytique et raisonnée des matieres ... du dictionnaire des sciences, des arts et des métiers, et dans son supplement*, vol. 1: A–H. A Paris: chez Panckoucke.
- Anonymous. N.d. *Do Słownika Angielsko-Polskiego Dodatek którego układanie przedsięwzięte zostało Roku 1799 Dnia 17 Czerwca*.
- Antonowicz, Julian. 1788. *Grammatyka dla Polaków chcących się uczyć angielskiego ięzyka...* Warszawa: W Drukarni Nadworney J. K. Mci i P. K. Edu.
- Bailey, Nathan & Anton E. Klausing (eds.). 1771. *A compleat English dictionary: Oder vollständiges englisch-deutsches Wörterbuch...* 4th edition. Leipzig-Züllichau: Wansenhau und Frommanninschen Handlung.
- Baker, Mona. 1992. *In other words: A coursebook on translation*. London: Routledge.
- Berglund, Lisa. 2010. Hester Lynch Piozzi's *British Synonymy* in Imperial France. *Dictionaries: Journal of the Dictionary Society of North America* 31. 69–86.
- Boyer, Abel. 1762. *The complete French master for ladies and gentlemen ...: For the use of His late Highness the Duke of Gloucester*. London: S. & E. Ballard.
- Boyer, Abel. 1780. *Dictionnaire royal françois-anglais et anglois-françois...* Nouvelle édition. A Lyon: chez Jean-Marie Bruyset père et fils.

- Brandes, George. 1903. *Poland: A study of the land, people, and literature*. New York: The Macmillan Company.
- Buczek, Katarzyna. 2016a. From the Royal Castle to school. King Stanisław August's legacy and the Volhynian Gymnasium collections. *Espacio, Tiempo y Educación* 4. 1–17. DOI: 10.14516/ete.2017.004.001.82.
- Buczek, Katarzyna. 2016b. *Zbiory dydaktyczne Gimnazjum i Liceum Wołyńskiego w Krzemieńcu (1805–1833)*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Butterwick, Richard. 1998. *Poland's last king and English culture: Stanisław August Poniatowski, 1732–1798*. Oxford: Clarendon Press.
- Cieśla, Michał. 1974. *Dzieje nauki języków obcych w zarysie: monografia z zakresu historii kultury*. Warszawa: Wydawnictwo Naukowe PWN.
- Coles, Elisha. 1677. *An English dictionary: Explaining the difficult terms that are used in divinity, husbandry, physick, phylosophy, law, navigation, mathematicks, and other arts and sciences ...* London: Peter Parker.
- Cormier, Monique C. 2010. New perspectives on seventeenth- and eighteenth-century European lexicography. *International Journal of Lexicography* 23(2). 133–135.
- Cormier, Monique C. & Heberto Fernandez. 2005. From the *Great French dictionary* (1688) of Guy Miège to the *Royal dictionary* (1699) of Abel Boyer: Tracing inspiration. *International Journal of Lexicography* 18(4). 479–507.
- Danet, Pierre & Dymitr F. Kola. 1743. *Nouveau grand dictionnaire ... françois, latin & polonois*. À Varsovie: dans l'imprimerie Royale de la Republique aux Ecolles Pieuses. <https://jbc.bj.uj.edu.pl/dlibra/doccontent?id=314599>.
- Danet, Pierre & Dymitr F. Kola. 1745. *Nouveau grand dictionnaire ... Francois, latin & polonois, enrichi des mellieurs françons [sic] de parler. Avec des notes de critique & de grammaire sur ces trois langues. Tome second*. À Varsovie: L'imprimerie Royale de la Republique aux Ecolles Pieuses. <https://jbc.bj.uj.edu.pl/dlibra/publication/329080/edition/314602/content>.
- Danilewicz-Zielińska, Maria. 1992. *Próby przywołań: szkice literackie*. Warszawa: Biblioteka „Więzi”.
- Dolezal, Fredric T. 2023. Review: Sarah Ogilvie, ed. 2020. *The Cambridge companion to English dictionaries*. Cambridge: Cambridge University Press, *International Journal of Lexicography* 36(1). 93–101. DOI: 10.1093/ijl/ecac006.
- Eberlein, Christian. 1763. *Entwurf eines polnischen Wörter-buchs ...* Vol. 2. Breslau: Wihelm G. Korn und Gampert. [https://www.google.pl/books/edition/Entwurf\\_eines\\_polnischen\\_W%C3%B6rter\\_Buchs/\\_mdpAAAAcAAJ?hl=pl&gbpv](https://www.google.pl/books/edition/Entwurf_eines_polnischen_W%C3%B6rter_Buchs/_mdpAAAAcAAJ?hl=pl&gbpv).

- Frączek, Agnieszka. 1999. *Zur Geschichte der deutsch-polnischen und polnisch-deutschen Lexikographie (1772–1868)* (Lexicographica. Series Maior 93). Tübingen: Niemeyer.
- Francoeur, Aline. 2010. Fighting Cotgrave with Father Pomey: Guy Miège's recourse to the *Dictionnaire royal augmenté* (1671) in the preparation of his *New dictionary French and English* (1677). *International Journal of Lexicography* 23(2). 137–155.
- Hartmann, R. R. K. & Gregory James. 2001. *Dictionary of lexicography*. London/New York: Routledge.
- Haussner, Friedrich W. 1798. *Phraseologia Anglo-Germanica. Oder: Sammlung von mehr als fünfzigtausend redensarte, aus den besten englischen Schriftstellern gezogen ...* Strassburg: Fr. Georg Levrault.
- Jajdelski, Wojciech. 2000. *Michał Wiszniewski's słownik angielsko-polski (1854): The scholar and the dictionary he made*. Wrocław University. (MA dissertation).
- Jakubczyk, Marcin. 2016. *Leksykografia polsko-francuska XVIII wieku: w perspektywie metaleksykograficznej*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Kaliszewski, Adam Piotr. 1795. *Nomenclator czterech językow. francuzkiego, polskiego, niemieckiego y łacinskiego. Recueil des mots ... Zebranie słow ...* Warszawa: W Drukarni J. K. Mci i Rzeczypospolitey u XX. Scholarum Piarum. <https://polona.pl/item-view/d8a8d51a-ccc1-49bf-ac99-6cdfdedcf7cd?page=0>.
- Konopczyński, Władysław. 1947. Anglia a Polska w XVIII wieku. *Pamiętnik Biblioteki Kórnickiej* 4. 93–129.
- Kowalski, Franciszek. 1859. *Wspomnienia. Pamiętnik Franciszka Kowalskiego*, vol. 1. Kijów: nakładem Leona Idzikowskiego.
- Kuryłowicz, Beata. 2021. Największe dzieła polskiej leksykografii XVIII wieku (*Nowy dykcjonarz* Michała Abrahama Troca oraz *Nowy wielki dykcjonarz* Piotra Daneta i Dymitra Koli) z perspektywy współczesnych badań historyczno-językowych: Nowe wyzwania i potrzeby badawcze. *Białostockie Archiwum Językowe* 21. 97–107. DOI: 10.15290/baj.2021.21.06. <https://crispa.uw.edu.pl/object/files/413708/display/JPEG>.
- Lach-Szyrma, Krystyn. 1828. *Anglia i Szkocja: Przypomnienia z podróży roku 1823–24 odbytey*, vol. 3. Warszawa: w Drukarni A. Gałęzowskiego i Komp.
- Lach-Szyrma, Krystyn. 1828. *Słownik angielsko-polski ułożony... dla użytku młodzieży...* Warszawa: Drukarnia Gałęzowskiego i Komp. <https://crispa.uw.edu.pl/object/files/413708/display/JPEG>.
- Landau, Sidney I. 2001. *Dictionaries: The art and craft of lexicography*. 2nd edn. New York/Cambridge: Cambridge University Press.

- Libiszowska, Zofia. 1972. *Życie polskie w Londynie w XVIII wieku*. Warszawa: Instytut Wydawniczy „PAX”.
- Linde, Samuel B. 1807–1814. *Słownik języka polskiego*. Warszawa: u Autora. <http://kpbk.umk.pl/publication/8173>.
- Lipoiński, Wojciech. 1978. *Polska a Brytania 1801-1830. Próby politycznego i cywilizacyjnego dźwignięcia kraju w oparciu o Wielką Brytanię*. Poznań: UAM.
- Łukaszewicz, Józef. 1849. *Historia szkół w Koronie i w Wielkim Księstwie Litewskim od najdawniejszych czasów aż do roku 1794*, vol. 1. Poznań: nakładem Księgarni Jana Konstantego Żupańskiego.
- Macaro, Ernesto, Suzanne Woore & Robert Graham. 2015. *Improving foreign language teaching: Towards a research-based curriculum and pedagogy*. London: Routledge.
- Miege, Guy & Abel Boyer. 1763. *Grammaire angloise-françoise, par Mrs. Miege et Boyer ...* Paris: chez Briasson & David l’Ainé.
- Nicoll, Josephine. 1923. *Shakespeare in Poland* (Published for the Shakespeare Association by Humphrey Milford). Oxford: Oxford University Press.
- Osiński, Alojzy. 1806. *Słownik mitologiczny, z przyłączeniem obrazo-pismu (jconologia)*, vol. 1: A–D. W Warszawie: w Drukarni Xięży Piarów.
- Peyton, V. J. 1767. *La nouvelle grammaire angloise ou se trouvent développés les vrais principes de cette langue et tout ce qui est nécessaire aux étrangers pour apprendre facilement à parler, lire, & écrire l’anglois*. A Londres: chez L. Nourse.
- Piechota, Marek. 2016. *Poliglotyzm wielkich romantyków polskich (Mickiewicz, Słowacki, Krasiński)*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Piotrowski, Wojciech. 2005. *Życie umysłowe Krzemieńca w latach 1805–1832*. Piotrków Trybunalski: Naukowe Wydawnictwo Piotrkowskie przy Filii Akademii Świętokrzyskiej.
- Podhajecka, Mirosława. 2011. Research in historical lexicography: Can Google Books collection complement traditional corpora? In Stanisław Gózdź-Roszkowski (ed.), *PALC 2009: Explorations across languages and corpora*, 529–546. Frankfurt am Main: Peter Lang.
- Podhajecka, Mirosława. 2016a. *A history of Polish-English/English-Polish bilingual lexicography (1788–1947)*. Opole: Wydawnictwo Uniwersytetu Opolskiego.
- Podhajecka, Mirosława. 2016b. Szkic z dziejów leksykografii dwujęzycznej: Paweł Sobolewski i jego Słownik Angielsko-Polski ... (1840). *Prace Filologiczne* 68. 323–344.
- Rykaczewski, Erazm. 1849–1851. *A complete dictionary English and Polish and Polish and English... / Dokładny słownik polsko-angielski...* Vol. 1–2. Berlin: B. Behr.

- Siekierska, Krystyna. 1985. Kijowskie polonika (rękopisy XVI–XIX wieku w Bibliotece AN USSR w Kijowie). *Przegląd Humanistyczny* 29(1). 129–142.
- Simpson, John & Michael Proffitt (eds.). 2000. *The Oxford English dictionary*. 3rd edn. Oxford: Clarendon Press.
- Słowikowski, Adam. 1863. Wspomnienia szkoły krzemienieckiej, śp. Adama Słowikowskiego. *Biblioteka Warszawska* 2. 341–359.
- Spencer, Philip & Howard Wollman. 2002. *Nationalism: A critical introduction*. London: Sage Publications.
- Teich, Elke. 2003. *Cross-linguistic variation in system and text. A methodology for the investigation of translations and comparable texts*. Berlin: Mouton de Gruyter.
- Trotz, Michał Abraham. 1744. *Nouveau dictionnaire françois, allemand et polonais...* Vol. 1. Leipzig: Gleditsch. <https://jbc.bj.uj.edu.pl/dlibra/doccontent?id=466391>.
- Trotz, Michał Abraham. 1764. *Nouveau dictionnaire polonois, allemande et françois ...* Vol. 3. Leipzig: Jean Frederic Gleditsch. <https://jbc.bj.uj.edu.pl/dlibra/doccontent?id=463227>.
- Trotz, Michał Abraham. 1772. *Vollständiges deutsches und polnisches Wörter-Buch, welsches die gebräuchlichsten Wörter und Redensarten ...* Leipzig: Johann Friedrich Gleditsch. <https://www.dbc.wroc.pl/dlibra/publication/7844/edition>.
- Trotz, Michel A. 1747. *Nouveau dictionnaire françois, allemand et polonois ...* Vol. 2: I–Z. Leipzig: Jean Frederic Gleditsch. <https://jbc.bj.uj.edu.pl/dlibra/doccontent?id=466386>.
- Walczak, Bogdan. 2018. Ruszczyzna: drugi język Rzeczypospolitej. In *Języki ruskie w rozwoju historycznym i kontaktach z polszczyzną*, 277–287. Białystok: Uniwersytet w Białymstoku.
- Zgusta, Ladislav. 1971. *Manual of lexicography*. Berlin/New York: De Gruyter Mouton.

## Chapter 17

# Lexicon Lapponicum Bipartitum....ungarice scriptum: Hungarian aspects of North Saami dictionary writing

Ivett Kelemen

University of Debrecen

The birth of North Saami lexicography in 18th-century Norway is closely related to the work of the Norwegian priest, Knud Leem (1697–1774). He was the first to publish a Danish–North Saami dictionary (*En lappisk Nomenclator*, 1756) to help the language learning of missionaries setting off for Saami territories. Later on, the “father” of the North Saami language continued his lexicographical work with a larger endeavour: he started to write a trilingual (North Saami–Danish–Latin) dictionary. The printing of the first volume of the 1600-page-long *Lexicon Lapponicum* started in 1768, but a year later, the work was suspended for a long period of time, because of unexpected “Hungarian happenings”. It is a well-known fact among Finno-Ugric scholars that during their 1768–1769 astronomical expedition to the Norwegian Island of Vardø, Maximilian Hell (1720–1792) and János Sajnovics (1733–1778) also studied the Saami language that was already considered by contemporary international linguists as potentially related to Hungarian. Sajnovics thought that the Hungarian way of spelling would have been more suitable for the recording of the Saami language and for the demonstration of linguistic affinity, as “the Hungarian and the Lapp languages must be identical just on the basis of pronunciation alone”. He put his theory into practice and with the help of the Saami-origin Anders Porsanger (1735–1780), he revised Leem’s lexicon. During their work, in addition to applying Hungarian spelling, they also complemented the entries with Hungarian data. The current paper presents the history of the Saami dictionary written in “Hungarian”.



## 1 Introduction

1769 was a very promising year for the international scientific and academic society, with opportunities and prospects of historical scientific significance. On 3 June of this year, Venus passed between the Sun and the Earth. Through the observation of this astronomical event, astronomers had the means to calculate the distance between the Sun and the Earth. This explains why the transit of Venus triggered the most prominent scientific collaboration in the history of the world until that time, which made it possible for 151 scientists to travel to 77 locations across the Earth to achieve international recognition and glory for the most exact measurement of the parallax<sup>1</sup> of the Sun. Furthermore, these expeditions were of paramount importance not only from the aspect of astronomy but also contributed to several other areas of humanities and science (e.g. linguistics, ethnography, botany, zoology, etc.)

This is how Maximilian Hell (1720–1792) and János Sajnovics (1733–1785), Hungarian Jesuit astronomers, also made close observations on the transit of Venus on the island of Vardø, and at the same time discovered the linguistic affinity between the Hungarian and Saami languages through the examination of the native local people's language.<sup>2</sup> In his work titled *Demonstratio. Idioma ungarorum et lapponum idem esse* (1770), János Sajnovics provided a scientific demonstration of the relationship between the two languages by comparing the respective linguistic levels. This book is still considered one of the fundamental works of Finno-Ugric studies and comparative linguistics. Sajnovics also emerged as a prominent figure of the history of North Saami<sup>3</sup> orthography and lexicography by making observations and suggestions on the Saami language.

---

<sup>1</sup>The angle at which an observer would see the radius of the Earth from the Sun, and which can be used to calculate the distance between the Sun and the Earth.

<sup>2</sup>Saami people (formerly known as Lapps or Laplanders) had never formed a separate, individual state. They have always lived in Sápmi (or Saamiland, formerly known as Lapland), a cultural region of 400,000 square kilometres, stretching across four countries (Norway, Sweden, Finland, and Russia). Saami belongs to the Uralic language family and it is not a unified language, but the sum of nine, separate variants: Southern Saami, Ume Saami, Pite Saami, Lule Saami, North Saami, Skolt Saami, Inari Saami, Kildin Saami, and Ter Saami. During their expedition, Sajnovics and his colleagues researched the North Saami dialect, which is the central Saami language variant, with the most speakers today.

<sup>3</sup>The North Saami language is the largest, central Saami language, spoken by ca. 20,000 people. As the borders of the neighbouring countries never coincided with the borders of the Saami language, North Saami is spoken in three countries (Northern Norway, Northern Sweden, and Northern Finland). From the linguistic point of view, this dialect is the most researched one, and the majority of Saami dictionaries and grammars were written in this dialect.

## 2 The beginnings of Saami lexicography (Sweden)

The beginning of the Christian missionaries' activity was an important milestone in the evolution of the Saami languages – similarly to a number of other languages. Only Protestant missions were able to show significant results in this respect (unlike Catholic ones). The underlying reason was Reformation's strategic emphasis on peoples' native languages, on the importance of making the word of God available for everybody in their mother tongues. This is why the very first written works of the Saami culture were created by translating ecclesiastical texts into Saami. The literacy of Saami languages is very closely linked to missionary activities, as missionaries working among the Saami were the first to write and publish texts in the Saami language. In the 17th and 18th centuries, the Saami texts published in Sweden (Finland included) and (Denmark–)Norway<sup>4</sup> were of three kinds. In addition to descriptions of the Saami people and their traditions (especially regarding their pagan beliefs), translations of religious texts (catechisms and Bible excerpts) constituted the second type of text, while the third type consisted of dictionaries, grammars, and alphabet books for educational purposes to assist missionaries in their linguistic training (Aspaas 2021: 8).

Christian missionary work was the most intense in Sweden, beginning to produce remarkable results in the 17th century. However, Saami lexicography only started to thrive in the 18th century. The very first instance of educational texts was the first Saami dictionary, published in Stockholm in 1738, by a Swedish school teacher and dean, Per Fjellström (1697–1764) from Lycksele.<sup>5</sup> The book was a Swedish-Saami dictionary titled *Dictionarium sueco-lapponicum*. It contained mainly Ume material, but also some Pite and Lule words,<sup>6</sup> not being identified as such in the entries. This dictionary marked the first milestone of Southern Saami lexicography, and was the sole help of Saami translators for over 40 years until the Norwegian Knud Leem (1697–1774) started his activity as a linguist.

---

<sup>4</sup>In 1536, the Kingdom of Denmark and the Kingdom of Norway entered into a personal union, which lasted until 1814. During this time, political and economic control was exercised in Copenhagen, Denmark.

<sup>5</sup>The oldest settlement in the Swedish Lapland.

<sup>6</sup>Although Fjellström travelled to Pite and Lule areas in 1737, his work was based mainly on the Ume dialect, which he presumably acquired as a child (Tóth 2015: 154).

### 3 The beginnings of Saami lexicology in (Denmark–)Norway: Knud Leem

In Norway, missionary activity intensified in the 1700s, partly as a result of the devastating reports of the church and closely related school conditions in Northern Norway (Kjølaas 1995: 26–27). In 1714, Frederick IV established the College of Missions (Missionskollegiet) in Copenhagen to finance and manage Protestant missions under royal patronage. The principal goal of the institution was to organize a mission to India, but from 1716, it also managed the Saami missions.

The first prominent representative of Saami lexicography in (Denmark–)Norway was the Norwegian scholar Knud Leem.

#### 3.1 The life and work of Knud Leem

Knud Leem was born in 1715. At the age of 18, he received a degree in theology from the University of Copenhagen. As a fresh graduate, he was interested in Finnmark<sup>7</sup> and the Saami people living there. In 1723, he requested the College of Missions to be appointed as a missionary among the Saami people of Finnmark. As it was important to him to be able to speak to the Saami in their own language, he started to learn it. Between 1723 and 1725, he studied North Saami at the Seminarium<sup>8</sup> in Trondheim. In 1725, he was granted permission to work as a missionary in Porsanger and Laksefjorden in Finnmark and thus had the opportunity to study the Saami language and culture more closely. Twenty-five years after its closure, the Seminarium Scholasticum was reopened in 1752, now called the Seminarium Lapponicum. Leem was appointed as head of the institution. He taught the Saami language at the Seminarium Lapponicum from 1752 to his death in 1774. In the meantime, he published a number of works. After his 400 page long grammar book,<sup>9</sup> his first paper as a professor<sup>10</sup> was a Danish–Saami glossary titled *En Lappisk Nomenclator efter den Dialect, som bruges af Fjeld-Lapperne i Porsanger-Fjorden* [Saami glossary on the basis of the Mountain Saami dialect spoken around the Porsanger Fjord] published in 1756 in Trondheim.

<sup>7</sup>It is the north-easternmost and largest (74 293 km<sup>2</sup>) county of Norway. It constitutes a part of Saamiland and a quarter of its population is made up of the Saami people.

<sup>8</sup>Between 1717 and 1727, the institution provided Saami linguistic training for future priests and missionaries, preceding their theological studies in Copenhagen.

<sup>9</sup>*En lappisk Grammatica efter den Dialect, som bruges af Field-Lapperne udi Porsanger-Fiorden: samt et Register over de udi samme Grammatica anførte Observationers Indhold, hvorhos er føyet et Blad af den berømmelige Historie-Skriveres Hr. Baron Ludvig Holbergs Kirke-Historie oversat i det Lappiske Tungemaal med en Analyse over hvert Ord.*

<sup>10</sup>In 1752, Leem was awarded the title „professor lingvæ lapponiæ” (for more details, c.f. Magga 2003).

## 4 About work on the *Lexicon Lapponicum*

From 1753, Leem was working on his great work, the trilingual (North Saami–Danish–Latin) dictionary titled *Lexicon Lapponicum*. The printing started in Trondheim<sup>11</sup> in June of 1768. The elderly Leem needed a permanent deputy due to his deteriorating health, which position was filled by Anders Porsanger (1735–1780).

### 4.1 The conflict between Anders Porsanger and Knud Leem

Anders Porsanger<sup>12</sup> was also a pioneer in a way, as he was the first Saami person to hold a degree in theology. He was 15 years old when he was invited by Johan Ernst Gunnerus (1718–1773), Bishop of Trondheim<sup>13</sup> to help Leem. He was studying at the local cathedral school at that time, completed his theological studies in Copenhagen in 1716, and became a missionary in Varanger<sup>14</sup> in the following year. In 1764, he returned to Trondheim and started to work as a senior lecturer at the Seminarium Lapponicum, and as a hospital priest at the local hospital. Porsanger had numerous duties in addition to dictionary writing since he was obliged to undertake other jobs to earn a living. This hindered work on the dictionary, which made Leem very dissatisfied. Although the manuscript of the second volume was already in progress during the publishing work on the first volume, the tension that was present between Leem and Porsanger from the beginning continuously heightened. The fact that Porsanger was expected to undertake more and more responsibilities added to the issue. Leem only taught a few classes in the Seminary by this time, and he also did little work on the dictionary. Furthermore, Porsanger was asked to write a report<sup>15</sup> for the College on Saami orthography in September of 1767, in which he criticised Leem’s difficult spelling system.<sup>16</sup> The conflict between them reached its peak in 1769 when the *Lexicon* suffered another “attack”.

---

<sup>11</sup>The first volume (ca. 42 arcs) was printed before the Sajnovics “incident”.

<sup>12</sup>His original name was Anders (Biret-Ánde) Andersen, and he took on the Porsanger name as a tribute to his homeland (Solbakk 1998: 9).

<sup>13</sup>Johan Ernst Gunnerus was a patron of sciences, a botanist himself, and professor of theology at the University of Copenhagen. He was the Bishop of Trondheim from 1758 until his death. He was also among the founding members of the Trondheim Society of Sciences.

<sup>14</sup>The easternmost administrative unit of Finnmark.

<sup>15</sup>The report is available in Martinussen (1992).

<sup>16</sup>For more details, see Kelemen (2019).

## 5 Hungarians in the North

The 1769 transit of Venus attracted the attention of Christian VII (1749–1808), the Danish-Norwegian king as well, who was interested in astronomy. He had an observatory built on the Norwegian Island of Vardø and organised an expedition, providing considerable financial support. Christian wrote a letter to Maria Theresa and asked her to nominate an expert from the Vienna Observatory who had the required skills. The empress was well-known as a patron of the sciences, so she supported the planned expedition and invited the director of the Vienna observatory, Maximilian Hell<sup>17</sup> for the mission. He gladly accepted the invitation, and as he needed an assistant, he chose a former colleague of his, who was also a Jesuit: János Sajnovics (for further details, see Maticsák & Kelemen 2021: 403–404).<sup>18</sup>

### 5.1 About János Sajnovics

János Sajnovics was born in Tordas,<sup>19</sup> Hungary, in 1733. He studied at the prestigious Jesuit Grammar School of Győr<sup>20</sup> and then in Buda<sup>21</sup> from 1747. He showed an early interest in science, especially mathematics and astronomy. In 1748, he joined the Jesuit Order and in 1751 he began his philosophical studies in Trnava.<sup>22</sup> After completing his studies, he started working as a teacher in Pozsony (modern-day Bratislava). A few years later, he started to study theology at the College of Vienna, also supervised by the Order, where he became acquainted with astronomy at a professional level. His superiors soon recognized his talent and moved him to the observatory of Vienna to work with Maximilian Hell (who was already

---

<sup>17</sup>Maximilian Hell was born in 1720, in Banská Štiavnica (hun. Selmecebánya), in today's Slovakia. He joined the Jesuit Order in 1738 and continued his studies in Vienna from 1740. He got acquainted with the science of astronomy in the Jesuit Observatory of Vienna. In the September of 1755, he was appointed as director of the newly founded Observatory and as professor of mechanics at the University of Vienna. After the dissolution of the Jesuit order in 1773, Hell joined the secular clergy and continued his versatile academic work. He died in Vienna, in 1792, at the age of 72.

<sup>18</sup>It is not known why the king needed an expert from Vienna. The assignment of Hell (and Sajnovics) was especially interesting, because the presence of Jesuits was forbidden at the time in (Denmark-)Norway. It seems likely that the prestige of this scientific exploration was more significant than the religious ban (c.f. Aspaas & Kontler 2019: 186).

<sup>19</sup>A village in Central Transdanubia, Hungary.

<sup>20</sup>Győr is the most important city of northwest Hungary, the capital of Győr-Moson-Sopron County.

<sup>21</sup>It was one of the historical capital cities of the Kingdom of Hungary. In 1873, Buda, Óbuda, and Pest were officially unified and given the new name Budapest.

<sup>22</sup>Nagyszombat in Hungarian.

well-known internationally), and later, in 1766, to the Trnava observatory to work with Ferenc Weiss (1717–1785), prominent mathematician, astronomer, and Jesuit priest. The 1773 dissolution of the Jesuit Order caused a break in Sajnovics's career as well: he became a deputy astronomer in the Buda observatory. Toward the end of his life, he became more and more withdrawn and died at the young age of 52, in 1785.

## 5.2 Hungarian and Saami are “identical”

After Hell's invitation, Sajnovics arrived in Vienna on 12 April 1768, and at the end of the month, they set off on the long journey of nearly six months. They traveled through Prague, Dresden, Leipzig, Hamburg, and Lübeck to reach their destination: Copenhagen, then called Hafnia. The rough sea voyage finally came to a happy end on 11 October: the two scientists arrived at the Island of Vardø. During the months spent in Vardø, they made preparations to observe Venus and did other types of scientific research as well (e.g., on the sparkling of the Arctic Ocean, auroras, and tidal patterns). In addition to scientific observations, they also started to study the Saami language.

Hell already had some preliminary information on the presupposed linguistic affinity between the Hungarian and the Saami languages. This is why it was thought for a long time that linguistic research was also planned even before the journey and, as Hell did not speak good Hungarian, he chose Sajnovics knowingly, because of his linguistic competence. However, now it has become evident that the two scholars started their linguistic research so “unprepared”, that it could not have been deliberately planned.<sup>23</sup>

## 5.3 “Decoding” Leem's spelling system

For his linguistic research, Sajnovics used Knud Leem's above-mentioned grammar book and Danish–Saami glossary; however, he had considerable difficulty with the interpretation of these works:

When I opened up the grammar book, I saw a long list of rules, however, they were in Danish, a language I was completely unfamiliar with ... When I took a look at the Nomenclator, I saw Lapp words, consisting almost exclusively of vowels and some curiously selected consonants (Sajnovics 1994: 36).

---

<sup>23</sup>They did not take any kind of literature along. They received all material – even the Hungarian ones – during their stay abroad.

He was very discouraged and sadly reported to Hell that he saw no connection between the two languages. However, they did not give up, and took the following steps to find useful language material. First, they collected words from the *Nomenclator* that were similar to Hungarian words in form. Then they looked for the Latin equivalents of their Danish meaning. For this step, they used historian Hans Gram's (1685–1748) Latin–Danish dictionary titled *Nucleus Latinitatis* [The nucleus of Latin] (published in Copenhagen in 1722). The extremely tedious work often ended in failure, either because the word chosen was not in the *Nucleus*, or because the meaning of the Saami word and the Hungarian word it was intended to be related to were completely different after all (Sajnovics 1994: 36). They were already experimenting with novel phonetic notations at this time, but their work only yielded reliable results when native Saami speakers arrived and, with the help of an interpreter, they started asking them about various words. In these words, they saw a such similarity – both in their pronunciation and their meaning – that Sajnovics went as far as to say that these were in fact Hungarian words (Sajnovics 1994: 37). During these personal encounters, Sajnovics put down the words heard in Hungarian spelling, and he managed to compile a very rich corpus. From this point on, he worked with much more enthusiasm, and the work progressed faster. As a next step, within a few weeks' time, he detected the Latin equivalents of all the headings of the *Nomenclator*, making it easier to observe the spelling system and provide the rules for converting it to Hungarian spelling.

The painstaking work was finally completed in 1770: as an argument for the linguistic affinity between the Hungarian and Saami languages, Sajnovics published his work titled *Demonstratio. Idioma Ungarorum et Lapponum idem esse* in Copenhagen. This work is considered one of the fundamental works of Finno-Ugric studies and comparative linguistics today.<sup>24</sup>

Sajnovics's work also concerned the field of North Saami orthography and lexicography: he confidently claimed that Hungarian spelling would facilitate the description of the North Saami language. This was not a mere suggestion on his part, as he also intended to realize his idea: he wanted to publish the *Lexicon Lapponicum*<sup>25</sup> with Hungarian spelling, even though it was already in press.

---

<sup>24</sup>In recognition of his work, he was admitted as a member of the Norwegian Academy of Sciences on 5 February 1770 and of the Danish Academy of Sciences on 9 February 1770 – as the second foreigner after Hell (Sajnovics 1990: 135, 137).

<sup>25</sup>For a history of the lexicon see Kelemen (2018): and for a detailed description see Djärv (2003)

## 6 The dictionary of the Hungarian Jesuit and the Saami theologian

In fact, the idea of introducing Hungarian orthography had already come up on the Island of Vardø, during the study of the Saami language and the description of the linguistic affinity. Hell (as the idea was his) decided to take concrete steps to reform Saami spelling:

Reverend Father HELL's main wish was to introduce the solid Hungarian spelling practice into the Lapp books and texts ... Reverend Father HELL considered two more things necessary. Firstly, LEEM's dictionary, which is about to be republished, and the Grammatica should be given a Hungarian appearance. Secondly, there should be a Laplander who, having learned the Hungarian spelling from me, would supervise the printing of the dictionary and the Lapp training of the missionaries and teachers to be sent to Lapponia. In addition, it is necessary to invite to Copenhagen ... a man who is well versed in the Lapp language, whom I could personally teach to the Hungarian spelling, and with whom I could adapt the LEEM dictionary to the Hungarian spelling (Sajnovics 1994: 42–43).

With this agenda, they approached Otto Thott (1703–1785),<sup>26</sup> then-president of The Royal Danish Academy of Sciences on 13 February 1770: “We handed him in writing the arguments that would convince him that the great Lapp lexicon ought to be printed with Hungarian orthography” (Sajnovics 1990: 135). Thott was very enthusiastic about the idea. For his request, Bishop Gunnerus appointed Leem's assistant, Anders Porsanger, to help Sajnovics re-edit the dictionary. The choice could not have been more perfect, not only because Porsanger was a native speaker of Saami, but also because he had worked with the professor since he was 15 years old, and had helped him edit the dictionary he was transcribing, so he knew the material well. Sajnovics was aware of how unique a Saami man his “colleague” was, having been raised “in a way very different from the ways of his homeland” and having chosen to become a scholar in a period when the majority of Saami people were undereducated or outright illiterate.<sup>27</sup>

---

<sup>26</sup>Danish count and statesman. He held several positions at the time, e.g. he was a member of the Privy Council, Highest Chancellor of the Danish Chancery, etc.

<sup>27</sup>“This man came from Porsanger and studied in Nidrosia (i.e. Trondheim) as a young man. He received an excellent education in the field of humanities there. At the University of Copenhagen, he studied philosophy and theology. He spent three years as a missionary among his fellow Saami people, and then was appointed as a priest at the church of the Trondheim hospital. [...] This excellent scholar spoke not only his Saami mother tongue but also Danish, German, Latin, Greek and Hebrew, to the glory of the talented Saami people” (Sajnovics 1994: 43).

Sajnovics provided a very detailed description of the nature of changes they made in Leem's material in Chapter IV of the *Demonstratio's* second, Trnava edition.<sup>28</sup>

Porsanger himself also read the *Demonstratio* and thus regarded the linguistic affinity of the two languages as established. As Sajnovics reported, Porsanger was quick to get acquainted with the Hungarian way of spelling. They spent entire days working together, to have the dictionary ready for printing. Porsanger was reading the Saami words aloud while Sajnovics was recording in writing. Afterwards, the glossary was organized in alphabetical order. In Sajnovics's version of the dictionary, the headwords were spelt according to Hungarian orthography, each followed by a Hungarian synonym (if available), then the Danish and Latin equivalents. At the end of each entry such Hungarian words were included in italics (together with their Latin meanings) as were similar to the headword but had different meanings (see Table 1).<sup>29</sup>

Table 1: A comparison of dictionary entries by Sajnovics (1994: 44) and by Leem (1768: 7)

|                    |                                                                                           |
|--------------------|-------------------------------------------------------------------------------------------|
| Leem's entry:      | ABOM, syder over, efferveo. 3 sing. præ. ind. Abbo. 2 sing. præ. imp. Abo.                |
| Sajnovics's entry: | ABOM forrok, Syder Ower, efferveo. <i>hab</i> (spuma, unda), <i>habzom</i> (spumo, undo). |

Finally, Sajnovics wrote the preface of the book, providing extensive information on a) the grounds for changing the original spelling, b) the changes themselves, c) the status and rules of the contemporary Hungarian orthography, d) the pronunciation of the vowels and the consonants, and e) the benefit of the endeavor in general Sajnovics (1994: 44).

They completed their work in May 1770 and handed over the manuscript to Otto Thott, who promised to send some copies to Hungary when the Saami dictionary with the "Hungarian spelling" was printed. Sajnovics and Hell set off from Copenhagen to Hungary on 22 May 1770, leaving behind the Danish-Norwegian academic community in turmoil because of the rewritten dictionary.

<sup>28</sup>The first edition, published in Copenhagen in 1770, did not contain the chapter describing the circumstances of the dictionary's transcription.

<sup>29</sup>Sajnovics noted these for future reference, in case "another meaning of the Lapp word should be found" (Sajnovics 1994: 44).

## 7 The fate of the manuscript

The fate of the dictionary manuscript was documented in Bishop Gunnerus's correspondence: in his letter written on 13 December 1770, the bishop asked for Leem's opinion on Sajnovics's spelling proposal. Two days later, he also notified the Missionary College about handing the "material" to Leem, who was still reviewing the work (Dahl 1906: 49–50).

On 22 April 1771, Gunnerus sent two copies of the dictionary, one of Leem's and one of Sajnovics's version, to Christian Weldingh (1728–1801), a Danish-born missionary. His task was, similarly to that of another missionary, Gerhard Sandberg (1741–1805) (who published the book at the end), to evaluate which way of spelling would be ideal in the dictionary independently of each other.

The issue was left unresolved. Gunnerus made no more mention of Saami orthography or the issue of the dictionary in his surviving letters. The next letter touching upon this subject informed the readers that the publication of the dictionary was moving forward with the letter M, in Leem's original version Dahl (1906: 54). The full dictionary was published as late as 1781, together with its reverse version (Danish–Latin–Saami).

The manuscript of Sajnovics's dictionary disappeared without a trace. Neither the (Danish–)Norwegian nor the Hungarian side had any more information about it. However, it is obvious that everybody who was involved abandoned the subject. Sajnovics and Hell returned home, and Porsanger's career took an unfortunate turn after their joint project (as he was sent on a mission again and his academic work ended for good). In December 1770, Christian VII's personal physician, Johann Friedrich Struensee (1737–1772) was appointed, and Otto Thott was removed from office. Thus, the idea of rewriting the dictionary remained without any local supporters. In addition, the Missionary College of Copenhagen intended to publish Leem's dictionary very quickly and in a cost-effective manner, and as the Hungarian diacritical marks were new to printing technology, the Hungarian spelling would have made the printing more expensive. Not to mention the fact that these signs would have seemed a bit out-of-place in the Scandinavian linguistic context, not accommodating well to the then popular Gothic script either.

## 8 Summary

Although research on Sajnovics's work is extensive both within Hungary and abroad, discussing almost all aspects of his work, the history of the Saami dictionary's manuscript, and Sajnovics's efforts to apply Hungarian spelling in it

have received very little attention. I was the first to present this topic to the Hungarian research community in 2020 and 2021, within the scope of the 250-year centenary lecture series on the *Demonstratio*. The present paper is intended to reflect on the 1770 Copenhagen events, the trilingual Saami dictionary, and János Sajnovics's dictionary writing activity in the international context as well.

This research opens up new perspectives on the Saami history of science: besides the presentation of the manuscript's history (which already addresses a research niche), my ultimate goal is to retrieve the manuscript through international, in particular Norwegian/Danish, scientific cooperation, thus adding to Sajnovics's legacy. If the manuscript exists, after a detailed linguistic (lexical, lexicological, orthographic, etc.) analysis, I intend to present it in a monograph, first to the Hungarian and then to the international researchers interested in this topic.

## References

- Aspaas, Per Pippin. 2021. *Fororder til Knud Leems Lexicon Lapponicum. Oversatt til norsk med en innledning og kommentarer av Per Pippin Aspaas*. (Bibliotheca Neolatina Upsaliensis XVII.). Uppsala: Peter Sjökvist och Krister Östlund. <http://uu.diva-portal.org/smash/record.jsf?pid=diva2%3A1622412&dswid=948>.
- Aspaas, Per Pippin & László Kontler. 2019. *Maximilian Hell (1720–92) and the ends of Jesuit science in Enlightenment Europe*. (Jesuit Studies). Leiden–Boston: Brill Academic Publishers.
- Dahl, Ove Christian. 1906. *Biskop Gunnerus' virksomhed fornemmelig som botaniker, tilligemed en oversigt over botanikerens tilstand i Danmark og Norge indtil hans død*. (Det Kongelige Norske Videnskabers Selskabs Skrifter. 4.). Trondhjem: Aktietrykkeriet i Trondhjem. [https://www.ntnu.no/ojs/index.php/DKNVS\\_skrifter/article/view/1119](https://www.ntnu.no/ojs/index.php/DKNVS_skrifter/article/view/1119).
- Djärv, Ulrika. 2003. *De första samiska ordböckerna—mission och upplysning* (Opuscula Uralica 7). Uppsala: Finsk-ugriska institutionen, Uppsala universitet.
- Kelemen, Ivett. 2018. A 18. Századi számi lexikográfia úttörője: Knud Leem. Vázlatok a Lexicon Lapponicum Bipartitum történetéről. *Folia Uralica Debreceniensia* 25. 99–108. [https://finnugor.arts.unideb.hu/fud/fud25/07\\_kelemenivett.pdf](https://finnugor.arts.unideb.hu/fud/fud25/07_kelemenivett.pdf).
- Kelemen, Ivett. 2019. Anders Porsanger jelentése Knud Leem (északi) számi helyesírásáról. *Folia Uralica Debreceniensia* 26. 123–136. [https://finnugor.arts.unideb.hu/fud/fud26/08a\\_kelemenivett.pdf](https://finnugor.arts.unideb.hu/fud/fud26/08a_kelemenivett.pdf).
- Kjølaas, Per Oskar. 1995. *Bibelen på samisk. En bok om samisk bibeloversettelse*. Oslo: Det Norske Bibelselskap.

- Leem, Knud. 1768. *Lexicon Lapponicum bipartitum, Lapponico–Danico–Latina & danico–latino–lapponica cum indice latino. Pars prima Lapponico–Danico–Latina*. Trondheim: Impensis Seminarii Lapponici Fridericiani. <https://www.nb.no/nbsok/nb/f7147c85118952098ba3af3b2b513745?index=5#3>.
- Magga, Ole Henrik. 2003. Knud Leem som utforsker av samisk språk og kultur. In Jan Ragnar Hagland & Steinar Suphellen (eds.), *Knud Leem og det samiske*. Vol. 2 (Det Kongelige Norske Videnskabers Selskab), 29–43. Trondheim: Tapir Akademisk Forlag.
- Martinussen, Bente. 1992. Anders Porsanger – teolog og språkforsker fra 1770-tallets Finnmark. *Nordlyd* 18. 15–59.
- Maticsák, Sándor & Ivett Kelemen. 2021. Csillagászat és nyelvészet. *Debreceni Szemle* 29(4). 399–413. <https://ojs.lib.unideb.hu/szemle/article/view/10659>.
- Sajnovics, János. 1990. *Sajnovics naplója 1768–1769–1770* (Bibliotheca Regulyana 1). Budapest: ELTE Finnugor Nyelvtudományi Tanszéke.
- Sajnovics, János. 1994. *Demonstratio* (Bibliotheca Regulyana 2). Budapest: ELTE.
- Solbakk, Aage. 1998. Anders Porsanger–Biret Ánde. In Harald Gaski & John Trygve Solbakk (eds.), *Cafe boddu* (Essayöakkáldat 3), 7–19. Karasjok: Davvi Girji OS.
- Tóth, Anikó. 2015. A számi (lapp) lexikográfia évszázadai. *Nyelvtudományi Közlemények* 111. 151–183. DOI: 10.15776/NyK.2015.111.7. <http://archive.nytud.hu/nyk/nyk111.pdf>.



# Chapitre 18

## Les exemples dans les dictionnaires français–hongrois à travers les siècles

Gábor Tillinger

Uppsala University, Sweden

The aim of this chapter is to give an overview of the most important French–Hungarian general dictionaries, published since 1844, as well as to present a comparative study of their microstructure, especially from the point of view of the different types of examples provided in the entries. The study focuses on entries of frequent and semantically overloaded verbs, namely *aller*, *avoir*, *devoir*, *être*, *faire*, *laisser*, *mettre*, *pouvoir*, *prendre*, *rendre* and *venir*, the use of which is often illustrated by numerous examples – due to the polysemy and high frequency of these verbs in different contexts. Among other things, it can be noted that the evolution of the entries over time does not automatically mean improvement. The study also comments on the choice and the quantity of examples in the dictionaries, while analysing the functions and the purpose of the different examples. As for the quantity of examples, the analyses include a comparison with other bilingual dictionaries having French as the source language. Finally, some entries are analysed in more detail to illustrate the difficulties arising from the fact that Hungarian and French are not related languages.

### 1 Introduction

#### 1.1 L'arrière-plan

De nos jours, les Hongrois francophiles, ainsi que les Francophones intéressés par la langue hongroise, connaissent de manière générale trois dictionnaires français–hongrois volumineux (de grand format ou de taille moyenne) : le grand dictionnaire ECKHARDT & OLÁH (1999), le dictionnaire de PÁLFY (1999) et celui de BÁRDOSI & SZABÓ (2007). Les philologues et les générations plus âgées



connaissent également au moins une des nombreuses éditions du grand dictionnaire d'Eckhardt, publié pour la première fois en 1953 – soit le prédécesseur du grand dictionnaire paru en 1999, une version complètement remaniée et complétée sous la direction de Tibor Oláh.<sup>1</sup> Le premier grand dictionnaire français–hongrois rédigé par Aurélien Sauvageot n'est connu désormais que par les érudits ou les francophiles les plus engagés. En ce qui concerne les dictionnaires plus anciens, publiés dans le XIXe siècle ou au début du XXe siècle, ils ont presque complètement disparu dans les oubliettes de l'Histoire. Bien évidemment, cela n'est pas un phénomène unique, en revanche, le fait que les dictionnaires français–hongrois les plus anciens soient quasi inconnus même parmi les philologues est tout de même dommage.

## 1.2 Les objectifs

Nous avons pour objectifs de faire le panorama des dictionnaires français–hongrois les plus importants – parus dans les XIXe, XXe et XXIe siècles, compte tenu seulement des dictionnaires généraux –, et de présenter notre étude comparative, en mettant l'accent sur la microstructure et surtout sur les différents types d'exemples fournis dans certaines entrées. Il est important de mettre en évidence que les dictionnaires hongrois–français ne font pas partie de cette étude. Notre but principal consiste donc à présenter les dictionnaires français–hongrois les plus significatifs – avec plus de détails pour ceux parus jusqu'au début du XXe siècle –, énumérer leurs caractéristiques, tout en décrivant les changements constatables au fil du temps. Ce faisant, nous pourrons aussi montrer comment certains traits typiques d'une langue de « l'Ouest » sont rendus par des équivalents dans une langue de « l'Est ».

## 2 Matériels et méthodes

### 2.1 Matériels : Les dictionnaires français–hongrois

Dans cette partie, nous présentons séparément tous les dictionnaires français–hongrois dont nous nous sommes servis dans nos analyses. Nous nous concentrons sur les dictionnaires généraux (c'est-à-dire les ouvrages n'ayant pas pour but de couvrir seulement un lexique spécialisé), et les plus importants du point

---

<sup>1</sup>En Hongrie, le nom de famille précède le prénom. Puisque notre texte est écrit en français, nous avons changé l'ordre des noms de personne hongrois.

de vue de leur volume, de leur contenu et de leur rôle dans l'histoire de la lexicographie franco-hongroise. En fin de cette partie, nous précisons aussi les types d'ouvrages exclus de notre étude.

### 2.1.1 Kiss 1844

Celui de Mihály Kiss (Michel Kiss sur la page de titre française) est le premier véritable dictionnaire bilingue avec le français en tant que langue source et le hongrois comme langue cible. Avant 1844, il existait quelques grammaires françaises destinées aux Hongrois, ainsi que d'autres ouvrages plurilingues, qui, entre autres, comprenaient des listes de vocabulaire avec des équivalents français et hongrois, mais c'est le dictionnaire de Kiss qui est le premier à inclure seulement le français et le hongrois. L'année 1844 est par ailleurs très importante dans l'histoire de la langue hongroise, étant donné que c'est cette année-là qu'elle est devenue la langue officielle unique en Hongrie.

Sur la page de titre, on trouve le titre « Nouveau dictionnaire de poche français–hongrois et hongrois–français ». L'épithète « nouveau » peut paraître surprenant, compte tenu du fait que c'est le premier dictionnaire français–hongrois. Dans ce cas – tout comme dans le cas de plusieurs ouvrages publiés avant le XXe siècle –, le mot « nouveau » indique la nouveauté de l'ouvrage dans le sens où il s'agit soit d'une œuvre auparavant inexistante, soit d'une publication récente. Cette épithète ne s'utilisait pas forcément pour indiquer la nouvelle édition d'un livre, et, en quelque sorte, c'est aussi un moyen de la part de la maison d'édition d'attirer l'attention des clients. Ce qui de plus peut être fallacieux, d'après les informations de la page de titre, c'est que l'on s'attendrait à deux parties différentes (français–hongrois et hongrois–français) parues dans un même volume, ou bien au moins publiées parallèlement, en même temps, en la même année. En revanche, ce que l'on trouve dans l'ouvrage publié en 1844, c'est seulement la partie français–hongrois (rédigée donc par Kiss), alors que la partie hongrois–français a paru en 1848, rédigée par Ignác Karády (Ignace Karády sur la page de titre française).

Karády mentionne dans sa préface courte, écrite seulement en hongrois, que l'éditeur lui a confié la rédaction de la partie hongrois–français, puisque Kiss n'avait pas continué son travail – peut-être en raison de certains empêchements, ajoute-t-il – et l'éditeur souhaitait publier la seconde partie le plus tôt possible. On peut donc supposer que – selon l'intention originelle de l'éditeur – Kiss aurait aussi dû rédiger la partie hongrois–français, et la parution de la seconde partie n'aurait pas dû tarder quatre ans.

Dans la partie français–hongrois, on ne trouve aucune préface, aucune information sur le travail de la rédaction. On ne peut que supposer que Kiss s’est servi de dictionnaires français–allemand en tant que point de départ, mais cela n’est pas sûr du tout. D’ailleurs, on ne peut rien savoir sur les sources de Karády non plus.

Il est aussi intéressant de remarquer qu’aucun exemplaire de la première édition n’est accessible en Hongrie, même la Bibliothèque nationale de Hongrie ne possède que la deuxième édition de 1852 – qui est d’ailleurs une simple réimpression, le contenu des deux éditions est identique. Nous avons travaillé avec une copie de la première édition issue du British Museum.

Selon certaines sources (voir p.ex. PERROT 1992 : 95), il existerait une troisième édition aussi, parue en 1865, mais nous n’en avons trouvé aucune trace. Ce fait a été également souligné par SZABÓ (2011 : 80) qui se demandait s’il s’agissait du même dictionnaire que celui de Babos. Perrot en écrit ainsi :

Il avait existé plusieurs dictionnaires hongrois–français et français–hongrois dont certains étaient très estimables, notamment celui de KISS et KARÁDY qu’avait publié l’éditeur Heckenast à Pest et dont il faut en particulier citer la troisième édition « revue et considérablement augmentée », publiée en 1865 par BABOS Kálmán sous le titre de *Nouveau dictionnaire hongrois–français et français–hongrois d’après les ouvrages publiées* [sic] *jusqu’à ce jour*. (PERROT 1992 : 95)

D’après les données fournies par Perrot, nous supposons que Perrot a pris – incorrectement, à notre avis – le dictionnaire rédigé (et non seulement « publié ») par Babos pour une troisième édition du dictionnaire de Kiss et de Karády. En revanche, si l’on lit attentivement la préface du dictionnaire de Babos, on peut mieux comprendre la situation – pour notre solution, voir Section 2.1.3. Il faut aussi remarquer que Perrot ne cite pas correctement le titre du dictionnaire paru en 1865 – pour le titre correct, voir nos références bibliographiques ou Section 2.1.3. Malheureusement, on ne peut pas savoir pour quelles raisons Perrot considérait le dictionnaire publié en 1865 le plus estimable parmi tous les dictionnaires parus avant celui de Sauvageot en 1932.

Si ce dictionnaire s’intitule « dictionnaire de poche », cela ne s’explique pas par le nombre des entrées (dont il y a plus de 20 mille), mais surtout par la simplicité (la pauvreté) des entrées – en ce qui concerne le nombre des exemples. Nous avons décidé d’inclure cet ouvrage dans nos analyses surtout parce que c’est le premier véritable dictionnaire français–hongrois.

2.1.2 **BABOS 1865**

Le dictionnaire qui apparaît en 1865 est rédigé par Kálmán Babos, mais son nom ne figure pas sur la page de titre de la partie français–hongrois (bien que ce soit le « tome premier »), on ne le trouve que dans le « tome second » (ainsi que la forme francisée de son nom : Coloman Babos), soit la partie hongrois–français. Curieusement, c’est aussi dans la partie hongrois–français que l’on peut trouver une préface, en langue hongroise seulement, écrite par Babos. C’est grâce à cette préface que l’on peut apprendre qu’il était le rédacteur des deux parties. La préface révèle également la solution du mystère autour de la troisième édition supposée du dictionnaire de Kiss. La page de titre du dictionnaire de Babos inclut l’inscription « Troisième édition revue et considérablement augmentée », ce qui est aussi une source de confusion, sachant que le dictionnaire en question est le premier par Babos. Dans sa préface, Babos explique que son œuvre n’a rien à voir avec les deux éditions du dictionnaire de Kiss (et de Karády, qui était le rédacteur de la partie hongrois–français, publiée d’abord en 1848, puis en 1852). Comme Babos l’explique, la formulation « troisième édition » est due au fait que l’éditeur de son dictionnaire (notamment Gustave Heckenast, en hongrois Heckenast Gusztáv) est le même que celui des deux éditions de Kiss. Selon cette explication, la troisième édition mystérieuse du dictionnaire de Kiss est en réalité la première édition du dictionnaire de Babos – soit la troisième édition d’un dictionnaire français–hongrois et hongrois–français chez Heckenast. Même au niveau du nombre des entrées, le dictionnaire de Babos avec ses plus de 30 mille entrées est un ouvrage plus volumineux que celui de Kiss.

Sur la page de titre de l’ouvrage de Babos, on trouve également l’information suivante, soit le titre du dictionnaire : « Nouveau dictionnaire français–hongrois et hongrois–français d’après les meilleurs ouvrages publiés [sic] jusqu’à ce jour et spécialement d’après le dictionnaire français de Molé ». Dans sa préface, Babos mentionne deux dictionnaires dont il s’est servi au cours de la rédaction de la partie hongrois–français : la deuxième édition du dictionnaire hongrois et allemand de Mór Ballagi (nous précisons que c’est le dictionnaire hongrois–allemand et allemand–hongrois de Mór Ballagi, ouvrage publié en 1847–1848, les deux parties dans des volumes séparés, le nom de l’auteur dans le dictionnaire étant Móricz Bloch) et le dictionnaire allemand–français de Schuster (il s’agit probablement de l’édition de 1848). Babos explique qu’il a employé ces dictionnaires de façon presque exclusive au cours de la rédaction de la partie hongrois–français de son dictionnaire, et il a complètement réécrit le dictionnaire, indépendamment du dictionnaire de Kiss et de Karády. En revanche, Babos n’écrit rien sur les ouvrages utilisés au cours de la rédaction de la partie français–hongrois, ce qui est

plus que surprenant si l'on considère l'information sur la page de titre : « spécialement d'après le dictionnaire français de Molé ». Ce que l'on ne peut pas donc savoir après avoir examiné les différentes parties informatives du dictionnaire (surtout la préface), c'est qu'à quel point Babos s'est servi de l'œuvre de Molé – mentionné seulement sur la page de titre. En même temps, d'autres questions surgissent : De quel dictionnaire de Molé s'agit-il exactement ? Qui est-ce qui se cache derrière le nom de Molé ? On ne trouve aucune réponse directe à ces questions dans l'ouvrage de Babos.

Si l'on considère la préface du dictionnaire français–hongrois de THEISZ (1902) – dans laquelle Theisz prétend que tous les dictionnaires français–hongrois précédant le sien, excepté le dictionnaire de Spanraft, ont été basés sur des dictionnaires allemands (voire même qu'ils étaient les copies de certains dictionnaires allemands) – « le dictionnaire français de Molé » doit être un dictionnaire français–allemand. Le nom « Molé » apparaît parfois sous la forme « A. Molé » en tant que l'auteur de plusieurs dictionnaires français, dont le plus remarquable (en raison de son volume et de son contenu) du point de vue de l'allemand est le *Nouveau dictionnaire français–allemand et allemand–français à l'usage de tous les états*, paru pour la première fois en 1841 (BRAY 2000 : 175). Ce dictionnaire de Molé a eu 23 éditions entre 1841 et 1865 (la 42<sup>e</sup> édition a paru en 1900 !), il est donc impossible de savoir de quelle édition Babos s'est servi au cours de son travail – si c'est bien l'ouvrage de Molé que Babos a utilisé. En même temps, il faut aussi remarquer que les rééditions de ce dictionnaire (comme c'était souvent le cas dans la seconde moitié du XIX<sup>e</sup> siècle) était des réimpressions par stéréotypage (BRAY 2000 : 168), le contenu n'ayant pas ainsi été altéré. Il existe une version abrégée de ce dictionnaire (*Nouveau dictionnaire de poche français–allemand et allemand–français à l'usage des écoles*), dont la première édition a été publiée en 1844 (BRAY 2000 : 175), mais, à notre avis, il est peu probable que ce soit l'ouvrage utilisé par Babos.

Le Babos – Molé n'a eu qu'une seule édition.

### 2.1.3 MÁRTONFFY 1879

Le dictionnaire de Frigyes Mártonffy (ou Frédéric Mártonffy dans la partie écrite en langue française de la page de titre) semble être un travail plus autonome. Au sujet de ses sources, Mártonffy n'écrit dans sa préface qu'une seule phrase que nous citons ici :

Notre ouvrage renferme dans ses deux parties la nomenclature de tous les mots usuels des deux langues, puisée pour la partie française dans les

meilleurs ouvrages anciens et modernes, et surtout dans le *Dictionnaire de l'Académie* et celui *des difficultés de la langue française*; et pour la partie hongroise dans les œuvres de M. Dr. MAURICE BALLAGI.

Mártonffy écrit donc qu'il a utilisé les meilleurs ouvrages français et hongrois, mais on ne peut pas savoir comment et à quel point. De plus, outre la question des différentes éditions qu'il ne précise pas non plus, on ne peut pas être sûr de quel dictionnaire « des difficultés de la langue française » il s'agit. On ne peut que se demander si Mártonffy pense à l'ouvrage de Pierre Boiste (paru en 1828) ou à celui de Jean-Charles Laveaux (dont la première édition a été publiée en 1818, suivie par des éditions revues, corrigées et augmentées, la deuxième en 1822, la troisième en 1846 et la quatrième en 1873). On peut aussi constater qu'il ne fait mention d'aucun dictionnaire français–allemand employé.

En ce qui concerne le nombre des entrées, nous estimons que ce dictionnaire contient plus de 40 mille mots-vedettes.

Le dictionnaire de Mártonffy n'a eu qu'une seule édition. La partie hongrois–français a été rédigée par Babos – soit la 2e édition de la partie hongrois–français de son dictionnaire de 1865.

#### 2.1.4 POKORNY 1880

L'auteur du dictionnaire français–hongrois paru en 1880 est Jenő Ármin Pokorny (Eugène Armand Pokorny selon la page de titre française).

Dans sa préface (écrite seulement en hongrois), Pokorny écrit que les autres dictionnaires français–hongrois, parus avant le sien, n'offrent que de simples « nomenclatures », plus ou moins volumineuses. Comme il écrit, présenter des mots-vedettes accompagnés de simples équivalents peut rendre l'utilisation d'un dictionnaire confuse. Selon lui, c'est aussi la raison pour laquelle les dictionnaires français–allemand étaient tellement populaires en Hongrie à l'époque. Son but était donc de créer un dictionnaire français–hongrois et hongrois–français grâce auquel les dictionnaires allemands ne seraient plus indispensables. Pokorny met en relief la présence d'exemples d'usage (ayant pour but d'illustrer aussi bien des expressions figées que de différents styles) et des informations grammaticales dans son dictionnaire.

Pokorny mentionne également qu'il s'est servi des ouvrages français « les plus récents et les plus renommés », tels que « le Dictionnaire de l'Académie, les dictionnaires de Littré, de Poitèvin et de Landais, mais surtout le Bescherelle ». Malheureusement, Pokorny n'indique pas les données exactes de ces ouvrages, même pas leurs titres. Il écrit aussi que c'était à l'aide du « dictionnaire de Sachs

(Toussaint-Langenscheit [sic]), tant recherché en Allemagne », qu'il a structuré son dictionnaire, mais il ne précise pas à quel point et de quelle façon il l'a employé, et, tout comme dans le cas des autres ouvrages, il ne précise pas les données bibliographiques. On ne peut que deviner que Pokorny se réfère au *Dictionnaire encyclopédique français-allemand et allemand-français* par Karl Sachs et Césaire Villatte.

Le nombre des entrées du dictionnaire de Pokorny dépasse les 50 mille.

Ce dictionnaire de Pokorny a été publié en trois volumes, et il a connu quatre éditions, celle de 1880, une deuxième en 1881, une troisième en 1885 et une quatrième en 1892 (à l'occasion de la parution des deux volumes de la partie hongrois-français en 1891-1892).

### 2.1.5 THEISZ 1902

Gyula Theisz (Jules Theisz), le rédacteur du dictionnaire paru en 1902,<sup>2</sup> prétend dans sa préface (écrite seulement en langue hongroise) que son dictionnaire est le premier ouvrage qui se base sur le propre travail du rédacteur, c'est-à-dire lui-même. En même temps, Theisz ne mentionne même pas le nom de Pokorny (ni celui de Babos, ni celui de Mártonffy d'ailleurs) quand il parle des dictionnaires français-hongrois précédents, il fait seulement mention des noms des rédacteurs allemands et français dont les dictionnaires – selon lui – servaient de bases pour les dictionnaires français-hongrois. Theisz parle de Molé et de Sachs. Dans le cas de Molé, il est facile de deviner qu'il s'agit du dictionnaire de BABOS (1865), mais dans le cas de Sachs, ce n'est pas si évident qu'il s'agirait du dictionnaire de Pokorny. Theisz donne quelques exemples des fautes que l'on peut trouver dans les dictionnaires français-hongrois précédents, mais les fautes d'équivalence qu'il présente pour le dictionnaire de Sachs, ne sont que partiellement retrouvables dans le dictionnaire de Pokorny. De plus, en lisant la préface de Theisz, on ne peut identifier le dictionnaire de Mártonffy non plus.

Ce qui est sûr et certain, c'est que le dictionnaire de Theisz est le fruit d'un travail autonome se concentrant sur les particularités du hongrois et du français, et sur les particularités qui s'émergent si l'on va depuis le français directement vers le hongrois. En même temps, il faut remarquer que le dictionnaire de Theisz, quant au nombre des entrées – n'étant qu'environ 26 mille –, n'est pas parmi les ouvrages les plus riches.

Ce dictionnaire de Theisz n'a eu qu'une seule édition.

---

<sup>2</sup>L'année de la parution n'est pas indiquée dans le dictionnaire de Theisz. Ce n'est qu'à la fin de la préface de Theisz que l'on trouve en hongrois « au mois de décembre 1902 ».

### 2.1.6 SAUVAGEOT 1932

C'est avec le grand dictionnaire d'Aurélien Sauvageot que la lexicographie franco-hongroise prendra une tout autre direction, notamment une bonne. Pour la première fois, le rédacteur est un véritable savant,<sup>3</sup> et la taille du dictionnaire a considérablement augmenté par rapport aux dictionnaires précédents – non seulement quant à la nomenclature – le nombre des entrées y atteint les 70 mille –, mais aussi au niveau de la richesse des entrées en exemples. Même du point de vue lexicographique, il s'agit d'un grand pas en avant : la microstructure des entrées a beaucoup évolué. Ce qui est très surprenant, c'est que ce dictionnaire n'a eu que deux éditions : la première en 1932 et la seconde en 1942.

### 2.1.7 ECKHARDT 1953

Le dictionnaire qui a pris la place de celui de Sauvageot, qui l'a complètement remplacé, c'est celui de Sándor Eckhardt (Alexandre Eckhardt), dont la première édition date de 1953. À notre avis, cela peut s'expliquer par la situation politique après la seconde guerre mondiale. Sous le régime communiste, les représentants de la vie académique cherchaient – ils étaient obligés de faire ainsi – à atteindre des résultats significatifs, à créer « de grandes choses » dans tous les domaines possibles, tout en rompant avec les réalisations et les réussites liées à des représentants de l'Ouest. Eckhardt avait l'intention d'écrire son propre dictionnaire depuis les années 1920 (voir p.ex. SAUVAGEOT 1932 : 165). Il a finalement créé son dictionnaire énorme, avec 135 mille entrées, avec des entrées très-très riches en exemples, mais avec un grand inconvénient : le manque total de collaborateurs français a automatiquement entraîné un certain manque de fiabilité dans le cas de certaines entrées.<sup>4</sup>

---

<sup>3</sup>Sauvageot était le premier titulaire de la chaire des langues finno-ougriennes à l'École nationale des langues orientales vivantes, aujourd'hui appelée INALCO – Institut national des langues et civilisations orientales. Il a occupé ce poste depuis la création de la chaire en 1931 jusqu'à sa retraite en 1967 (avec une interruption entre 1941 et 1943). Sauvageot raconte l'histoire de son dictionnaire dans son livre intitulé *Souvenirs de ma vie hongroise*. On y apprend – entre autres – qu'il était choqué de voir ses élèves hongrois utiliser un dictionnaire français–allemand, les dictionnaires français–hongrois étant jugés tellement insuffisants. Dans un premier temps, il n'était pas spécialement tenté d'assumer la tâche de la rédaction d'un dictionnaire, qui lui a été proposée par Marcell Benedek – avec la collaboration de qui, et avec celle de József Balassa, il a finalement rédigé son grand dictionnaire. Sauvageot raconte également qu'il avait plus d'une fois parlé de la question du dictionnaire avec Sándor Eckhardt (voir Section 2.1.8) qui lui disait toujours qu'il envisageait de produire un jour un dictionnaire français–hongrois, mais Sauvageot comprenait au ton d'Eckhardt que ce n'était pas pour lui un projet très pressant.

<sup>4</sup>Dans le cas de plusieurs entrées, Eckhardt n'avait pas le choix, quand il ne connaissait pas l'équivalent, il essayait de deviner et d'inventer des choses, ce qui est encore plus frappant

Ce dictionnaire d'Eckhardt a eu une édition revue et augmentée en 1960, suivie par une troisième édition complétée par une annexe en 1973, réimprimée en 1981, en 1987, en 1990 et finalement en 1992.

### 2.1.8 ECKHARDT & OLÁH 1999

À la fin des années 1990, de la part de la maison d'édition liée à l'Académie hongroise des sciences, une équipe d'une trentaine de chercheurs a eu la tâche de renouveler les dictionnaires d'Eckhardt sous la direction de Tibor Oláh. Le problème, c'est que les rédacteurs n'avaient que très peu de temps, un an seulement, à leur disposition. Ils ont donc enlevé des entrées, ils en ont rajouté d'autres. Malheureusement, ils ont aussi enlevé des mots-vedettes qu'ils n'auraient pas forcément dû enlever (surtout des termes techniques, des mots liés à la faune et à la flore), alors qu'ils n'ont pas réussi à corriger un grand nombre d'erreurs et de fautes au niveau des équivalents, et ils en ont même rajouté quelques-unes. Le nombre total des entrées y était à peu près conservé, soit toujours environ 135 mille.

La première édition a été réimprimée en 2001.

### 2.1.9 PÁLFY 1999

1999 est l'année de la parution d'un autre dictionnaire français–hongrois, qui a vu le jour sous la direction de Miklós Pálffy, professeur à l'Université de Szeged en Hongrie à l'époque. Il s'agissait alors d'un dictionnaire tout nouveau, dont les travaux de rédaction sont partis de zéro. Le résultat en a été un dictionnaire très bien réussi, de taille moyenne, avec un bon choix de nomenclature comprenant 35 mille entrées assez riches en exemples. De plus, ce dictionnaire peut être considéré à la fois actif et passif, incluant des informations supplémentaires utiles destinées non seulement aux utilisateurs ayant le hongrois comme langue maternelle, mais aussi aux utilisateurs francophones.

Des éditions revues ont paru à partir de 2001, la septième en 2009.

### 2.1.10 BÁRDOSI & SZABÓ 2007

Le dictionnaire paru en 2007 est celui rédigé par Vilmos Bárdosi et Dávid Szabó. Il s'agit, bien évidemment, d'un dictionnaire moins riche que celui d'Eckhardt,

---

dans le cas de son dictionnaire hongrois–français. Si l'on regarde par exemple l'équivalent du mot hongrois *sütőpor*, on trouve *poudre gazifère* / *gazogène* (quelque chose qui s'utilisait dans la deuxième moitié du XIXe siècle et au début du XXe pour faire de l'eau gazeuse, pétillante), alors qu'il s'agit de ce que l'on appelait à l'époque en France *la levure chimique* (!) – aujourd'hui on préfère *poudre à lever*, peut-être parce que le mot *chimique* n'a pas de bonnes connotations.

mais plus riche que celui de Pálffy quant au nombre des entrées, soit à peu près 51 mille mots-vedettes. Au niveau du nombre des entrées, il s'agit également d'un dictionnaire de taille moyenne, mais avec un excellent choix de nomenclature, comprenant des mots et des expressions issues de toute la francophonie. Tout comme dans le dictionnaire de Sauvageot, on y trouve la prononciation de chaque mot-vedette – quelque chose qui manque dans les Eckhardt et le Pálffy. Les articles comprennent suffisamment d'exemples pour illustrer l'usage général des mots-vedettes. Il faut remarquer que Bárdosi est un grand expert de la phraséologie et de la parémiologie (il a publié plusieurs recueils de proverbes et des dictionnaires phraséologiques français–hongrois), Szabó est expert de l'argot, ce qui explique la quantité élevée d'expressions de tels types parmi les exemples. En même temps, il faut aussi préciser qu'il ne s'agit pas d'une quantité exagérée, les rédacteurs ont donc réussi à trouver le juste milieu.

Ce dictionnaire a été réimprimé quatre fois jusqu'à présent, la dernière fois en 2018 (soit une cinquième édition).

### 2.1.11 Les ouvrages exclus de nos analyses

Les ouvrages dont notre étude ne traite pas sont d'une part des dictionnaires français–hongrois que nous avons choisi d'exclure soit en raison de leur volume et de leur format (avec une quantité modérée d'entrées), soit parce qu'ils se concentrent sur un lexique spécialisé. D'autre part, nous avons également exclu les ouvrages qui contiennent seulement des listes de vocabulaire avec des lexèmes français et hongrois, comme c'est le cas, par exemple, de certaines grammaires. La raison en est simple : nous avons choisi d'examiner les véritables dictionnaires en fonction de la microstructure de leurs entrées et les exemples fournis, alors que les simples listes de vocabulaire se caractérisent – entre autres – par l'absence des exemples d'usages. En ce qui concerne les dictionnaires spécialisés, étant donné qu'ils ne traitent que d'une certaine couche définie du lexique, leurs contenus et leurs structures seraient difficilement comparables à ceux des dictionnaires généraux. Finalement, les dictionnaires généraux de petite taille ont pour caractéristique d'être restreints non seulement au niveau de la taille de la nomenclature, mais aussi pour ce qui est de la richesse des entrées en exemples d'usage. (Comme nous l'avons déjà précisé dans Section 1.2, les dictionnaires hongrois–français ne font pas non plus partie de cette étude.)

Pour une liste complète des dictionnaires français–hongrois consulter JEZERNICZKY 1933 (pour les parutions jusqu'à 1848), SÁGI 1922 et MAGAY 2011. Dans ces bibliographies, on trouve tous les dictionnaires français–hongrois et hongrois–français, et même d'autres ouvrages (p.ex. des grammaires) contenant des listes

de vocabulaire, en revanche, ces derniers sont loin d'être inventoriés d'une manière exhaustive.

## 2.2 Méthodes : À la recherche des exemples

### 2.2.1 Qu'est-ce qu'un exemple ?

En ce qui concerne la définition du terme « exemple », on peut constater une grande variété d'interprétations, et beaucoup ont déjà remarqué qu'il n'existe pas de convention unanime. Cela a été souligné par Jacobsen et ses coauteurs (1991) aussi, qui expliquent en même temps pourquoi les exemples que l'on a l'habitude d'appeler ainsi ne sont pas des exemples. Selon eux, suivant les différentes fonctions des différentes informations fournies par les entrées, on peut catégoriser et nommer ces informations par certains termes, mais un « exemple » n'est qu'un supplément à un équivalent de traduction, fournissant des informations implicites sur l'équivalent ou le mot-vedette.

Nous continuons cependant à nous servir d'une terminologie plus classique, plus proche de celle suivie par SVENSÉN (2009). Pour nous, il s'agit donc de tous les éléments supplémentaires dans la langue source qui comprennent une certaine forme du mot-vedette, ayant pour but d'illustrer l'usage du mot-vedette en question par rapport à un sens donné, ou bien de présenter une unité lexicale complexe, une collocation, une locution, etc., comprenant une certaine forme du mot-vedette.

Quels peuvent être les différents types d'exemples ?

- phrases (propositions) / citations
- collocations / locutions / expressions
- unités phraséologiques / phrasèmes (expressions figées : proverbes, dictons, devises, slogans)
- unités lexicales complexes

Dans plusieurs cas, Jacobsen et ses coauteurs (1991) préfèrent le terme « sous-entrée ».

Dans nos analyses que nous présentons ici, nous nous concentrons sur certains verbes (voir Section 2.2.2). Pour ce qui est de cette catégorie, on peut faire la distinction entre deux types principaux d'exemples dans les entrées de dictionnaire :

Type 1 : Formes neutres à l’infinitif

- Type 1a : *aller en voiture* (non-figé)
- Type 1b : *s’en aller* (figé)

Type 2 : Formes conjuguées / propositions (syntagmes verbaux)

- Type 2a : *je vais à Lyon* (non-figé)
- Type 2b : *ça va ?* (figé)

FIGURE 1 : Schématisation des différents types d’exemples pour la catégorie des verbes.

Pour d’autres aspects de la définition et du choix des exemples, voir p.ex. al-KASIMI 1983, PIOTROWSKI 2000, NIELSEN 2014, KARPINSKA 2015 et BELLEFLAMME 2019.

### 2.2.2 Nos méthodes

Pour nos analyses, nous avons choisi d’examiner des entrées présentant des verbes fréquents et sémantiquement surchargés, notamment *aller*, *avoir*, *devoir*, *être*, *faire*, *laisser*, *mettre*, *pouvoir*, *prendre*, *rendre* et *venir*, soit des mots-vedettes dont l’emploi est le plus souvent illustré par de nombreux exemples d’usage – aussi bien en raison de la polysémie qu’en raison de la fréquence élevée de ces verbes dans de différents contextes.

Nous avons compté les exemples dans les onze entrées mentionnées de tous les dix dictionnaires présentés dans Section 2.1. Nous avons également compté combien d’exemples appartiennent au type T1 et combien au T2 par entrée et par dictionnaire. En tant que référence, nous avons aussi analysé le nombre des exemples dans certaines de nos entrées choisies dans d’autres dictionnaires bilingues. Les statistiques de nos résultats seront présentés dans la partie 3.

Nous analysons la microstructure des entrées par dictionnaire (voir Section 3.1–3.10), décrivant surtout la structuration des sens et la présentation des exemples et des équivalents, mais nous tenons également compte de la prononciation éventuellement indiquée des mots-vedettes, ainsi que des autres types d’informations fournies par les entrées. Nous étudions également des exemples concrets surtout issus de l’entrée *aller* (typiquement riche en exemples, ce verbe ayant de nombreux sens) de tous les dictionnaires analysés.

Finalement, nous présentons les problèmes des équivalents et des exemples dans les entrées *avoir*, *devoir* et *pouvoir* de nos dictionnaires – des verbes français dont le contenu sémantique peut s'exprimer dans la langue hongroise par des constructions grammaticales très différentes de celles utilisées avec ces verbes français (3.11).

### 3 Étude comparative des dictionnaires

Le Tableau 1 présente le nombre des exemples dans les entrées choisies dans les différents dictionnaires parus entre 1844 et 1902, alors que le Tableau 2 indique les chiffres pour les dictionnaires parus entre 1932 et 2007.

On peut voir que le nombre des exemples a augmenté avec le temps de manière constante dans les dictionnaires du XIXe siècle – en parallèle d'ailleurs avec le nombre des entrées. Cette tendance continue même au XXe siècle avec le dictionnaire de Sauvageot et celui d'Eckhardt – le dictionnaire de Theisz représentant une exception. La fin du XXe et le début du XXIe siècle se caractérisent par la parution de dictionnaires qui ne sont pas de grand format (Pálfy, Bárdosi – Szabó), mais qui sont très riches en exemples. Le dictionnaire de Pálfy représente un cas exceptionnel : le nombre de ses entrées représente à peu près un quart du dictionnaire d'Eckhardt, mais cela ne se reflète pas du tout dans le nombre des exemples dans ses entrées.

Le Tableau 3 montre le nombre des exemples par types dans les dictionnaires de la période 1844–1902. Le premier chiffre indique la quantité des exemples de T1, alors que le second chiffre représente le T2 (voir aussi Figure 1 sous Section 2.2.1). Nous avons ombré les cases avec du gris clair quand il y a plus d'exemples avec des infinitifs (T1) et avec du gris foncé quand il y a plus de propositions ou phrases avec des formes conjuguées (T2). Le blanc (sans ombrage) signifie qu'il n'y a pas de différence remarquable entre les deux types, ou bien il n'y a aucun type d'exemple pour le verbe donné. On peut constater que – de manière générale – les différents dictionnaires ont les mêmes tendances pour un verbe donné, c'est-à-dire la préférence pour un type varie en fonction des verbes. Si cette tendance est rompue quelques fois, cela se constate toujours dans les deux premiers dictionnaires qui montrent une préférence générale pour les exemples de T1 – si des exemples sont fournis du tout. Ainsi, les dictionnaires préfèrent T1 pour les verbes *faire*, *laisser*, *mettre*, *prendre* et *rendre*, alors qu'ils préfèrent le T2 pour *aller*, *avoir*, *devoir*, *être*, *pouvoir* et *venir*. Nous avons trouvé le même schéma dans les autres dictionnaires de la période 1932–2007 (sans donner les chiffres concrets ici).

TABLE 1 : Le nombre des exemples dans les entrées choisies des dictionnaires parus entre 1844 et 1902.

|                | KISS<br>(1844) | BABOS<br>(1865) | MÁRTONFFY<br>(1879) | POKORNY<br>(1880) | THEISZ<br>(1902) |
|----------------|----------------|-----------------|---------------------|-------------------|------------------|
| <i>aller</i>   | 4              | 32              | 39                  | 99                | 74               |
| <i>avoir</i>   | 1              | 9               | 13                  | 33                | 20               |
| <i>devoir</i>  | 0              | 0               | 5                   | 15                | 6                |
| <i>être</i>    | 0              | 0               | 6                   | 77                | 54               |
| <i>faire</i>   | 0              | 2               | 46                  | 172               | 31               |
| <i>laisser</i> | 0              | 0               | 4                   | 15                | 15               |
| <i>mettre</i>  | 1              | 1               | 27                  | 83                | 32               |
| <i>pouvoir</i> | 0              | 2               | 7                   | 22                | 11               |
| <i>prendre</i> | 0              | 5               | 61                  | 140               | 39               |
| <i>rendre</i>  | 1              | 5               | 22                  | 35                | 15               |
| <i>venir</i>   | 0              | 15              | 17                  | 65                | 29               |

TABLE 2 : Le nombre des exemples dans les entrées choisies des dictionnaires parus entre 1932 et 2007.

|                | SAUVAGEOT<br>(1932) | ECKHARDT<br>(1953) | ECKHARDT &<br>OLÁH (1999) | PÁLFY<br>(1999) | BÁRDOSI &<br>SZABÓ (2007) |
|----------------|---------------------|--------------------|---------------------------|-----------------|---------------------------|
| <i>aller</i>   | 125                 | 196                | 237                       | 134             | 58                        |
| <i>avoir</i>   | 70                  | 120                | 127                       | 77              | 59                        |
| <i>devoir</i>  | 35                  | 49                 | 29                        | 42              | 19                        |
| <i>être</i>    | 118                 | 201                | 161                       | 52              | 69                        |
| <i>faire</i>   | 282                 | 240                | 357                       | 136             | 131                       |
| <i>laisser</i> | 74                  | 98                 | 117                       | 52              | 44                        |
| <i>mettre</i>  | 135                 | 300                | 328                       | 206             | 60                        |
| <i>pouvoir</i> | 18                  | 41                 | 40                        | 23              | 19                        |
| <i>prendre</i> | 169                 | 234                | 197                       | 138             | 91                        |
| <i>rendre</i>  | 69                  | 103                | 85                        | 44              | 31                        |
| <i>venir</i>   | 85                  | 155                | 148                       | 61              | 35                        |

Pour pouvoir comparer les chiffres des dictionnaires hongrois avec d'autres dictionnaires bilingues ayant le français comme langue source, nous avons fait le Tableau 4, avec des dictionnaires qui sont parmi les plus complets pour les paires de langues en question. Pour le norvégien, nous avons pris en compte deux dictionnaires qui ont à peu près le même nombre d'entrées : le premier chiffre correspond à ELLIGERS & REICHBORN-KJENNERUD 1995 et le second à JACOBSEN 2014.

Entre autres, on peut voir que les dictionnaires français–hongrois sont beaucoup plus riches en exemples que les dictionnaires français–norvégien ou français–suédois. Le cas du français–danois est une véritable exception pour les langues scandinaves : dans ce cas il s'agit des chiffres du fameux dictionnaire d'Andreas Blinkenberg et Poul Høybye, qui est un des dictionnaires bilingues les plus volumineux pour le français. Quant au norvégien, au fait, nous avons choisi d'inclure les deux dictionnaires, puisqu'il peut y avoir des différences considérables au niveau de la richesse en exemples, et cela se montre parfaitement dans notre tableau p.ex. pour les entrées *avoir*, *mettre* ou *prendre*.

### 3.1 KISS 1844

Dans le dictionnaire de Kiss, les mots-vedettes français commencent par une majuscule, et ils sont suivis par une abréviation (française) en italique indiquant la catégorie grammaticale du mot-vedette. Le mot-vedette et l'abréviation se séparent en général par une virgule, mais l'ouvrage n'est pas conséquent, parfois on y trouve un point ou rien du tout. Ensuite viennent les équivalents hongrois, séparés par une virgule ou par un point-virgule – il est difficile de trouver la logique du choix entre ces deux.

Les équivalents du mot-vedette sont éventuellement suivis par des exemples d'usages de différents types avec leur traduction hongroise. Chez Kiss, on ne trouve que des exemples de T1.

En ce qui concerne la forme des équivalents, les verbes hongrois sont indiqués à l'indicatif présent (de la conjugaison appelée « indéfinie » ou « générale ») de la 3e personne du singulier<sup>5</sup> – c'est ce qui est de norme dans les dictionnaires hongrois de nos jours depuis une centaine d'années, mais à l'époque, ce n'était pas ainsi. Dans les tout premiers dictionnaires hongrois, surtout dans les bilingues en paire avec le latin, on trouve les verbes hongrois à la 1ère personne du singulier, alors que dans les ouvrages où le hongrois est en paire avec l'allemand,

---

<sup>5</sup>Cette forme du verbe n'est marquée par aucun suffixe, c'est pourquoi elle sert de point de départ dans la conjugaison. L'utilisation de cette forme au lieu de l'infinitif peut ainsi aider les apprenants du hongrois.

TABLE 3 : Le nombre des exemples par types (T1 / T2) dans les entrées choisies des dictionnaires parus entre 1844 et 1902.

|                | KISS<br>(1844) | BABOS<br>(1865) | MÁRTONFFY<br>(1879) | POKORNY<br>(1880) | THEISZ<br>(1902) |
|----------------|----------------|-----------------|---------------------|-------------------|------------------|
| <i>aller</i>   | 4 / 0          | 16 / 16         | 15 / 24             | 48 / 51           | 30 / 44          |
| <i>avoir</i>   | 1 / 0          | 5 / 4           | 4 / 9               | 5 / 28            | 0 / 20           |
| <i>devoir</i>  | 0 / 0          | 0 / 0           | 1 / 4               | 4 / 11            | 2 / 4            |
| <i>être</i>    | 0 / 0          | 0 / 0           | 1 / 5               | 20 / 57           | 8 / 46           |
| <i>faire</i>   | 0 / 0          | 1 / 1           | 31 / 15             | 117 / 55          | 20 / 11          |
| <i>laisser</i> | 0 / 0          | 0 / 0           | 3 / 1               | 10 / 5            | 11 / 4           |
| <i>mettre</i>  | 1 / 0          | 1 / 0           | 27 / 0              | 82 / 1            | 32 / 0           |
| <i>pouvoir</i> | 0 / 0          | 0 / 2           | 1 / 6               | 1 / 21            | 0 / 11           |
| <i>prendre</i> | 0 / 0          | 3 / 2           | 37 / 24             | 100 / 40          | 32 / 7           |
| <i>rendre</i>  | 1 / 0          | 4 / 1           | 11 / 11             | 26 / 9            | 14 / 1           |
| <i>venir</i>   | 0 / 0          | 8 / 7           | 6 / 11              | 20 / 45           | 13 / 16          |

TABLE 4 : Le nombre des exemples dans certaines entrées choisies dans certains dictionnaires bilingues.

|                | anglais | danois | catalan | norvégien | suédois |
|----------------|---------|--------|---------|-----------|---------|
| <i>aller</i>   | 240     | 144    | 60      | 34/35     | 38      |
| <i>avoir</i>   | 148     | 102    | 35      | 20/7      | 23      |
| <i>être</i>    | 138     | 200    | 94      | 10/11     | 28      |
| <i>faire</i>   | 198     | 290    | 137     | 21/32     | 56      |
| <i>mettre</i>  | 179     | 67     | 92      | 45/25     | 35      |
| <i>prendre</i> | 206     | 221    | 145     | 57/29     | 54      |

les verbes hongrois se présentent d'habitude à l'infinitif. Ce sont aussi des infinitifs qui apparaissent dans les dictionnaires français-hongrois du XIXe siècle, mais non chez Kiss. En même temps, il y a une entrée, notamment *avoir*, où Kiss donne un infinitif hongrois aussi (*lenni* outre la forme conjuguée *van*) – mais tout cela en vain, car les équivalents restent défectifs, étant donné qu'ils signifient simplement < être > (voir Section 3.11). D'ailleurs, justement dans le cas d'*avoir*, Kiss donne aussi l'infinitif allemand *haben* en tant qu'équivalent.

L'infinitif de plusieurs des verbes français de notre étude s'utilise également en tant que nom. Chez Kiss, on trouve parfois des entrées séparées pour les mots homonymiques (*aller*, *avoir* et *pouvoir*), alors que d'autres fois les homonymes sont traités à l'intérieur d'une même entrée (*devoir* ou *être*). Justement dans le cas d'*avoir* et de *devoir*, il serait fort souhaitable de fournir des exemples de T2, puisque le hongrois exprime les sens principaux de ces verbes français d'une manière complètement différente (voir Section 3.11).

Parmi les entrées étudiées, c'est *aller* où l'on trouve le plus d'exemples chez Kiss, il y en a quatre. Ces exemples sont les suivants : *aller à cheval*; *aller en carrosse* [sic]; *aller faire qc.*; *se laisser aller sur....* Le choix des trois premiers est tout à fait compréhensible du point de vue de l'usage de l'époque, en revanche, le quatrième exemple nous semble trop spécifique – vu le nombre total des exemples. Si l'on étudie les équivalents hongrois de ces quatre exemples, les deux premiers sont facilement compréhensibles. Le troisième, qui exemplifie surtout l'usage d'*aller* en tant qu'auxiliaire, est cependant expliqué par un équivalent hongrois qui signifie < commencer à faire qc >. Le quatrième exemple est le plus problématique même au niveau des équivalents : on n'y trouve qu'un seul verbe hongrois « *roskad* » qui peut avoir plusieurs sens, et nous avons du mal à en identifier un qui pourrait refléter l'exemple français. Dans ce cas, l'utilisateur du dictionnaire aurait vraiment besoin d'un exemple de T2, et une bonne traduction hongroise comme équivalent – qu'il s'agisse de quelqu'un apprenant le hongrois ou d'un(e) Hongrois(e) apprenant le français.

Le Kiss est un dictionnaire pauvre en exemples. Parmi les articles traitant des onze verbes de notre étude, il n'y a que quatre contenant des exemples d'usage, et trois sur ces quatre entrées ne présentent qu'un seul exemple.

L'insuffisance des équivalents reflétant les différents sens d'un mot-vedette donné, et l'absence presque totale d'exemples d'usage rendent l'utilisation de ce dictionnaire vraiment difficile, même aux Hongrois à qui cet ouvrage a probablement été destiné.

Finalement, il est intéressant de remarquer que le style typographique des mots français et celui des mots hongrois ne sont pas identiques, ce qui facilite l'usage

du dictionnaire. La différence se constate surtout au niveau des espaces entre les lettres, les lettres étant plus éloignées les unes des autres dans les mots français.

### 3.2 BABOS 1865

Tout comme dans le dictionnaire de Kiss, les mots-vedettes du dictionnaire commencent par une majuscule, et sont suivis par une virgule, ensuite par une abréviation (française) en italique indiquant la nature du lemme. Les équivalents hongrois sont tous en italique, séparés par de simples virgules s'il s'agit de synonymes (représentant à peu près le même sens), alors que l'équivalent appartenant à un autre champ sémantique se sépare des autres équivalents par un point-virgule. On trouve un point-virgule devant les exemples aussi.

Chez Babos, l'homonymie éventuelle est toujours traitée à l'intérieur d'une même entrée.

Les verbes hongrois (servant d'équivalents) sont indiqués à l'infinitif, ce qui suit la norme des dictionnaires allemand–hongrois de l'époque.

En ce qui concerne les exemples, parmi les entrées étudiées, c'est le verbe *aller* qui est le mieux décrit par le plus d'exemples dans le Babos aussi – on y trouve 32 exemples (il y en avait 4 dans le Kiss). En revanche, il n'est pas toujours clair comment il faut comprendre certains exemples, car Babos n'y ajoute aucune marque d'usage désignant le champ d'emploi de l'exemple. Par exemple, dans l'article *aller*, on trouve *de combien allez-vous* [?], mais cette question est difficilement interprétable, si l'on ne sait pas qu'il faudrait penser à un jeu (auquel on joue en argent). Quant à *avoir*, le sens principal < posséder > est aussi mal rendu par les simples équivalents insuffisants que dans le Kiss. De plus, les exemples *avoir chaud* et *avoir froid* sont mal traduits – les équivalents hongrois signifient < être chaud/froid >.

Dans le Babos, la proportion des exemples de T1 et de T2 est assez en balancée, mais T1 reste dominant.

Tout compte fait, du point de vue de la microstructure, le dictionnaire de Babos est beaucoup plus conséquent que celui de Kiss. On peut trouver plus d'exemples d'usage dans le Babos, mais leur nombre est loin d'être suffisant. Parmi les onze entrées de notre étude, il y a trois dans lesquelles on ne trouvera aucun exemple, et dans six articles le nombre des exemples n'atteint pas les dix. Les exemples choisis ne reflètent pas toujours les sens les plus fréquents ou les plus problématiques à expliquer du point de vue de la langue hongroise. Le Babos a été probablement destiné aux utilisateurs hongrois, mais même eux, ils devaient avoir du mal à « décoder » des sens fréquents en raison de l'insuffisance des équivalents et des exemples dans certains cas.

### 3.3 MÁRTONFFY 1879

Dans le Mártonffy, les mots-vedettes commencent par une majuscule – comme chez Kiss et chez Babos. Ce qui est novateur et fort appréciable dans ce dictionnaire, c'est que la prononciation des lemmes français est indiquée entre parenthèses après les mots-vedettes. Même s'il ne s'agit que d'une prononciation approximative rendue à l'aide d'une transcription proche de l'écriture de la langue hongroise, cela devait représenter une grande aide aux utilisateurs hongrois. Une abréviation (française, cette fois-ci non en italique) indique la catégorie grammaticale du lemme. Les équivalents hongrois sont tous en italique, comme dans le Babos.

Une autre innovation du Mártonffy est que les différents sens et/ou les différents types de fonction du lemme sont numérotés, ce qui facilite l'orientation à l'intérieur de l'entrée. Par exemple, dans l'article *aller*, on trouve 15 différents « sens » de ce verbe. Les sens et les exemples sont toujours précédés d'un point-virgule qui les sépare de l'information précédente, alors que les exemples et leurs traductions se séparent par de simples virgules.

Les verbes hongrois (servant d'équivalents) sont indiqués à l'infinitif.

Chez Mártonffy, tout comme chez Babos, l'homonymie est traitée dans un même article. P.ex. après les 15 sens du verbe *aller*, le sens du nom *aller* a reçu le numéro 16.

Pour ce qui est des exemples, ils sont beaucoup plus nombreux que dans le Babos. Pour le verbe *aller*, on y en trouve 39, mais pour deux de nos verbes étudiés il y en a encore plus : 46 pour *faire* et 61 pour *prendre*.

Dans le Mártonffy, la proportion des exemples de T1 et de T2 n'est pas homogène du tout, elle varie en fonction des entrées – ce qui caractérisera d'ailleurs les dictionnaires français–hongrois du XXe siècle aussi. Le T1 reste prédominant dans le cas de *faire*, *laisser*, *mettre* et *prendre*, tandis que les exemples de T2 sont plus fréquents dans les articles *aller*, *avoir*, *devoir*, *être*, *pouvoir* et *venir*. Chez Mártonffy, dans le cas de *rendre*, le nombre des T1 et celui des T2 sont en égalité, mais ce verbe se caractérise plutôt par l'usage dominant du T1 dans les dictionnaires postérieurs.

Quant à l'innovation de numéroté les différents sens / types de fonction (que nous avons déjà mentionnée), sa réalisation n'est pas la mieux réussie, étant donné que sa logique reste souvent obscure. Il arrive que l'on tombe sur des exemples et des équivalents sémantiquement bien différents, mais présentés sous le même numéro. Pour ne citer que quelques cas, dans l'article *aller* seul, on trouve trois « sens » problématiques. Sous le numéro 7, on trouve les exemples suivants (que nous allons marquer par des minuscules différentes ici : a) *il y va*

avec prudence; b) *je vais lui écrire*; c) *voyons ce qu'il va dire*; d) *aller voir q.*; e) *n'allez pas croire*. On voit bien que ce sont seulement b et c qui exemplifient le même type de fonction du mot-vedette. Sous le numéro 14, on trouve trois exemples : a) *il s'en va 2 heures*; b) *comment vous va ?*; c) *cela va bien*. On constate alors que b et c sont liés l'un à l'autre, mais a exemplifie quelque chose de très différent. Finalement, nous mentionnons les exemples sous le numéro 15 : a) *aller par dessus le marché*; b) *aller selon le vent*; c) *allez-vous par terre ou par mer ?*. Dans ce cas, le verbe *aller* s'utilise dans des expressions très différentes, et il peut être difficile de trouver ce qui est commun dans les différents usages.

Nous voudrions attirer l'attention à l'expression c, dont la traduction hongroise signifie « mangez-vous de la viande ou non ? ». En effet, on peut identifier le sens métaphorique de la partie « par terre ou par mer » de l'expression : c'est comme demander à quelqu'un « mangez-vous de la viande ou du poisson ? ». En revanche, cet exemple du dictionnaire présente une expression qui n'est pas du tout répandue en France, et, selon nos sources, elle ne l'a jamais été. En fait, nous n'avons réussi à retrouver aucune trace de cette « pseudo-expression » française, à notre avis, il s'agit d'un hapax intéressant.

### 3.4 POKORNY 1880

Dans le Pokorny, comme dans les dictionnaires antérieurs, les mots-vedettes commencent par une majuscule, et, ce qui est nouveau et très efficace du point de vue du visuel, ils sont composés en gras. Les mots-vedettes sont suivis par une prononciation approximative entre parenthèses comme dans le Mártonffy, et d'une abréviation indiquant la catégorie grammaticale.

Pokorny – tout comme Mártonffy – a numéroté les sens / les types de fonctions, p.ex. de 1 à 25 pour *aller*. Les homonymes éventuels, ainsi que la forme pronominale des verbes, apparaissent comme des entrées à part. C'est la raison pour laquelle *s'en aller* apparaît comme une entrée indépendante après *aller*, avec ses différents sens numérotés de 1 à 9, et, ensuite, on trouve *aller* en tant que nom. La typographie des verbes pronominaux est un peu bizarre : ils apparaissent notamment avec le pronom *se* en minuscules, mais le verbe de base commence toujours par une majuscule : p.ex. *s'en aller* apparaît sous la forme *s'en Aller*.

On peut aussi constater un nouveau phénomène lexicographique dans le Pokorny : les équivalents hongrois sont précédés d'un signe d'égalité. On trouve ce signe toujours devant le numéro du premier sens, et après tous les exemples d'usage. Les équivalents hongrois qui apparaissent après les numéros des différents sens sont en italique, on peut alors se demander pourquoi Pokorny trouvait nécessaire de marquer le passage depuis le français vers le hongrois d'une

telle manière, si la différence entre les mots français et les mots hongrois est conséquemment signalée par les différents styles typographiques. En revanche, si l'on contemple tous les types d'équivalents, on constate que ceux qui suivent les exemples d'usages ne sont pas en italique, et, dans ces cas, le texte hongrois a exactement le même style typographique que le texte français – c'est ce qui rend nécessaire l'usage des signes d'égalité. Si tous les équivalents hongrois avaient été composés en italique, il n'y aurait pas eu besoin de ces signes d'égalité, mais il paraît que Pokorny voulait séparer visuellement les deux types d'équivalents.

Dans certaines entrées, pour certains sens du mot-vedette français, on trouve non seulement des équivalents hongrois, mais aussi des explications courtes en hongrois, qui peuvent aider les utilisateurs hongrois du dictionnaire à mieux décoder un sens donné du mot-vedette. Ces explications en langue hongroise apparaissent cependant en italique, tout comme les équivalents des différents sens, et, ainsi, elles sont fort trompeuses du point de vue des utilisateurs francophones qui cherchent à se servir de cet ouvrage en tant que dictionnaire d'encodage. Voir p.ex. le sens numéro 8 dans l'entrée du verbe *aller* : on y trouve l'explication *Állapot, helyzet* < état, situation > en hongrois, et rien ne montre que ce ne sont pas de simples équivalents. Dans le même article, on trouve des explications trompeuses de ce type sous les sens numéros 12, 16, 17, 19, 20, 23 et 24. En même temps, ces explications – qui manquaient dans les ouvrages antérieurs – jouent un rôle très important dans le décodage correcte des mots-vedettes ou des exemples d'usage et dans l'interprétation des équivalents. Voir p.ex. le sens numéro 20 sous *aller*. On y rencontre l'explication *Játék* < jeu(x) > en hongrois, qui est indispensable pour comprendre le contexte des exemples et de leurs équivalents :

- Va! = marad! tartom! tartom a tételt.
- Va banque! = tartom a bankot.
- Tout va! = minden játszik. minden tételt elfogadunk.
- De combien allez-vous? = mennyit tart ön? meddig megy ön?

Un autre élément nouveau et unique de certaines entrées est l'indication des antonymes éventuels entre parenthèses, tout de suite après la catégorie grammaticale ou le numéro du sens dont il s'agit. Par exemple, dans le cas du verbe *aller*, Pokorny donne *venir* en tant qu'antonyme. Pokorny donne parfois même des synonymes et des homonymes (homophones) éventuels, mais ceux-ci se trouvent toujours à la fin d'un article, précédés d'un signe de double ligne verticale.

Même si Pokorny est novateur à plusieurs égards, les verbes hongrois en tant qu'équivalents des mots-vedettes verbaux français apparaissent à l'infinitif – comme chez Babos et chez Mártonffy.

Quant à la quantité des exemples, le Pokorny est beaucoup plus riche que ces prédécesseurs. Dans ce dictionnaire, on trouve bien plus d'exemples d'usage dans les entrées. P.ex. sous *aller* on en trouve 99, mais pour d'autres verbes, ce chiffre dépasse les cent – il y en a 140 pour *prendre* et 172 pour *faire*, une quantité comparable à celle dans certains dictionnaires de grand format de nos jours.

À l'égard des types d'exemples, le Pokorny présente les mêmes schémas que l'on rencontre déjà dans le Mártonffy : la proportion de T1 et de T2 varie en fonction des entrées. Le T1 est prédominant dans le cas de *faire*, *laisser*, *mettre*, *prendre* et *rendre*, alors que le T2 est plus fréquent dans les articles *aller*, *avoir*, *devoir*, *être*, *pouvoir* et *venir*.

Comme nous avons montré, la microstructure a beaucoup évolué, et les articles de ce dictionnaire fournissent beaucoup plus d'informations sur les mots-vedettes et leur usage que les prédécesseurs de Pokorny. P.ex. les verbes sémantiquement surchargés – tels que ceux dans notre étude – sont présentés et décrits d'une manière bien plus nuancée (voir l'augmentation du nombre des sens indiqués dans les entrées, ou la richesse en exemples).

Les deux points les plus faibles de ce dictionnaire que nous avons pu identifier sont la structuration des sens (surtout quand ils sont nombreux) dans le cas de certaines entrées et la typographie trompeuse des explications.

### 3.5 THEISZ 1902

Les entrées de Theisz se caractérisent par une microstructure extrêmement complexe qui paraît quelque fois impossible à déchiffrer.

Dans le Theisz, les mots-vedettes sont doublement mis en relief : ils apparaissent à la fois en majuscules et en gras. Les articles continuent par des abréviations françaises en italique indiquant la catégorie grammaticale. Parfois, on y rencontre d'autres abréviations (françaises ou hongroises – p.ex. (Rh) = *rendhagyó* < irrégulier >) fournissant des informations supplémentaires. Le mot-vedette peut même être précédé d'un chiffre qui indique que certaines entrées consécutives appartiennent l'une à l'autre. P.ex. dans le cas d'*aller*, on trouve trois entrées liées (« verbe neutre » ; « substantif masculin » ; verbe pronominal) :

- 1, ALLER (Rh) v/n.
- 2, ALLER s/m.
- 3, S'en ALLER

Theisz ne donne pas la prononciation du mot-vedette, ce qui représente un pas en arrière dans l'histoire des dictionnaires français–hongrois. D'autre part, la structuration des sens et des équivalents est devenue trop complexe, notamment à trois niveaux dans une même entrée – quatre si l'on y rajoute le numérotage devant les mots-vedettes –, les différents niveaux étant marqués par des chiffres romains, des chiffres arabes (en gras devant les mots-vedettes et en forme régulière après) et des lettres en minuscule. Parmi les équivalents, les verbes hongrois apparaissent à l'infinitif.

Ce qui peut être considéré comme un pas en avant, c'est la typographie et son usage conséquent. Les exemples, tout comme le mot-vedette, sont en gras, tandis que les équivalents et les traductions en hongrois se présentent en caractères réguliers, et ce n'est que les explications et les informations supplémentaires qui apparaissent en italique.

Dans beaucoup d'entrées (plus souvent que chez Pokorny), on rencontre aussi des synonymes et des antonymes, présentés entre crochets.

Quant au nombre des exemples, le Theisz est considérablement moins riche que le Pokorny, mais il contient en moyenne plus d'exemples que le Mártonffy – même s'il y a parmi nos entrées analysées quelques-unes (notamment trois sur onze) où les exemples sont moins nombreux dans le Theisz.

Au sujet des types d'exemples, le Theisz suit les mêmes schémas que le Pokorny et le Mártonffy : la proportion de T1 et de T2 varie en fonction des entrées.

Dans le cas des mots-vedettes qui sont difficiles à rendre par de simples équivalents hongrois (p.ex. *avoir*), le Theisz recourt à des moyens d'explication inhabituels dans les dictionnaires bilingues et qui caractérisent plutôt les livres d'enseignement, tel que l'indication de structures / constructions similaires dans d'autres langues qui peuvent être connues par les utilisateurs de l'ouvrage. Dans l'entrée *avoir*, entre crochets directement après le mot-vedette, on trouve les équivalents latin *habēre* et allemand *haben*. Même à l'intérieur de la même entrée, après l'exemple *j'ai une lettre à écrire*, on peut aussi trouver la traduction allemande *ich habe einen Brief zu schreiben*, outre la traduction hongroise. Vu les intérêts des utilisateurs, cela pouvait représenter une aide significative à l'époque.

### 3.6 SAUVAGEOT 1932

La microstructure des entrées dans le Sauvageot ressemble déjà à celles dans les dictionnaires de grand format modernes. Les mots-vedettes apparaissent en minuscules en gras, suivis éventuellement par un chiffre arabe en exposant pour séparer les mots homonymiques qui constituent donc des entrées à part (p.ex. *aller*<sup>1</sup> pour le verbe et *aller*<sup>2</sup> pour le nom).

La prononciation des mots-vedettes est indiquée entre crochets, suivie par des abréviations hongroises en italique entre parenthèses qui fournissent des informations sur la nature (surtout la catégorie grammaticale) des mots-vedettes. Si le mot-vedette est un verbe régulier, une lettre en gras entre parenthèses indique le type de conjugaison.

Les équivalents hongrois sont composés en caractères réguliers. Les verbes hongrois apparaissent (pour la première fois depuis le dictionnaire de Kiss) à l'indicatif présent de la 3e personne du singulier, et non à l'infinitif.

Les entrées plus complexes sont organisées à deux niveaux hiérarchiques : les chiffres romains indiquent de différents types d'usage/de fonction du mot-vedette (p.ex. dans le cas des verbes, le dictionnaire fait ainsi différence entre « verbe transitif », « verbe intransitif », « verbe impersonnel », « verbe pronominal », etc.), alors que les différents sens sous un tel type d'usage sont indiqués par des chiffres arabes. P.ex. *aller* est structuré de la manière suivante : **aller**<sup>1</sup> I. intransitif (16 sens), II. impersonnel (1 sens), III. *s'en aller* (6 sens); **aller**<sup>2</sup> nom masculin (1 sens).

Quant à la richesse en exemples, nous avons trouvé cinq entrées parmi les onze analysées dans lesquelles le nombre des exemples dépasse les cent – sous *faire*, on en trouve 282. En même temps, il faut aussi remarquer que certains parmi les nombreux exemples sont superflus, puisque plusieurs reflètent le même sens ou le même type d'usage. Parmi les 125 exemples sous *aller*, nous avons trouvé les suivants (nous avons mis une barre oblique entre les exemples de type identique) :

- I. 6. je vais lui écrire / je vais vous le raconter
- I. 10. de combien y allez-vous ? / il y est allé de dix francs
- I. 16. le malade va mieux / elle va de mieux en mieux
- III. 1. va-t'en ! / allez-vous-en ! / allons-nous-en !

Surtout dans le cas de I. 6. et de I. 16., on peut se demander si plus d'un exemple est vraiment nécessaire.

### 3.7 ECKHARDT 1953

La microstructure des entrées dans le Eckhardt ressemble beaucoup à celle dans le Sauvageot, mais les entrées sont plus détaillées, plus nuancées – avec plus de sens pour les mots-vedettes et avec plus d'exemples. Nous pouvons mettre en relief trois changements importants : 1) la prononciation des mots-vedettes n'est

pas indiquée ; 2) après les mots-vedettes, Eckhardt indique entre parenthèses une vingtaine de formes conjuguées des verbes irréguliers ; 3) il n'y a pas de numérotage en exposant après certains mots-vedettes pour différencier les mots homonymiques, car ils apparaissent à l'intérieur de la même entrée, différenciés par des chiffres romains – tout comme les différents types d'usage. P.ex. *aller* est structuré de la manière suivante : I.<sup>6</sup> (19 sens), II. auxiliaire (5 sens), III. impersonnel (1 sens), IV. transitif (1 sens), V. *s'en aller* (9 sens), VI. nom masculin (1 sens).

L'Eckhardt est extrêmement riche en exemples (voir Tableau 2), mais – tout comme dans le Sauvageot – il y en a des superflus. Nous en citons ici quelques-uns dans l'entrée *aller* :

- (I.) 16. de combien y allez-vous ? / il y est allé de mille francs
- II. 1. je vais vous le dire / je vais faire ce qu'on veut / son imprudence va le perdre
- V. 1. va-t'en ! / allez-vous-en ! / allons-nous-en !

Il arrive qu'un exemple de T2 est trop simple et/ou mal choisi, et ne fournit plus d'information qu'un exemple similaire de T1. Si le verbe conjugué de l'exemple (de T2a) est donné à l'indicatif présent de la 3e personne du singulier (soit également la forme de base pour les équivalents des verbes hongrois) – de plus, avec le pronom « il » en tant que sujet, qui ne se traduit pas en hongrois si le sujet n'est pas mis en relief – on aura la même traduction hongroise qu'avec l'infinitif en français. Voici deux exemples dans l'entrée *aller* :

- (I.) 2. aller à cheval
- (I.) 2. il va à pied

Dans ces deux cas, *aller* et *il va* ont pour traduction hongroise la même forme, *megy*.

Nous avons également identifié des exemples qui n'exemplifient pas du tout l'usage du mot-vedette, mais plutôt celui d'un autre constituant de l'exemple. Aussi dans l'entrée *aller*, on rencontre l'exemple suivant qui – à notre avis – devrait figurer dans l'article *avoir* :

(I.) 1. ils n'avaient qu'à ne pas y aller

---

<sup>6</sup>Par erreur, le chiffre romain « I. » manque à sa place dans l'entrée *aller*, mais il apparaît dans la deuxième édition de 1960.

### 3.8 ECKHARDT & OLÁH 1999

Du point de vue de la structuration des sens, la microstructure des entrées dans l'Eckhardt renouvelé est semblable à celle dans l'Eckhardt ancien. Les formes conjuguées des verbes placées derrière certains mots-vedettes ont été enlevées. À leur place, un chiffre arabe indique le type de conjugaison. Les exemples ne sont plus en italique mais en gras, ce qui rend la lecture des entrées beaucoup plus facile.

Une partie des exemples est identique dans les deux ouvrages, mais il y a eu tout de même de nombreux changements à cet égard : certains exemples (surtout parmi les superflus, vieillis, trop spécifiques ou problématiques) ont été retirés, d'autres ont été rajoutés. Quant à la richesse des entrées étudiées, le nombre des exemples a considérablement augmenté dans certaines entrées (*aller, faire, laisser, mettre*), alors qu'il a baissé dans d'autres (*devoir, être, prendre, rendre*).

Beaucoup d'exemples superflus ont été cependant retenus, p.ex. ceux que nous avons mentionnés dans Section 3.7. De plus, un bon nombre de nouveaux exemples superflus ont été rajoutés, voir p.ex. dans l'entrée *aller* :

I. 8. b) ces vêtements vont dans le placard / cette housse va sur le matelas

### 3.9 PÁLFY 1999

Visuellement, la microstructure du Pálffy ressemble à celle des entrées de l'Eckhardt & Oláh, car les exemples y sont composés en caractères gras, et la structuration des sens suit une logique semblable. Ce qui est différent, entre autres, c'est le traitement des mots homonymiques : les homonymes apparaissent dans des entrées à part, et l'on trouvera un chiffre arabe devant le mot-vedette (comme dans le Theisz).

Les entrées du Pálffy comprennent beaucoup plus d'informations supplémentaires (en langue française, entre parenthèses) qui suivent les différents exemples et qui aident les utilisateurs à mieux comprendre de quelles manières et dans quels types de contexte les structures présentées par les exemples sont utilisées, fournissant ainsi des informations aussi bien sur les formalités grammaticales que sur la sémantique. Les équivalents hongrois sont aussi formulés d'une telle façon que cet ouvrage peut être considéré comme un dictionnaire de décodage et d'encodage à la fois.

Les exemples d'usage dans les entrées sont plutôt bien choisis et bien formulés, mais nous en avons aussi trouvés quelques-uns qui – pour des raisons diverses – sont problématiques. Si l'on regarde l'entrée *aller*, on tombe sur l'exemple *tout*

*son argent s'en va en cassettes*. Dans ce cas, le choix du mot *cassette* dans un tel exemple est discutable – les cassettes étant en voie de disparition en 1999. Toujours dans la même entrée, on trouve des exemples parfaitement choisis pour illustrer le futur proche, tels que *elle va accoucher dans six mois* ou *je vais y aller*, en revanche, le rajout des exemples « ça va faire des histoires !/ça va barder ! *fam* » [sic – le manque de l'espace devant les signes doubles est un problème général dans le Pálffy] n'est pas justifié. D'une part, ces deux expressions (familières) n'apportent rien de plus par rapport aux exemples précédents quant à l'illustration de la fonction grammaticale en question, d'autre part la nature familière des expressions n'est pas due à l'utilisation du verbe *aller* (soit le mot-vedette dans ce cas), mais aux expressions *faire des histoires* et *barder*.

Finalement, nous voudrions mentionner le problème des nouvelles éditions dans lesquelles les différents types de fautes peuvent être conservés, même s'il s'agit de plusieurs éditions revues/corrigées (la 7<sup>e</sup> édition revue et corrigée a paru en 2009). P.ex. dans l'entrée *aller*, l'exemple « on sonne, va ouvrir » et même sa traduction hongroise figure deux fois. Ce qui est encore plus surprenant, c'est que de telles fautes ne sont corrigées dans la version en ligne du dictionnaire non plus – comme s'il s'agissait d'une simple réimpression.

### 3.10 BÁRDOSI & SZABÓ 2007

Du point de vue de la clarté de la microstructure des entrées, ainsi qu'au niveau du visuel et de la maniabilité, ce dictionnaire est celui qui nous paraît le mieux organisé.

Les mots-vedettes apparaissent en minuscules en gras (comme dans les autres dictionnaires du XX<sup>e</sup> siècle depuis Sauvageot), suivis éventuellement par un chiffre arabe en exposant pour séparer les mots homonymiques qui sont traités dans des entrées à part. Ce qui est une nouveauté remarquable et très utile, c'est la réapparition de la prononciation des mots-vedettes, cette fois-ci (pour la première fois!) avec les symboles de l'alphabet phonétique international (API). La transcription phonétique des mots-vedettes est suivie par de différents types d'informations sur le mot-vedette (p.ex. type de conjugaison pour les verbes, classe grammaticale, etc.).

Les équivalents hongrois sont composés en caractères réguliers, souvent complétés d'informations supplémentaires entre crochets.

Quant à la richesse en exemples (qui apparaissent en gras), leur nombre n'atteint pas la même quantité que l'on trouve dans le Pálffy. En même temps, les exemples sont assez nombreux pour pouvoir bien illustrer les différents sens présentés des mots-vedettes, et ils sont très bien choisis. Ce qui ne paraît pas toujours

logique, c'est la structuration de la présentation de certains exemples – quelque chose qui peut être remarqué par des linguistes, mais qui ne gêne pas l'utilisateur moyen du dictionnaire, surtout qu'il s'agit de cas rares et sporadiques. Voyons un exemple parmi les exemples d'usage de l'entrée *aller* :

- **en bus/tram/méto** busszal/villamossal/metróval megy
- **en avion** repülőgépen utazik
- **en voiture** kocsival megy

Il n'est pas évident pourquoi *en voiture* n'apparaît pas ensemble avec *bus/tram/méto*, surtout que la construction hongroise reflétée par les traductions est la même pour les deux. Bien sûr, on peut reconnaître que les éléments du trio « bus/tram/méto » représentent le même type de transport (transport en commun), alors que *voiture* n'en fait pas partie, mais on peut aussi se demander si une telle différenciation sémantique est vraiment nécessaire. De plus, *en avion* pourrait également être mentionné avec les transports en commun (soit des transports collectifs), si son équivalent hongrois suivait la même structure (avec le verbe *megy* et le mot signifiant le mode de transport à l'instrumental-comitatif) – ce qui serait tout à fait possible en hongrois : *repülőgéppel megy* au lieu de *repülőgépen utazik* (où le mot pour « avion » est au superessif avec le verbe *utazik*). Il se peut que les auteurs aient eu l'intention de montrer que le verbe *utazik* (< voyager >) s'utilise souvent en hongrois en cas de voyages en avion, mais ce verbe s'utilise aussi avec les autres modes de transport. Il serait peut-être préférable de les présenter ensemble, avec deux équivalents possibles (tous les deux étant souvent utilisés). L'équivalent comprenant le verbe *megy* pourrait même être complété avec le verbe *utazik* en tant que variante, puisque ce dernier s'emploie aussi bien avec l'instrumental-comitatif qu'avec le superessif en hongrois : « aller en avion/bus/tram/méto » *repülőgéppel/busszal/villamossal/metróval megy/utazik, repülőgépen/buszon/villamoson/metrón utazik*.

### 3.11 « L'Ouest rencontre l'Est » : Ou le cas des entrées *avoir, devoir* et *pouvoir*

Puisque le hongrois est une langue finno-ougrienne (et non indo-européenne comme le français), ce n'est pas surprenant que certaines constructions grammaticales s'y forment de manières différentes. On pourrait nommer de nombreux cas dans lesquels un sens ou une fonction s'exprime par un seul mot dans une

langue, alors qu'il est impossible de définir le même sens par un simple équivalent dans une autre langue. Nous avons choisi de mettre en relief ici trois mots-vedettes français qui peuvent poser des problèmes à cet égard, et dont la présentation se montre insuffisante dans les premiers dictionnaires français–hongrois. Ce type de problème ne touche plus le dictionnaire de Sauvageot et les ouvrages parus depuis.

### 3.11.1 *avoir*

Dans la langue hongroise, il n'y a pas de verbe *avoir* pour exprimer la possession. C'est l'équivalent du verbe *être* qui s'utilise en tant que prédicat, et le sujet sera l'objet directe (le possédé) de la construction française, marqué en hongrois par un suffixe possessif, tandis que le sujet du verbe français sera utilisé en hongrois au datif. Ce type de construction est d'ailleurs typique pour les langues finno-ougriennes. Voyons quelques phrases concrètes :

- (1) hongrois  
Van (nekem) (egy) szótár-am.  
est (DAT.1SG) (un) dictionnaire-1SG.POSS  
« J'ai un dictionnaire. » (litt. « Il y a (à moi) mon dictionnaire. »)
- (2) hongrois  
Péter-nek van (egy) szótár-a.  
Pierre-DAT est (un) dictionnaire-3SG.POSS  
« Pierre a un dictionnaire. » (litt. « À Pierre il y a son dictionnaire. »)

Une telle construction s'explique mieux par des exemples d'usage dans un dictionnaire, mais les premiers dictionnaires français–hongrois n'en donnent aucun pour ce sens de « posséder ». En tant qu'équivalent hongrois, Kiss donne les mots *lenni* (l'infinitif « être ») et *van* (forme conjuguée – voir Section 3.1), et il les complète avec l'infinitif allemand *haben*. Outre *lenni*, Babos donne aussi *birni*, un verbe polysémique qui peut signifier « posséder », mais sans exemples d'usage, il est difficile d'interpréter cet équivalent. Mártonffy donne aussi *bírni*<sup>7</sup> et *lenni*, et il complète ce dernier entre parenthèses avec les suffixes *-nak*, *-nek* indiquant le datif, mais – comme il n'en écrit plus – il n'est pas évident quelle partie de l'expression devrait être mise au datif (et il existe d'autres constructions en hongrois avec *lenni* et le datif). Pokorny donne seulement *birni* (avec *i*), et il le complète avec un seul exemple, un « proverbe » formellement assez complexe, ce qui ne

<sup>7</sup>Chez Mártonffy, ce mot s'écrit déjà avec *í* (et non avec *i* comme chez Babos) – soit la forme utilisée aujourd'hui.

facilite pas vraiment la tâche de l'utilisateur du dictionnaire : *Il n'est rien de tel que d'en avoir*. C'est d'abord chez Theisz que l'on trouve plus d'informations sur ce sens en question. Theisz ne donne que les infinitifs latin (*habēre*) et allemand (*haben*) comme équivalents, et il met un mot hongrois (*birtok* < possession >) entre parenthèses en tant qu'information supplémentaire. Il continue ensuite avec des exemples qui sont déjà plus utiles : *j'ai une plume*; *il a des amis*; *je l'ai pour ami* et *il a de quoi (vivre)*.

### 3.11.2 *devoir*

Ce verbe français peut avoir plusieurs sens, parmi lesquels nous voudrions mentionner son usage avec un autre verbe à l'infinitif pour exprimer une sorte d'obligation : « quelqu'un doit faire quelque chose ». Dans la langue hongroise, il n'y a pas de verbe correspondant qui pourrait être conjugué en fonction de l'agent – soit la personne qui doit agir, identique avec le sujet de la construction française. Avec ce sens, c'est le verbe *kell* qui s'utilise en hongrois, dont le sujet est l'infinitif (désignant l'activité à accomplir), et l'agent (comme une sorte de complément) se met au datif. Très souvent, l'infinitif utilisé dans une telle construction peut être marqué par un suffixe possessif indiquant la personne de l'agent – soit une forme personnelle conjuguée de l'infinitif.

- (3) hongrois  
Péter-nek vár-ni-a kell.  
Pierre-DAT attendre-INF-3SG.POSS falloir-3SG.PRES  
< Pierre doit attendre. >
- (4) hongrois  
(Nekem) vár-n-om kell.  
(DAT.1SG) attendre-INF-1SG.POSS falloir-3SG.PRES  
< Je dois attendre. >

Dans le Kiss, on ne trouve que trois équivalents simples qui seraient censés de couvrir tous les sens (principaux) du verbe *devoir* : *adós*, *kell* et *tartozik*. Le premier est un adjectif (utilisé également comme substantif) signifiant < ayant une dette; débiteur >, le deuxième (mentionné plus haut) est un verbe qui s'emploie d'une manière complètement différente au niveau de la syntaxe par rapport au verbe *devoir*, tandis que le troisième – également un verbe – signifie < avoir une dette >.

Babos donne les équivalents suivants, sans exemples : *adósnak lenni* (< être débiteur >), *tartozni* (< avoir une dette >), *köszönhetni* (< être redevable de qc. à qu. >), *kelleni* (l'infinitif de *kell*), et *kötelesnek lenni* (< être obligé >).

Mártonffy différencie cinq sens du verbe *devoir*, parmi lesquels le sens que nous mettons ici en relief est représenté par l'équivalent *kelleni*. Heureusement, l'auteur donne deux exemples, parmi lesquels c'est surtout le premier qui peut être utile avec sa traduction hongroise : *il doit parti* [sic], en hongrois *el kell utaznia*. Malheureusement, la forme erronée de l'infinitif *partir* peut tromper l'utilisateur hongrois du dictionnaire, car la forme « parti » existe en tant que participe passé. Le second exemple est moins utile, puisqu'il comprend une négation, ce qui peut changer la sémantique du verbe en fonction du contexte : *vous ne devez pas rester* dont la traduction dans le Mártonffy est *nem szabad maradnia*. Cette traduction suggère justement un autre sens du verbe *devoir*, elle signifie « Vous ne pouvez pas rester ».

Chez Pokorny, on trouve deux équivalents à l'égard du sens en question : *kell* (curieusement, l'auteur donne la forme conjuguée du verbe, alors qu'ailleurs il donne l'infinitif des verbes hongrois) et *köteles* (adjectif qui signifie « obligé »). Les exemples dans le Pokorny sont plutôt compliqués et moins utiles : *Les choses devaient se venir là*; *Dussais-je y périr*; *Fais ce que dois, advienne que pourra* (proverbe); *Vas où tu peux, mourir ou* [sic] *tu dois* (proverbe).

Dans le Theisz, une petite remarque d'usage éclaircit que le sens en question s'exprime si l'infinitif d'un autre verbe suit *devoir*. Au lieu de simples équivalents, Theisz ne donne que deux exemples pour ce sens : *je dois servir...* et *il doit savoir que...* Le second exemple peut être considéré comme problématique, puisqu'il est sémantiquement ambigu.

### 3.11.3 *pouvoir*

En français, on a l'habitude de faire la différence entre *pouvoir* (*faire qc*) et *savoir* (*faire qc*) – si quelqu'un est capable de faire quelque chose ou bien parce que les circonstances le rendent possible (*pouvoir*), ou bien en raison de ses connaissances/compétences (*savoir*). En hongrois, on ne fait pas cette distinction, ces deux verbes français peuvent être traduits par le même verbe hongrois : *tud*. On aura donc besoin de certains compléments en hongrois pour pouvoir faire la différence. Les phrases « je ne peux pas nager » et « je ne sais pas nager » peuvent donc se traduire en hongrois par une même phrase : *nem tudok úszni* (NEG pouvoir/savoir-1SG.PRS nager-INF). En revanche, en fonction de la nature des circonstances, *pouvoir* peut être exprimé de manières différentes. Le verbe hongrois *tud* représente donc l'une des possibilités. L'autre possibilité consiste à se servir d'un suffixe potentiel – une forme qui n'est pas considérée par les grammaires hongroises comme un mode verbal à part. (Le verbe *pouvoir* peut être traduit en hongrois de plusieurs manières, ici nous ne mentionnons que les

constructions les plus fréquentes pour le sens « avoir la possibilité de faire qc ».) Ce potentiel est utilisé en hongrois surtout pour exprimer que l'acte (indiqué par l'infinitif qui suit le verbe *pouvoir* en français) est permis, mais il peut aussi se référer à d'autres circonstances – voir les deux phrases suivantes.

- (5) hongrois  
 Nem utaz-hat-nak Párizs-ba.  
 NEG voyager-POT-3PL Paris-ILL  
 < Ils/Elles ne peuvent pas voyager à Paris. >
- (6) hongrois  
 Nem tud-nak Párizs-ba utaz-ni.  
 NEG pouvoir-3PL.PRS Paris-ILL voyager-INF  
 < Ils/Elles ne peuvent pas voyager à Paris. >

Dans le Kiss, on ne trouve qu'un seul équivalent hongrois : *lehet*. À l'origine, ce mot hongrois est le verbe *van* « être » (*lenni* à l'infinitif) au potentiel (au présent de la 3e personne du singulier). Cette forme signifie donc « pouvoir/peut être », et, en effet, c'est aussi un des équivalents possibles de l'expression française « peut-être ». Il faut aussi remarquer que *lehet* a un troisième type d'usage : il peut être utilisé avec d'autres verbes à l'infinitif. P.ex. *lehet menni* (être-POT.3SG.PRS aller-INF) « on peut / il est possible d'y aller ». Il est donc difficile d'interpréter cet équivalent hongrois sans aucun exemple d'usage.

Babos donne les équivalents suivants : ...*hatni*, ...*hetni*, *tehetni*, *megtehetni*, *birni* et *birhatni*. Les deux premiers représentent les deux variantes du suffixe potentiel avec la terminaison de l'infinitif, les trois points indiquant la place d'un verbe quelconque. Le troisième et le quatrième sont, en revanche, des verbes concrets au potentiel ayant le sens « pouvoir faire ». Le cinquième est le même verbe polysémique que nous avons déjà mentionné à propos du verbe *avoir* (« posséder » et « pouvoir » étant seulement deux des traductions possibles) – difficile à interpréter sans exemples d'usage. Finalement, le sixième est le même que le précédent, mais au potentiel – une forme tautologique.

Les équivalents de Mártonffy sont moins nombreux et moins compréhensibles. Il donne *hatni...*, *hetni...* – on peut se demander pourquoi les trois points suivent les suffixes. Ensuite, il continue avec un exemple de T1 : *pouvoir marcher* avec l'équivalent *mehetni* – exemple simple, sans faute et assez efficace. Son exemple suivant se base en revanche sur la forme pronominale (impersonnelle) *se pouvoir*, ce qui est très curieux, étant donné qu'il traite ce sens séparément dans l'entrée avec d'autres exemples semblables. Plus loin dans l'entrée, il donne également les équivalents *tehetni* « pouvoir faire » et *birni* (voir plus haut).

Pokorny donne les équivalents suivants : *hatni*, *hetni*, *birni* et *tehetni*. Les deux derniers ont déjà été présentés plus hauts. Quant aux deux premiers, ils apparaissent comme s'ils étaient des infinitifs indépendants, mais, en réalité, ce ne sont que des suffixes – comme nous l'avons expliqué à propos des équivalents de Babos. Les exemples – comme c'est souvent le cas chez Pokorny – sont trop compliqués et peu efficaces : *On ne peut être plus mal servi* ; *Je ne puis qu'y faire*.

Chez Theisz aussi, on trouve *-hatni* et *-hetni*, mais grâce à l'usage des traits d'union (tout comme avec les trois points chez Babos), il est clair comment on doit interpréter ces équivalents. Theisz continue ensuite avec *tudni* (pour la première fois!), *bírni* (complété par *vmit tenni* < faire qc >) et *képesnek lenni* < être capable >. L'auteur précise que ce sens se base sur l'usage de *pouvoir* avec un autre verbe à l'infinitif, ensuite, il continue avec ses exemples (peu nombreux et assez simples) : *il peut écrire* ; *si je peux parler ainsi*.

## 4 Conclusion

Les mots-vedettes et les équivalents suivent et reflètent en général l'époque des différents ouvrages, mais ce n'est pas toujours le cas des exemples d'usage. Le choix et la présentation des exemples représentent un problème constant. Dans les dictionnaires du XIXe siècle, les exemples ont été choisis sans logique transparente, plutôt au hasard, pourrait-on dire – un peu comme la structure des entrées au sujet des sens présentés.

Le nombre des entrées ne reflète pas automatiquement la richesse des entrées en exemples. Dans le cas du PÁLFY, nous avons montré qu'un dictionnaire à 35 mille entrées peut fournir presque autant d'exemples que les dictionnaires de grand format ayant au moins deux ou trois fois plus d'entrées.

En raison de ses erreurs, l'Eckhardt renouvelé de 1999 (l'ouvrage ECKHARDT & OLÁH) devrait être renouvelé encore une fois. Il est cependant fort douteux qu'un tel projet puisse avoir lieu dans un avenir proche. Dans un premier temps, il suffirait de corriger les erreurs de l'ouvrage et de publier une version numérisée (en ligne), qui serait plus facile à compléter selon les besoins. Cela est aussi valable pour les autres dictionnaires actuellement en vente.

Le grand dictionnaire de SAUVAGEOT mériterait d'être numérisé en raison de ses qualités pour faciliter le travail des philologues.

Les dictionnaires de PÁLFY et de BÁRDOSI & SZABÓ représentent le plus haut niveau de la lexicographie. Les articles sont très clairement structurés, d'une façon très efficace, ce qui fait qu'il s'agit de deux dictionnaires facilement maniables. De plus, les équivalents dans ses dictionnaires sont les plus fiables – même si l'on

peut y trouver quelques erreurs. Il serait très important d'« entretenir » constamment ces dictionnaires précieux, de publier des versions en ligne dont le contenu serait de temps en temps revu, corrigé et augmenté.

## Références

- BABOS, Kálmán. 1865. *Uj francia–magyar és magyar–francia szótár. A legjobb kútfők különösen Molé francia szótára szerint szerkesztve. Harmadik tetemesen bővített és javított kiadás. Első kötet. Francia–magyar rész. = Nouveau dictionnaire français–hongrois et hongrois–français d'après les meilleurs ouvrages publiés jusqu'à ce jour et spécialement d'après le dictionnaire français de Molé. Troisième édition revue et considérablement augmentée. Tome premier. Partie française–hongroise.* Pest : Heckenast.
- BÁRDOSI, Vilmos & Dávid SZABÓ. 2007. *Francia–magyar kéziszótár = Dictionnaire français–hongrois.* Budapest : Akadémiai Kiadó.
- BELLEFLAMME, Clémence. 2019. Quels exemples (lexicographiques) pour le traducteur? *Meta* 64(1). 3-19. DOI : 10.7202/1065326ar.
- BRAY, Laurent. 2000. Bibliographie des dictionnaires français–allemand et allemand–français de 1816 à 1899. *Cahiers de lexicologie* 76(1). 167-189.
- ECKHARDT, Sándor. 1953. *Francia–magyar szótár = Dictionnaire français–hongrois.* Budapest : Akadémiai Kiadó.
- ECKHARDT, Sándor & Tibor OLÁH. 1999. *Francia–magyar nagyszótár = Grand dictionnaire français–hongrois.* Budapest : Akadémiai Kiadó.
- ELLIGERS, Anne & Finn REICHBORN-KJENNERUD. 1995. *Fransk ordbok : Fransk-norsk, norsk-fransk* (Blå ordbøker). Oslo : Kunnskapsforlaget.
- JACOBSEN, Tove. 2014. *Stor fransk ordbok : Fransk-norsk, norsk-fransk.* Oslo : Fagbokforlaget.
- JEZERNICZKY, Margit. 1933. *Les impressions en français de Hongrie (1707–1848)* (Études françaises 8). Szeged : Institut français de l'Université de Szeged. [http://acta.bibl.u-szeged.hu/38815/1/francia\\_tanulmanyok\\_008.pdf](http://acta.bibl.u-szeged.hu/38815/1/francia_tanulmanyok_008.pdf).
- KARPINSKA, Laura. 2015. *English-Latvian Lexicographic Tradition : A Critical Analysis.* Berlin : De Gruyter. DOI : 10.1515/9783110365764.
- KASIMI, Ali M. al-. 1983. *Linguistics and bilingual dictionaries.* Photomechanical reprint. Leiden : Brill.
- KISS, Mihály. 1844. *Francia–magyar és magyar–francia zsebszótár. Első kötet. Francia–magyar rész. = Nouveau dictionnaire de poche. Tome premier. Partie française.* Pest : Heckenast.
- MAGAY, Tamás. 2011. *A magyar szótárirodalom bibliográfiája* (Lexikográfiai füzetek 1). Budapest : Akadémiai Kiadó.

- MÁRTONFFY, Frigyes. 1879. *Teljes francia és magyar szótár. Első kötet : Francia–magyar rész.* = *Dictionnaire complet des langues françaises et hongroises. Tome premier : Partie française–hongroise.* Budapest : Franklin Társulat.
- NIELSEN, Sandro. 2014. Example sentences in bilingual specialised dictionaries assisting communication in a foreign language. *Lexikos* 24(1). 198-213. DOI : 10.5788/24-1-1259.
- PÁLFY, Miklós. 1999. *Francia–magyar kéziszótár = Dictionnaire français–hongrois.* Szeged : Grimm Kiadó.
- PERROT, Jean. 1992. Soixante ans après : Encore « enfler des mots ». *Cahiers d'études hongroises* 4. 93-97.
- PIOTROWSKI, Tadeusz. 2000. Examples : The Ragbag of Bilingual Lexicography? *Lexicographica : International Annual for Lexicography/Internationales Jahrbuch für Lexikographie* 16. 14-24. DOI : 10.1515/9783110244205.14.
- POKORNY, Jenő Ármin. 1880. *A francia–magyar és magyar–francia nyelv szótára. Francia–magyar szótár.* = *Dictionnaire français–hongrois et hongrois–français.* 3 t. Budapest : Rautmann.
- SÁGI, István. 1922. *A magyar szótárak és nyelvtanok könyvészete.* Budapest : Magyar Nyelvtudományi Társaság.
- SAUVAGEOT, Aurélien. 1932. *Francia–magyar és magyar–francia nagy kéziszótár. Francia–magyar rész.* = *Dictionnaire général français–hongrois et hongrois–français. Première partie français–hongrois.* Budapest : Dante.
- SVENSÉN, Bo. 2009. *A Handbook of Lexicography : The Theory and Practice of Dictionary-Making.* Cambridge : Cambridge University Press.
- SZABÓ, Dávid. 2011. Les dictionnaires hongrois–français et français–hongrois. In Zsuzsanna FÁBIÁN (éd.), *Hungarian Lexicography I : Bilingual Dictionaries (Lexikográfiai füzetek 5)*, 79-91. Budapest : Akadémiai Kiadó.
- THEISZ, Gyula. 1902. *Francia–magyar és magyar–francia szótár. I. Francia–magyar rész.* = *Dictionnaire français–hongrois.* Budapest : Lampel.

# Chapitre 19

## « Sul finir d'imparare la Grammatica Francese, fa d'uopo studiar il Dizionario delle Frasi » : Deux recueils phraséologiques bilingues franco-italiens de la première moitié du 19e siècle

Michela Murano

Università Cattolica del Sacro Cuore, Milan, Italy

In this paper we provide a metalexical analysis of two French-Italian phraséological dictionaries published in Naples, in Italy, in the first half of the 19th century : the *Fraseologia italiana-francese* by Giuseppe TRUCCHI (1825) and the *Dictionnaire des phrases françaises-italiennes* by Nicolas LANZA (1837). These two works are composite but appear to be devoted mainly to the description of phraséological units of various types : in fact, both authors are convinced that knowing phraseology is complementary to the study of grammar and is indispensable for perfecting the knowledge of the French language. Our aim is to detect the specificities of these works at the level of the target audience, the chosen nomenclature and the lexicographic techniques implemented, through the exploration of the paratext, the macrostructure and the microstructure.

### 1 Introduction

Au 19e siècle, la présence d'un nombre considérable de recueils qui mettent en regard les phraséologies française et italienne témoigne de l'intérêt que les didacticiens et les lexicographes portent à la partie figée du lexique. On repère aussi



bien des parties d'ouvrages de grammaire ou de conversation consacrées à la phraséologie que des recueils indépendants, dont certains relèvent de véritables projets lexicographiques.

En Italie, l'enseignement de la langue française devient à cette époque une discipline scolaire qui touche un vaste public, comprenant des élèves issus de couches sociales même modestes (PELLANDRA 1997, MINERVA 1998). Les recueils phraséologiques, qui font partie du matériel didactique complémentaire élaboré par des professeurs, tout comme les manuels de lecture ou les recueils de dialogues, n'ont pas encore fait l'objet d'études approfondies pour le couple de langues français-italien<sup>1</sup> : nous avons récemment constitué un premier corpus de recueils (MURANO 2017) et entamé l'analyse par deux des rares ouvrages publiés en France (MURANO 2020) ; des études pointues sur un manuscrit de Goudar (LILLO 2014) et sur les dictionnaires phraséologiques commerciaux (SCLAFANI 2019) ont également vu le jour ; la corrélation entre la terminologie phraséologique adoptée et la typologie des phrasèmes recensés a été étudiée par AUTELLI (2022) dans les *phraséologies* de la fin du 19<sup>e</sup> siècle.

Dans cet article, nous nous proposons de présenter deux recueils phraséologiques parus en Italie dans la première moitié du siècle, dans les décennies 1820 et 1830 et publiés chez un même éditeur à Naples. Tout en étant des textes composites, ils paraissent consacrés en priorité à la phraséologie, si l'on en juge les titres : Giuseppe Trucchi a publié une *Fraseologia italiana-francese* en 1825, Lanza un *Dictionnaire des phrases françaises-italiennes* en 1837.

Après quelques considérations sur la phraséodidactique du Français Langue Étrangère en Italie au 19<sup>e</sup> siècle par le biais des études de manuels et un bref aperçu sur les recueils phraséologiques bilingues publiés à la même époque, les recueils de Trucchi e Lanza seront situés dans la production didactique et lexicographique de chaque auteur ; ensuite, nous analyserons la composition des ouvrages et le paratexte (titres, frontispices, préfaces), afin d'établir les motivations qui ont poussé les auteurs à les écrire, les objectifs pédagogiques, le public visé et les sources ; puis nous estimerons la nomenclature et nous essaierons de déterminer les critères de rangement des entrées. Enfin, l'analyse de la microstructure permettra de vérifier la présence éventuelle de rubriques lexicographiques.

---

<sup>1</sup>Un ouvrage magistral a par contre vu le jour pour le couple de langues français-hongrois (BÁRDOSI 2017).

## 2 La phraséologie dans le matériel pédagogique (1820–1830)

Dans les années 1990, les historiens de l'enseignement du Français Langue Étrangère (FLE) membres du CIRSIL (*Centro Interuniversitario per la Ricerca sugli Insegnamenti Linguistici*<sup>2</sup>) se sont intéressés à la phraséologie dans le cadre d'études sur les manuels de FLE en Italie. L'étude la plus significative est celle de Lillo en 1994, dans laquelle, en prenant en compte la structure générale des manuels de l'époque, la phraséologie est définie d'un point de vue pragmatique : « ce qui n'est ni phonétique, ni morphologie, ni syntaxe, ni lexique, ni dialogues, ni correspondances ou morceaux choisis, ni exercice, ni versification » (LILLO 1994 : 70–71). Sur la base de cette définition et se référant aux manuels publiés entre 1625 et 1860, Lillo repère trois façons différentes de présentation de la phraséologie : les deux premières se réfèrent à la présentation à l'intérieur des manuels, dans une section à part bien distincte ou disséminée dans les autres sections (« phraséologie périphérique », (LILLO 1994 : 71); la troisième correspond à l'organisation en recueils ou en dictionnaires.

Concernant ce troisième mode de présentation du matériel phraséologique, dans une étude précédente (MURANO 2017) nous avons essayé d'évaluer le nombre de recueils phraséologiques bilingues franco-italiens en croisant les données des répertoires existants pour les dictionnaires bilingues (LILLO 2019) et les manuels de FLE (PELLANDRA 1997, MINERVA 1996), auxquelles nous avons ajouté les résultats de recherches dans les catalogues nationaux *Opac Servizio Bibliotecario Nazionale* et *Catalogue Collectif de France* (s. d.). Une augmentation exponentielle du nombre de recueils peut être observée au 19e siècle, au cours duquel 66 recueils<sup>3</sup> sont publiés, contre 8 au total dans les deux siècles précédents. La publication de la plupart des recueils se situe dans la deuxième moitié du siècle, mais dans la période 1820–1839 on enregistre déjà la parution d'un nombre appréciable de recueils (15).

Si l'on prend en compte la périodisation proposée par Lillo pour les manuels de FLE, les ouvrages que nous allons examiner dans cette étude, publiés en 1825 et 1837, se situent au tournant entre une période (1780–1830) dans laquelle les manuels privilégient la partie syntaxique et accordent peu de pages à la phraséologie, située surtout en position périphérique, et une période (1830–1860) dans

---

<sup>2</sup><https://cirsil.it>

<sup>3</sup>Nous adoptons ici la dénomination *recueil* en tant qu'hyperonyme, car à l'étape actuelle de nos recherches, n'ayant pu consulter qu'une partie des ouvrages du corpus, il ne nous est pas possible de savoir dans quels ouvrages les entrées sont rangées selon l'ordre alphabétique et combien d'ouvrages relèvent d'un projet lexicographique.

laquelle l'intérêt pour la phraséologie augmente et plus de pages y sont consacrées. L'apparition d'un grand nombre de recueils phraséologiques bilingues s'expliquerait donc jusqu'à 1830 par la nécessité de compenser « la pénurie de pages » consacrées à la phraséologie dans les manuels, et après 1830 par un regain d'intérêt pour ce sujet.

Dans une étude récente (MURANO 2020), nous nous sommes penchée sur deux des rares recueils phraséologiques bilingues publiés dans les années 1820 en France, le *Petit trésor de la langue française et de la langue italienne* de Giuseppe Filippo BARBERI (1821) et le *Dictionnaire des idiotismes Italiens-Français* de Giacomo POLESÌ (1829), qui, tout en poursuivant un même objectif, celui de faciliter la compréhension des textes littéraires italiens, correspondent à deux typologies de recueils très différentes. Le *Petit trésor* de Barberi se présente comme un *traité de tropes* composé de plusieurs parties, parmi lesquelles les « Figures dans l'association de plusieurs mots » et une « Phraséologie familière », qui ne suivent aucun critère de présentation des entrées. Les « Figures dans l'association de plusieurs mots » présentent cependant l'intérêt d'être structurées en blocs de texte qui suivent un ordre de présentation régulier des informations (citation tirée d'un texte littéraire, forme lemmatisée, explication du sens). Le *Dictionnaire des idiotismes de Polesi*, comme son titre l'indique, est un dictionnaire phraséologique au sens moderne du terme, qui suit un ordre alphabétique strict de présentation des entrées et différencie les informations fournies (traduction littérale, équivalent, explication, renvois synonymiques) par l'emploi de différents caractères typographiques.

Dans les paragraphes suivants, nous présenterons deux recueils publiés, comme ceux de BARBERI (1821) et POLESÌ (1829), après 1820 et avant 1840, au moment de la première hausse considérable dans le nombre de publications de recueils phraséologiques franco-italiens. Ce premier critère de sélection chronologique se double d'un critère d'unité spatiale, car il s'agit de deux ouvrages publiés dans une ville de l'Italie du Sud, et donc vraisemblablement destinés à un public italophone. Les objectifs didactiques seront donc foncièrement différents par rapport à ceux des recueils de Barberi et Polesi, destinés à un public francophone apprenant l'italien « sans maître » (DUBOIS 2015 : 16).

### 3 La *Fraseologia italiana-francese* (1825) et le *Dictionnaire des phrases françaises-italiennes* (1837)

#### 3.1 Les auteurs et leur production

Le recueil phraséologique de Giuseppe Trucchi, qui se définit *maître de la langue française* (« maestro di lingua francese ») a eu plus de deux éditions, ce qui est assez rare pour ce type d'ouvrages (MURANO 2017 : 171) : la *Fraseologia italiana-francese* publiée en 1825 chez Seguin, que nous analysons dans cet article, est en effet la deuxième édition augmentée (on lit dans le frontispice « Corredata di utili giunte ») d'un premier ouvrage publié en 1812 avec le même titre ; une troisième édition voit le jour en 1854. La *Prefazione* de 1825 fournit des renseignements précieux sur le succès de la première édition de 1812, imprimée et vendue en 1500 exemplaires, qui a poussé l'auteur à travailler à la deuxième édition plutôt qu'à réimprimer la première :

« La pubblica approvazione ch'ebbe allora la mia *Fraseologia*, e le continue istanze che n'ebbi, dopo averne esaurite più di mille cinquecento copie, per farla ristampare, mi hanno fatto decidere di procedere alla seconda edizione [...]»

Trucchi est également l'auteur d'autres ouvrages didactiques publiés dans les décennies 1820 et 1830 chez différents éditeurs à Naples : il s'agit d'un recueil de dialogues (TRUCCHI 1834) et deux suppléments aux grammaires ((TRUCCHI 1831) et (TRUCCHI 1837), qui annoncent dans le frontispice la présence d'une partie consacrée à la phraséologie (un « elenco di frasi particolari, proverbj e gallicismi », « una lunga lista di gallicismi »)<sup>4</sup>.

De son côté, Nicolas Lanza paraît n'avoir publié que le *Dictionnaire des phrases françaises-italiennes* en 1837.

Les deux ouvrages sont publiés chez Giovanni Battista (Jean Baptiste) Seguin à Naples<sup>5</sup>, une ville qui avait été, avec Venise et Turin, l'un des centres d'édition les plus actifs pour les manuels de français au 18e siècle (MINERVA 1996 : 102) : dans la première moitié du 19e siècle, Naples est encore un centre d'édition bien

<sup>4</sup>Une « liste de phrases particulières, proverbes et gallicismes » (TRUCCHI 1834), « une longue liste de gallicismes » (TRUCCHI 1825). Notre traduction.

<sup>5</sup>En croisant nos données avec celles du répertoire des dictionnaires bilingues (LILLO 2019), nous relevons que Seguin a édité en 1846-1847 un seul dictionnaire bilingue, le *Vocabolario militare di marineria francese-italiano* de Giuseppe Parrilli. Le catalogue SBN révèle en outre qu'au moins deux grammaires du français ont paru chez Seguin dans les mêmes années, celle de Merle en 1826 et celle de De Willot en 1829.

représenté en ce qui concerne les recueils phraséologiques, mais sera devancé par Milan dans la deuxième moitié du siècle.

Il est difficile d'estimer la diffusion de ces recueils sur le territoire, mais il nous semble intéressant de prendre en compte les exemplaires recensés dans les catalogues nationaux des bibliothèques italiennes et françaises ((*OPAC SBN, Catalogo del Servizio Bibliotecario Nazionale s. d.*) et (*Catalogue Collectif de France s. d.*)) : aucun exemplaire de ces recueils n'est repérable dans les catalogues français, ce qui laisse supposer qu'ils n'ont pas circulé en France ; au contraire, dans les bibliothèques italiennes on trouve 5 copies du *Dictionnaire* de Lanza (3 dans des villes du Sud de l'Italie, 2 dans le Centre) et 8 exemplaires de la *Fraseologia* de Trucchi de 1825<sup>6</sup> (à Naples, Palerme, Bari, Potenza). Ces chiffres sont comparables avec ceux d'un dictionnaire portatif publié dans les mêmes années, comme la quatrième édition du dictionnaire de Cormon et Manni publiée à Paris en 1823, dont on dénombre 6 exemplaires dans le catalogue OPAC SBN.

### 3.2 Le titre et la composition des ouvrages

Les titres des recueils pris en compte dans ce travail reflètent les deux tendances majeures repérées dans les recueils phraséologiques franco-italiens (MURANO 2017 : 174), à savoir l'emploi comme premier mot d'un hyperonyme ou d'un terme collectif : Lanza choisit, comme l'avait fait pour la première fois Giacomo Polesi en 1829, la dénomination *Dictionnaire* « qui ferait supposer l'existence d'un ordre (alphabétique ou autre) et d'un programme d'information régulier dans [la] microstructure » (MURANO 2017 : 174) ; quant à Trucchi, il utilise à partir de 1812, premier parmi les auteurs de recueils phraséologiques, le nom collectif *Fraseologia*, qui reflète le contenu de l'ouvrage. Le mot *fraseologia*, absent des dictionnaires de la Crusca jusqu'à la quatrième édition de 1729–1738<sup>7</sup>, avait désormais acquis au début du 19<sup>e</sup> siècle le sens de 'recueil de phrases' (« raccolta di frasi, o di espressioni eleganti in qualche lingua »<sup>8</sup>, (BONAVILLA 1819)).

Loin du modèle contemporain de dictionnaire bilingue, qui comprend deux parties principales de même ampleur, chacune ayant une nomenclature dans une

<sup>6</sup>Dans (*OPAC SBN, Catalogo del Servizio Bibliotecario Nazionale s. d.*) on repère également 4 exemplaires de l'édition de 1812 et 7 exemplaires de l'édition de 1854.

<sup>7</sup>Comme le signale le *Dizionario Etimologico della Lingua Italiana* (CORTELAZZO & ZOLLI 1980)), le mot se trouve dans un ouvrage (BERGANTINI 1745) qui recense les mots utilisés par des auteurs approuvés par la Crusca, mais non enregistrés dans le *Vocabolario* : sous la forme *frasilogía*, il avait été employé par G. Caracciolo au 18<sup>e</sup> siècle avec le sens de « Arte, che infegna la scelta delle buone frasi » (« Art qui apprend le choix des bonnes phrases », notre traduction).

<sup>8</sup>« Recueil de phrases, ou d'expressions élégantes dans une langue » (notre traduction).

langue et fournissant des équivalents dans une autre langue, ces ouvrages ne présentent pas une organisation spéculaire et sont formés de plusieurs parties, qui peuvent ne pas s'apparenter au dictionnaire. La *Fraseologia* de Trucchi contient d'abord trois parties qui recueillent des *phrases* : *Frasi / Phrases*<sup>9</sup> (pp.5–49); *Frasi particolari / Phrases particulières*, (pp.50–64); *Proverbi / Proverbes* (pp. 54 –69). Les autres parties constituent des compléments variés à la grammaire, dont le poids en nombre de pages (environ la moitié du total) ne permet pas de les considérer comme du simple paratexte. Il s'agit d'un traité d'orthographe française (*Trattato di ortografia*, pp.70–82), suivi de plusieurs listes de mots : des homonymes et paronymes français avec leurs équivalents (*Catalogo di voci di quasi simile pronuncia, ma di significato o d'ortografia dissomiglianti*, pp.83–94); des mots qui ont un genre différent en italien et en français (*Catalogo di parecchi nomi sostantivi che in italiano hanno un genere, ed un altro in francese [...]*, pp.95–101); des mots français qui ont deux genres (*Sostantivi di doppio genere del medesimo significato*, pp.102–103); une liste thématique de verbes français (*Raccolta dei verbi i più usati nel discorso*, pp.104–112); des mots italiens polysémiques ou homonymiques avec leurs équivalents (*Espressioni teoriche particolari della lingua francese*, pp.113–125).

Le *Dictionnaire* de Lanza est également un ouvrage composite, dont voici la liste des sept parties : *Phrases précédés des Verbes / Frasi precedute dai verbi* (pp.7–136); *Façons de parler / Modi di spiegarsi* (pp.137–273); *Expressions adverbiales / Espressioni avverbiali* (pp.274–300); *Termes élégants / Termini eleganti* (pp.301–352); *Phrases proverbiales / Frasi proverbiali* (pp.353–378); *Proverbes / Proverbi* (pp.379–386); *Mots d'usage / Voci di usanza* (pp.387–391). Excepté les *Termes élégants* et les *Mots d'usage*, les titres des autres parties renvoient à la présentation d'items polylexicaux relevant de catégories grammaticales spécifiques (*Phrases précédés des Verbes*; *Expressions adverbiales*) ou non (*Façons de parler*), ou bien d'unités phraséologiques atteignant la dimension de la phrase (*Proverbes*; *Phrases proverbiales*).

### 3.3 Destinataires, objectifs et sources des ouvrages

L'analyse du paratexte révèle les intentions pédagogiques des auteurs et les raisons qui les ont poussés à écrire les recueils.

---

<sup>9</sup>Nous rapportons ici les titres des trois premières parties tels qu'ils sont imprimés au début de chaque partie. Dans le texte de présentation de la *Fraseologia*, à la p.4 Trucchi emploie d'autres dénominations, toutes en italien, comme *Frasi per ordine alfabetico* qui correspond à *Frasi / Phrases*.

La *Fraseologia* de Trucchi est destinée aux jeunes amateurs des lettres napolitains (« [...] la studiosa gioventù amante delle lettere » p.3, « la virtuosa gioventù Napoletana » p.4). Dans sa *Prefazione* (TRUCCHI 1825 : 3–4), l’auteur affirme qu’en dépit du large nombre d’étudiants qui entreprennent l’étude de la langue française, rares sont ceux qui la maîtrisent parfaitement, car la plupart pensent connaître la langue après avoir étudié la grammaire et appris par cœur quelques mots. La publication d’un recueil des *phrases* les plus utilisées (« più usitate ») en français et en italien, entreprise à laquelle il se consacre à partir de 1810, s’avère d’autant plus nécessaire que ce *maître de la langue française*<sup>10</sup> constate que peu de précepteurs se donnent la peine d’enseigner les *phrases* du français (« rari sono i Precettori che si danno l’incomodo d’insegnarne le frasi », (TRUCCHI 1825 : 4).

Dans son texte de présentation (*L’Autore*, pp. 3–4), Lanza dénonce les risques de la traduction mot à mot (« alla lettera », « con semplici vocaboli », « parola per parola », p.4) comme technique d’apprentissage, car elle ne peut pas mener à une connaissance parfaite de la langue française pour des apprenants italiens et au contraire est source de malentendus en compréhension et en production :

« coloro che così imparato l’avessero, se trovansi poi a parlare co’ Francesi istessi, difficile li si rende d’intendere perfettamente il di loro linguaggio, e maggior difficoltà li si appresenta se esprimer si volessero co’ medesimi traducendo parola per parola il proprio discorso » (*ibidem*)

La maîtrise de la phraséologie parfait la connaissance du français et son apprentissage par cœur (« se n’impara il difficile a memoria », *ibidem*) doit survenir vers la fin de l’apprentissage de la grammaire :

« È perciò che sul finir d’imparare la Grammatica francese, fa d’uopo studiar il Dizionario delle Frasi »

L’objectif de son Dictionnaire est de faciliter et de rendre plus rapide l’étude de la langue française (« un largo capo ho aperto nello studio di questa bella lingua, in breve tempo e con pochissima fatica », *ibidem*), car dans les textes français et dans les autres ouvrages à visée didactique, comme les dictionnaires et les dialogues, l’apprenant est obligé de « mendier » les expressions, qui s’y trouvent éparpillées sans aucun ordre.

Concernant les sources des recueils, Trucchi affirme avoir puisé dans les meilleurs auteurs classiques, alors que le frontispiece de Lanza cite « les dialogues les plus accrédités, et [...] les plus beaux traits d’éloquence de la rhétorique française ». On ne repère cependant aucune citation datée ou signée, qui nous permette de remonter à des auteurs ou ouvrages précis.

<sup>10</sup>« Maestro della lingua francese », dans le frontispiece.

Les deux auteurs affirment en outre avoir puisé dans le *Dictionnaire de l'Académie française* : vu les dates de publication, il s'agit probablement de la cinquième édition de 1798 pour Trucchi, qui publie la *Fraseologia* en 1812 et en 1825, et de la sixième édition de 1835 pour Lanza, qui publie le *Dictionnaire des phrases* en 1837.

### 3.4 La macrostructure : taille de la nomenclature, critères de rangement des entrées

En prenant en compte uniquement les parties que ces ouvrages consacrent à la phraséologie (les trois premières dans Trucchi, toutes sauf les *Termes élégants* et les *Mots d'usage* dans Lanza), on peut estimer la nomenclature de Trucchi autour de 1800 *phrases* en 69 pages ; celle de Lanza s'élève à presque 3700 *phrases*<sup>11</sup> en 334 pages.

Du point de vue de l'organisation typographique de la page, dans les deux recueils le texte est imprimé sur deux colonnes, dont celle de gauche est en italiques et celle de droite en caractères droits. Une lettrine est présente au-dessus des colonnes de texte dans la *Fraseologia* de Trucchi, dont la densité du texte dans la colonne explique le nombre si élevé de phrases dans un nombre de pages limité.

En ce qui concerne les critères de rangement des entrées dans les recueils phraséologiques, Lillo a pu affirmer que « ces recueils sont le plus souvent agencés en fonction de regroupements plus ou moins arbitraires de phrases et constituent des amas d'éléments plutôt disparates » (LILLO 1994 : 76). En effet, comme nous avons pu le constater lors de notre étude sur les recueils de Polesi et Barberi (MURANO 2020), seule une analyse fine de chaque ouvrage et de chaque partie d'ouvrage permet de déceler quelques critères, qui peuvent être appliqués de manière plus ou moins systématique.

Dans la *Fraseologia* de Trucchi, la colonne de gauche présente des items en langue italienne, la colonne de droite leurs équivalents en français.

La section *Phrases* contient des séquences à base verbale, plus ou moins figées et compositionnelles, comprenant des locutions (« *Far castelli in aria* – Faire des châteaux en Espagne »), des collocations (« *Indur sonno* – Faire venir le sommeil » ; « *Innalzar una Statua* – Dresser une Statue »), des séquences libres (« *Giuocar alla palla* – Jouer à la paume »). Les entrées sont rangées en suivant un

---

<sup>11</sup>Ce calcul approximatif a été fait en considérant le nombre moyen des phrases présentes dans les dix premières pages, 28 environ dans la *Fraseologia* de Trucchi et 11 environ dans le *Dictionnaire* de Lanza, dans lequel les interlignes sont plus grands.

ordre alphabétique strict selon le premier mot, puis selon le deuxième mot<sup>12</sup> : par exemple à la page 30 (voir Figure 1) *Girar gli occhi di quà e di là* précède *Giugner in porto*, qui précède à son tour *Giugner nuova alcuna cosa*.

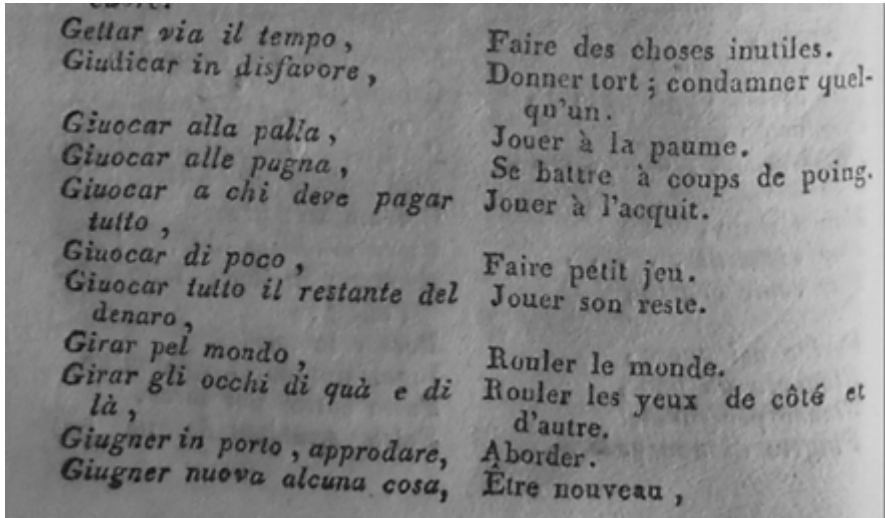


FIGURE 1 : Fraseologia italiana-francese, p.30

La section dénommée *Proverbes* contient en réalité des séquences figées aussi bien phrastiques et proverbiales (« *E' meglio un ovo oggi, che domani una gallina* – Il vaut mieux un tiens, que deux tu l'auras ») que verbales (« *Far vedere la luna nel pozzo* – En donner à garder »), ces dernières étant bien moins nombreuses. Par rapport à la section précédente, on ne repère ici que des séquences verbales dont le sens est figuré et non compositionnel. L'ordre alphabétique y est également adopté, mais de manière bien moins rigoureuse concernant le deuxième mot : *Cane che abbaja, non morde mai* précède tous les proverbes en *Chi...*, mais parmi ceux-ci, « *Chi compra terra, compra guerra* précède *Chi cerca d'ingannare resta ingannato*.

Aucun ordre, ni alphabétique, ni thématique, n'est repérable par contre dans la partie *Phrases particulières*, que l'on prenne en compte les entrées ou les équivalents. Cette section, appelée *Frase particolari, e gallicismi* dans la *Prefazione*, présente un contenu plutôt hétéroclite, allant des proverbes (« *Chi vuole acquistare bisogna che affatichi* – Nul bien sans peine ») aux séquences adverbiales (« *Nel maggior impeto della collera* – Dans le fort de la colère ») aux noms composés

<sup>12</sup>Quelques hésitations sont présentes concernant l'ordre alphabétique du deuxième mot : par exemple, *Giuocar alle pugna* précède *Giuocar a chi deve pagar tutto* (Voir Figure 1).

(« *Carta suga* – Papier qui boit »). Les séquences phrastiques sont toutefois les plus nombreuses : la plupart sont libres (« *Egli è uomo aspro* – C'est un homme brusque »), mais on repère quelques séquences figées verbales présentées en discours (« *Ha il cervello sopra la berretta* – Il a la cervelle au bout du pied »). Les items phrastiques libres et compositionnels en italien présentent souvent un équivalent français figé et non compositionnel, ce qui est probablement à l'origine de leur enregistrement dans le recueil :

*Sono tutto sudato, tutto bagnato di sudore* – Je suis tout en nage.

*E' molto vecchio, è cadente* – Il est sur le bord de la fosse.

*E' ridotto alla mendicITÀ* – Il est réduit au bâton blanc à la main.

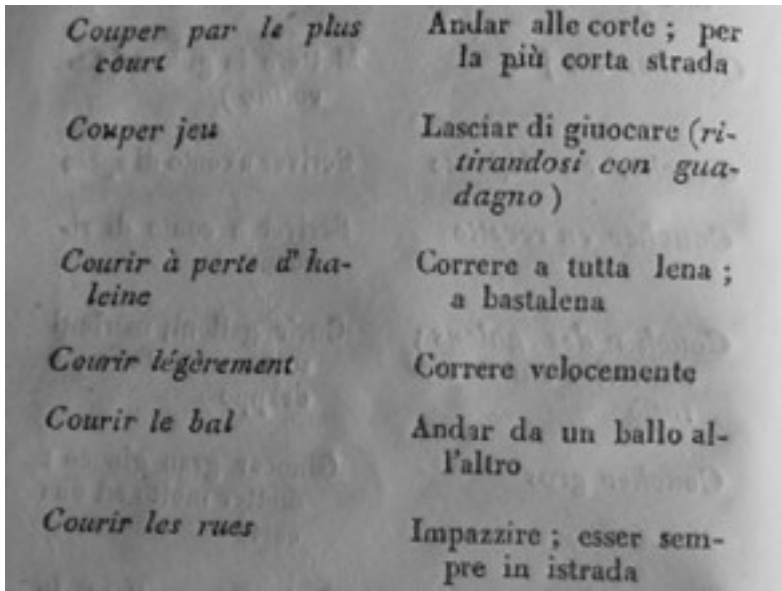
Dans le *Dictionnaire* de Lanza, les *phrases* françaises se trouvent dans la colonne de gauche et les équivalents italiens dans la colonne de droite.

Les *Phrases précédées des Verbes* sont des séquences verbales, rangées selon l'ordre alphabétique du premier mot, qui est toujours un verbe à l'infinitif, puis, avec quelques hésitations, selon l'ordre alphabétique du deuxième mot : dans la Figure 2 on peut voir que les séquences qui commencent par *Couper* précèdent celles qui commencent avec *Courir*, mais *Couper jeu* suit, au lieu de précéder, *Couper par le plus court*. Cette section contient des locutions verbales non compositionnelles (« *Couver des yeux* – *Mirar con occhio affettuoso e contento* »), des verbes simples qui ont un équivalent polylexical en italien (« *se Dégarnir* – *Alleggerirsi di panni* »), ainsi que des séquences dont le sens est compositionnel (« *Avoir le visage effilé* - *Aver il volto sottile; sfilato* »).

L'ordre alphabétique est suivi également dans toutes les autres parties, à commencer par les *Façons de parler*, qui regroupent des séquences phrastiques avec un verbe conjugué, aussi bien figées (« *Il l'a manqué belle* – *Egli ha schivato un gran pericolo* ») que libres (« *As-tu songé à cela?* – *Hai tu pensato a ciò?* »). Les *Phrases adverbiales* comprennent des items relevant de plusieurs catégories grammaticales : locutions adverbiales (*D'un seul coup*, *En abrégé*), prépositionnelles (*Depuis que*), pronoms indéfinis (*Qui que ce soit*), phrases impersonnelles (*Il s'en est peu fallu*).

Les *Termes élégants* sont pour la plupart des noms composés, comme *Heures perdues* (it. « *Ore libere, d'ozio* »), *Homme à projet* (it. « *Progettista* »), *Jours de Palais* (it. « *Giorni di Tribunali, giorni non feriat* »).

La section *Phrases proverbiales* contient des séquences figées verbales présentées avec le verbe à l'infinitif (« *Marcher à pas de loup* - *Caminar quatto quatto* » ; « *Mettre le doigt entre le bois et l'écorce* – *Trovarsi tra l'incudine e 'l martello* »),



|                                  |                                                         |
|----------------------------------|---------------------------------------------------------|
| <i>Couper par le plus court</i>  | Andar alle corte ; per la più corta strada              |
| <i>Couper jeu</i>                | Lasciar di giuocare ( <i>ritirandosi con guadagno</i> ) |
| <i>Courir à perte d' haleine</i> | Correre a tutta lena ; a bastalena                      |
| <i>Courir légèrement</i>         | Correre velocemente                                     |
| <i>Courir le bal</i>             | Andar da un ballo all'altro                             |
| <i>Courir les rues</i>           | Impazzire ; esser sempre in istrada                     |

FIGURE 2 : Dictionnaire des phrases françaises-italiennes, p.22.

alors que les *Façons de parler proverbiallement* sont des séquences figées verbales présentées avec le verbe conjugué (« *C'est un roseau qui plie à tout vent* – Egli è un uomo volubile ») ou des séquences figées phrastiques (« *Votre affaire est faite* – Ella è finita per voi ; non bisogna darvi più speranza »).

La section *Proverbes* contient de véritables proverbes comme « *Le loup mourra dans sa peau* – Il lupo cangia il pelo, ma non il vizio » ou « *À brebis tondue, Dieu mesure le vent* – Dio manda il freddo, secondo i panni ».

Les pages finales de la section *Mots d'usage* contiennent des formules utilisées dans différentes situations de communication, comme la chasse ou le jeu :

*Chou chou là* – Badalo, badalo, badalo (*termine, che si usa da'cacciatori, per eccitar il cane alla ricerca, che significa Piglia, piglia (detto popolarmente)*)

*Coup de malheur* – Colpo disgraziato (*espressione usata da' giuocatori allorché perdono*)

*Je m'y tiens* – Non ne voglio più (*risposta ad alcuni giochi di carte che si danno una per volta*)

La présentation des sections qui composent la *Fraseologia italiana-francese* et le *Dictionnaire des phrases françaises-italiennes* met en évidence l'hétérogénéité

des items qu'elles recensent et la volonté de rendre disponible aux lecteurs l'ensemble des unités phraséologiques du français, quel que soit leur degré de figement formel, sémantique ou pragmatique.

Les dénominations des différentes sections nous paraissent délimiter de manière imparfaite leurs contenus. Toutefois, il se peut qu'elles reflètent les emplois de l'époque 1820–1840 et s'avèrent pour cela précieuses : sous l'étiquette *Façons de parler*, Lanza réunit des séquences avec un verbe conjugué, ce qui les oppose aux *Phrases*, qui présentent le verbe à l'infinitif aussi bien dans le *Dictionnaires des phrases* de Lanza que dans la *Fraseologia* de Trucchi<sup>13</sup>.

Tout comme Lillo l'avait repéré dans les manuels, dans ces recueils les proverbes « sont souvent confondus avec les expressions imagées » LILLO (1994 : 71). La dénomination *Proverbe* avait probablement un sens plus large que celui que nous lui attribuons aujourd'hui, car elle désigne chez Trucchi des séquences aussi bien phrastiques que verbales : la preuve en est que chez Lanza, l'adverbe *proverbialmente* et l'adjectif *proverbial* ont le sens de 'figé et non compositionnel' lorsqu'ils sont ajoutés aux dénominations *Phrases et Façons de parler*.

### 3.5 La microstructure

Les recueils phraséologiques de Trucchi et Lanza ne sont pas composés d'articles de dictionnaire au sens moderne du terme : ils se limitent à mettre en regard les séquences plus ou moins figées italiennes et françaises, sans qu'un véritable bloc de texte ayant une structure hiérarchique et typographique interne soit consacré à chaque item. Cette composition typographique fondée sur le dénombrement des adresses était typique de la lexicographie ancienne, ainsi que l'a montré QUEMADA (1967 : 467), et avait été employée dans la lexicographie bilingue français-italien jusqu'au 18e siècle (MURANO 2013 : 85–6).

Dans la *Fraseologia* de Trucchi, on repère dans de rares cas quelques éléments typiques de la microstructure lexicographique, aussi bien dans la colonne de gauche qui contient les entrées que dans celle de droite qui concerne les équivalents :

- Des variantes formelles introduites par la conjonction *o*, comme dans « *Darsi intieramente in preda, o in balia alle passioni* – Lâcher la bride à ses passions » ou dans le proverbe « *Tanto va la gatta al lardo, che vi lascia lo zampino* [...] – Tant va la cruche à l'eau qu'à la fin elle se casse ou qu'enfin elle se brise ».

---

<sup>13</sup>Sauf dans la section *Phrases particulières*, comme nous l'avons vu plus haut.

- Des gloses qui expliquent le sens des entrées, éventuellement introduites par la formule métalinguistique d'explication *vale*<sup>14</sup> : par exemple « *Far la nanna, vale dormire* – Faire dodo » et « *S'ella è Rosa fiorirà, s'ella è spina pungerà, vale il tempo scuopre la verità* – La fin découvrira le vrai ».
- Des notes en bas de page, aux contenus divers : la note reliée à *Fare la nanna* contient la glose « *Frase usata dalle balie per addormentar i bambini, dicendo ninna nanna.* » ; pour *Essere un ciarlone*, la note reliée à l'équivalent « être bavard » indique la connotation négative de l'expression : « *Prendesi sempre in cattiva parte* ».
- Des locutions ou des équivalents synonymiques de forme proche ou éloignée, introduits par la conjonction *o* ou tout simplement séparés par le point virgule, par exemple « *Aver poco sale in zucca, o aver la zucca vuota* – N'avoir pas un grain de sel dans sa caboche » ou « *Esser fuori di sé; essere trasportato dalla collera* – Ne se connaître point ; n'être pas à soi ».

De même, des éléments de microstructure sont présents dans le *Dictionnaire des phrases* de Lanza. Dans la colonne de gauche qui contient les *phrases* - entrées en français, on repère essentiellement des variantes, comme dans « *Conter des sornettes; conter des fagots; ou conter de belles; en conter bien* – Contar baja [...] ».

Dans la colonne de droite, qui contient les équivalent(s) en italien, les éléments de microstructure sont plus nombreux et variés :

- des gloses explicatives, par ex. « *se Marier par amourette* – prender moglie per innamoramento (*matrimonio ineguale*) » ,
- des variantes, par ex. « *Mettre quelqu'un en cervelle; le tenir en cervelle* – Metter altrui il cervello a partito; recarlo in dubbio; farlo star sospeso » ; « *Nager entre deux eaux* – Star sospeso; in dubbio » ,
- des restrictions de sujet, par ex. « *Avoir un beau toucher, un toucher délicat, brillant* – Suonar con eccellenza (*si dice de' sonatori d'istrumenti di tasti*) » ; « *Mettre bas* – partorire (*parlando di animali*) ». Ces indications s'avèrent particulièrement utiles dans le cas de *phrases* polysémiques, comme *Flotter au gré; à la merci du vent*, enregistrée deux fois, une première fois avec l'équivalent « *Essere portato in balia del vento (per un vascello mal governato)* », puis avec l'équivalent « *Ondeggiar sulle spalle agitati dal vento (parlando de' capelli sciolti)* »,

---

<sup>14</sup>Dans un cas, on repère la formule métalinguistique *Si dice anche*, qui paraît suivre l'explication et précéder la locution verbale : « *Accertarsi di checchessia con dimostrazione certissima; si dice anche toccar con mano* – Toucher au doigt et à l'œil ».

- des marques diatechniques (caccia, giuoco di carte / di bigliardo, guerra, marina, maneggio, scherma etc.), par ex. « Ménager bien sa voix – Regular bene la sua voce (*musica*) »; « Méler un cheval – imbrogliare un cavallo (*maneggio*) ».

## 4 Conclusion

L'analyse de deux recueils publiés à Naples entre 1825 et 1837 nous a permis d'approfondir notre réflexion récente (MURANO 2017 et 2020) sur les recueils phraséologiques franco-italiens. Si le critère de choix des recueils a été éminemment chronologique au départ, car nous voulions étudier la production phraséographique dans les décennies 1820 et 1830, il s'est avéré intéressant d'examiner deux ouvrages publiés chez le même éditeur qui ont vraisemblablement circulé simultanément dans les mêmes régions d'Italie.

Nous avons d'abord présenté la production de chaque auteur, qui dans le cas de Lanza paraît limitée au seul *Dictionnaire des phrases françaises-italiennes*, comme c'était le cas pour Polesi (cf. MURANO 2020). Des informations précieuses dans le paratexte et une recherche dans les catalogues des bibliothèques nationales et universitaires laissent supposer une bonne diffusion des recueils phraséologiques, ou tout du moins comparable à celle de quelques dictionnaires portatifs, même si nous ne possédons aucune indication précise sur le public visé (écoles ou étude en autonomie).

La composition de ces ouvrages nous confirme dans le choix du mot *recueil* pour les désigner, car l'absence d'unité textuelle et la fragmentation en sections hétérogènes éloignent ces ouvrages de la conception traditionnelle de dictionnaire : la dénomination *recueil* nous sert alors d'hyperonyme et embrasse la variété des ouvrages qui ont pour objectif de mettre en regard les phraséologies française et italienne. Les deux ouvrages analysés dans cette recherche montrent bien ce statut hybride et fluctuant entre texte de lecture et texte de consultation : alors que la structuration typographique de la page en séries d'items successifs suggérerait une lecture suivie, l'ordre alphabétique est adopté dans presque toutes les sections, ce qui permet une recherche ponctuelle des informations.

Concernant les raisons qui ont motivé l'écriture des recueils, les auteurs constatent l'absence d'enseignement de la phraséologie par les maîtres et la nécessité de l'apprentissage de la phraséologie pour la traduction et la conversation. Complément nécessaire de l'étude de la grammaire au niveau avancé de l'apprentissage, le recueil phraséologique est l'instrument qui permet d'apprendre par cœur ce qui est *difficile* dans la langue, pour que l'apprenant ne soit pas tenté de faire une traduction mot à mot.

Du point de vue des items linguistiques présentés dans la *Fraseologia* et dans le *Dictionnaire des phrases*, on remarque que le contenu des recueils n'est pas entièrement phraséologique au sens moderne du terme, ce qui impliquerait la dimension du figement : les *phrases* peuvent être figées ou libres, leur sens peut être compositionnel ou non compositionnel. Par contre, l'emploi du nom *proverbe* (et de ses dérivés *proverbial* et *proverbialement*), qui s'étend au-delà de la désignation de séquences phrastiques, indique sans hésitation des séquences caractérisés par le figement, syntaxique et/ou sémantique.

Dans ces deux recueils, on repère des éléments de microstructure ponctuant les textes de manière irrégulière : les variantes des *phrases* et de leurs équivalents sont nombreuses dans les deux ouvrages ; au contraire, les paraphrases en langue source sont rares car l'objectif premier est de poser la relation d'équivalence entre les *phrases* des deux langues ; dans le *Dictionnaire des phrases* de Lanza on peut repérer également des restrictions de sujet et de nombreuses marques diatechniques qui témoignent de l'ouverture vers la phraséologie des langues de spécialité. Ce dernier trait avait déjà repéré par Lillo, qui affirmait que les recueils phraséologiques ne dédaignaient plus « les expressions burlesques, familières, ni les langues de spécialité » (LILLO 1994 : 76). Cette citation nous amène à constater l'absence quasi-totale<sup>15</sup> de marques diastratiques et diaphasiques ou d'une section consacrée à la *phraséologie familière*.

Les recherches ultérieures que nous comptons conduire sur d'autres recueils de la même période et postérieurs nous permettront de dresser un état de lieu plus complet sur la phraséographie bilingue franco-italienne au 19e siècle et sur le rôle que les recueils phraséologiques pouvaient jouer dans l'enseignement de la phraséologie (phraséodidactique) et en général dans l'enseignement du français à un public pour lequel la maîtrise de cette langue était essentielle, comme le rappelle Nicolas Lanza :

« lo studio di essa potrebbe con ragion dirsi esser una essenziale parte della buona educazione, facendo imparare a' Giovani una lingua che potrebbe loro ben servire nelle diverse vicende della vita in estero paese, e nella propria patria esser li può di gran risorsa perché li facilita il mezzo di trattar con esteri, ed apprendere le di loro cognizioni, come pure far loro intendere le proprie idee parlando e scrivendo in tal idioma »<sup>16</sup> (LANZA 1837 : 3).

<sup>15</sup>Voir plus haut dans le *Dictionnaire des phrases*, « detto popolarmente » qui s'applique à l'équivalent s.v. *Chou chou là*.

<sup>16</sup>« Son étude [de la langue française] pourrait à juste titre être considérée comme une composante essentielle de la bonne éducation, car on ferait apprendre aux jeunes une langue qui pourrait leur servir dans les situations de la vie dans un pays étranger, et qui, dans leur pays,

## Références

- AUTELLI, Erica. 2022. Les débuts de la phraséologie et les premières « Phraséologies historiques » italo-françaises. *Linguistik online* 113. 17-43. DOI : 10.13092/lo.113.8316.
- BARBERI, Giuseppe Filippo. 1821. *Petit trésor de la langue française et de la langue italienne, ou des différentes figures, appelées tropes, de la Langue Française et de la Langue Italienne, les unes correspondantes aux autres*. Paris : J.-P. Aillaud.
- BÁRDOSI, Vilmos. 2017. *Du phrasème au dictionnaire : Études de phraséographie franco-hongroise*. Budapest : Eötvös Kiadó.
- BERGANTINI, Gian Pietro. 1745. *Voci italiane d'autori approvati dalla Crusca, nel Vocabolario d'essa non registrate [...]* Venezia : Pietro Bassaglia.
- BONAVILLA, Aquilino. 1819. *Dizionario etimologico di tutti i vocaboli usati nelle scienze, arti e mestieri, che traggono origine dal greco*. Milano : Tipografia di Giacomo Pirola.
- Catalogue Collectif de France*. s. d. <https://ccfr.bnf.fr/portailccfr/jsp/public/index.jsp?success=%2Fjsp%2Fpublic%2Findex.jsp&failure=%2Fjsp%2Fpublic%2Ffailure.jsp&profile=public>.
- CORTELAZZO, Manlio & Paolo ZOLLI. 1980. *Dizionario Etimologico della Lingua Italiana*. Bologna : Zanichelli.
- DUBOIS, Jérémie. 2015. *L'enseignement de l'italien en France (1880-1940). Une discipline au cœur des relations franco-italiennes*. Grenoble : ELLUG.
- LANZA, Nicolas. 1837. *Dictionnaire des phrases françaises-italiennes*. Napoli : Se-guin.
- LILLO, Jacqueline. 1994. La phraséologie dans les manuels de français publiés en Italie de 1625 à 1860. *Documents pour l'histoire du Français Langue étrangère et seconde* 14. 70-81. DOI : 10.3406/docum.1994.1067.
- LILLO, Jacqueline. 2014. *Un corpus phraséologique bilingue (français-italien) d'Ancien Régime : Le manuscrit de Ludovico Goudar*. Franck NEVEU, Peter BLUMENTHAL, Linda HRIBA, Annette GERSTENBERG, Judith MEINSCHAEFER & Sophie PRÉVOST (éd.). [http://www.shs-conferences.org/articles/shsconf/abs/2014/05/shsconf\\_cmlf14\\_01197/shsconf\\_cmlf14\\_01197.html](http://www.shs-conferences.org/articles/shsconf/abs/2014/05/shsconf_cmlf14_01197/shsconf_cmlf14_01197.html).
- LILLO, Jacqueline (éd.). 2019. *1583-2010 : Quattro secoli e più di lessicografia italo-francese. Repertorio analitico di dizionari bilingue*. 2<sup>e</sup> éd. Bern : Peter Lang.
- MINERVA, Nadia. 1996. *Manuels, maîtres, méthodes. Repères pour l'histoire de l'enseignement du français en Italie*. Bologna : CLUEB.

---

peut s'avérer très utile parce qu'elle facilite les relations avec les étrangers et la compréhension des connaissances de ces derniers, et leur [= aux jeunes] permet également de faire comprendre leurs idées en parlant et en écrivant dans cette langue » (notre traduction).

- MINERVA, Nadia. 1998. Les dialogues dans les manuels italiens. Enquête quantitative (1625-1860). *Documents pour l'histoire du français langue étrangère ou seconde* 22. 83-98. DOI : 10.3406/docum.1998.1251.
- MURANO, Michela. 2013. *Des phrases aux séquences figées : La phraséologie dans les dictionnaires bilingues français-italien (1584-1900)*. Bologna : CLUEB.
- MURANO, Michela. 2017. Les recueils phraséologiques bilingues franco-italiens au 19e siècle. In Cosimo DE GIOVANNI (éd.), *Fraseologia e paremiologia : Passato, presente, futuro*, 165-184. Milano : Franco Angeli.
- MURANO, Michela. 2020. Connaître la phraséologie « pour bien entendre les auteurs » : Le 'Petit trésor' de Barberi (1821) et le 'Dictionnaire des idiotismes de Polesi' (1829). In Joanna SZERSZUNOWICZ (éd.), *Intercontinental Dialogue on Phraseology 4 (IDP), Reproducible language units in an interdisciplinary perspective*, 283-301. Białystok : University of Białystok Publishing House.
- OPAC SBN, *Catalogo del Servizio Bibliotecario Nazionale*. s. d. <https://opac.sbn.it/web/opacsbn>.
- PELLANDRA, Carla. 1997. Que lisait-on dans les classes de français d'autrefois ? (Les contenus culturels de quelques manuels pour l'enseignement du français publiés en Italie de 1846 à 1908). *Documents pour l'histoire du français langue étrangère ou seconde* 12. 32-38.
- POLESI, Giacomo. 1829. *Dictionnaire des idiotismes Italiens-Français et Français-Italiens*. Paris : Baudry.
- QUEMADA, Bernard. 1967. *Les dictionnaires du français moderne : 1539-1863*. Paris : Didier.
- SCLAFANI, Marie-Denise. 2019. Le figement dans le 'Nuovo vocabolario commerciale fraseologico de Caricati et Mariotti'. *Les cahiers du dictionnaire* 11. 201-215.
- TRUCCHI, Giuseppe. 1825. *Fraseologia italiana-francese*. 2<sup>e</sup> éd. Napoli : Seguin.
- TRUCCHI, Giuseppe. 1831. *Supplemento alla Grammatica del Goudar ed a parecchie altre [...]* Napoli : Raffaele Miranda.
- TRUCCHI, Giuseppe. 1834. *Raccolta di varj dialoghi tratti dall'illustre penna de' più eleganti gramatici francesi [...]* Napoli : presso Velizza e Ravallese-Gargiulo.
- TRUCCHI, Giuseppe. 1837. *Il Filogallo zelante ossia Ampliato supplemento alla maggior parte delle grammatiche italiane-francesi [...]* Napoli : da' torchi del Tramater.

## Chapter 20

# The *Lexicon of Oriental Words in Ancient Greek*

Rafał Rosół

Adam Mickiewicz University in Poznań

The aim of this chapter is to present the *Lexicon of Oriental words in Ancient Greek*, which currently being prepared in both electronic and book form. The lexicon will contain common words (loanwords and glosses) that derive from Anatolian, Armenian, Austroasiatic, Cushitic, Dravidian, Egyptian, Hurro-Urartian, Indo-Aryan, Iranian, Kartvelian, Phrygian, Semitic, Sino-Tibetan and other Eastern languages. It will also include vocabulary of uncertain provenance (especially *Wanderwörter*). The data will cover two millennia of the history of the Greek language, i.e. from the Mycenaean period to the 6th cent. CE.

The *Lexicon of Oriental Words in Ancient Greek*, which is currently in preparation, will have an electronic and a printed version. The entries will look the same in both versions, but there will be some differences largely due to their specific nature. The former will provide tools for searching and grouping words, while the latter will contain a much expanded introduction and indexes of linguistic data. The online lexicon will be fully completed in 2025, but is already partially available (<https://orientalwordsgivingreek.org/>). Shortly afterwards, work will begin on preparing the book for publication.

The goal of the project is to study words borrowed by the Greeks from Oriental languages, i.e. languages used in Asia (along with the European part of Scythia and the Caucasus region), Egypt and, because of the ancient maritime trade route from India to Europe, on both sides of the Gulf of Aden and the Red Sea. Accordingly, the vocabulary under investigation will originate in Anatolian, Indo-Aryan, Iranian, Phrygian, Armenian, Semitic, Egyptian, Cushitic, Hurro-Urartian, Kartvelian, Dravidian, Austroasiatic, Sino-Tibetan, etc. It will also encompass vocabulary without a determined source (especially *Wanderwörter*). The



data will cover two millennia of the history of the Greek language, i.e. from the Mycenaean period (the Linear B tablets) to the 6th cent. CE.

The lexicon will generally contain only common nouns. However, it will also include words that are not proper nouns in principle, but were treated as such by Greek authors and therefore usually appear in capital letters in modern editions, e.g. Βραχμᾶνες m. pl. ‘Indian sages, philosophers (often treated as a tribe)’, which goes back to Sanskrit *brāhmaṇa-* ‘one who has divine knowledge; a Brahman’ (cf. *brahman-* ‘the class of men who are the repositories and communicators of sacred knowledge, the Brāhmanical caste as a body’). Moreover, words for which it is difficult to determine whether they are proper nouns or common nouns will also be taken into account.

It should be noted that common nouns derived from proper nouns, if they are formed according to the rules of Greek morphology, will be excluded, e.g. Δαρεικός m. ‘a Persian gold coin’ (from Δαρειός). On the other hand, words like Μόσυλον (also μόσυλλον) n. ‘a kind of cinnamon of fine quality’ (cf. Μοσυλίτις f. ‘a kind of cinnamon’) will be put into the lexicon; this term derives from the name of the city (and the cape) Mosyl(l)on located on the northern coast of Somalia and attested in Greek sources as Μόσυλον or Μόσυλλον (cf. Μόσυλοι, the name of a people living nearby),<sup>1</sup> but morphologically it does not show a Greek derivation with any of the possible suffixes.

An important part of the project are Oriental glosses that appear in Greek sources. We are referring to words that ancient and early Byzantine authors explicitly identify as Egyptian, Phoenician, Carian, etc., e.g. the plant name μίθ ‘celery (*Apium graveolens* L.)’ that is quoted in the appendix (the so-called *Recensio Vindobonensis*) to *De materia medica* by Pedanius Dioscorides (3.64 RV):

σέλινον κηπαῖον· [...] Αιγύπτιοι μίθ [...].

*Selinon kēpaion* (celery): [...] the Egyptians [call it] *mith* [...]

This gloss comes from Egyptian *m3t.t* ‘id.’ (cf. Coptic *mit* (Sahidic and Bohairic), *emit* (Bohairic) ‘parsley, celery’). Another example is the gloss ἄγλυ ‘swan’ transmitted by Hesychius of Alexandria (α 621):

<sup>1</sup>That cinnamon was imported from Mosyl(l)on is attested to by a reference in the *Periplus of the Erythraean Sea* (10): Ἐξάγεται δὲ ἀπὸ τῶν τόπων [scil. ἀπὸ τοῦ Μοσύλλου] κασίας χύμα πλεῖστον – (<δι>ὸ καὶ μειζόνων πλοίων χρήζει τὸ ἐμπόριον – καὶ ἄλλη εὐωδία καὶ ἀρώματα καὶ χελωνάρια ὀλίγα καὶ θυμίαμα μοκρότου [...] καὶ λίβανος ὁ περατικός, ἐλέφας δὲ καὶ σμύρνα σπανίως. – “Exports from this area [scil. from Mosyllon] are: a great quantity of cassia (for this reason the port requires bigger ships); other spices and aromatics; a little low-quality tortoise shell; *mokrotu* incense [...]; “far-side” frankincense; ivory and myrrh but only on rare occasions.” (trans. L. Casson).

ἄγλυ· ὁ κύκνος, ὑπὸ Σκυθῶν.

*agly*: swan; [it is called so] by Scythians.

The word still has no convincing etymology. A large group of glosses are transliterations of Hebrew or Aramaic words found in the Septuagint (and in other Greek translations of the Old Testament), the New Testament, in the writings of Joseph Flavius and in other texts related to Judaism and Christianity. The following words may be cited as examples: γαι ‘valley’ (LXX, Ez. 39.11, 39.15) from Biblical Hebrew *gay* ‘id.’ and ταλιθα κουμ ‘little girl, get up!’ (NT, Mar. 5.41) from Aramaic \**ṭaliṭā qūm* ‘id.’

To date, no attempt has been made to create a lexicon containing all the possible Oriental loanwords and glosses found in Greek sources. Although there are many important books and meticulously researched papers devoted to the subject, there are still significant gaps. Of course, we can rely on some works concerning loanwords and glosses from Semitic,<sup>2</sup> Indo-Iranian,<sup>3</sup> Egyptian,<sup>4</sup> Anatolian<sup>5</sup> or Phrygian.<sup>6</sup> However, many aspects of Oriental borrowings and glosses in ancient Greek have not been sufficiently studied and described. There is still no up-to-date monograph on Semitic words attested in the Hellenistic and Roman periods.<sup>7</sup> The lack of a monograph concerning Egyptian words in Greek must also be emphasised, since the papers mentioned above comprise only a partial material. There is also no comprehensive work on “wandering” words in antiquity or on Oriental words with undefined source language. Moreover, it is worth noting that the existing etymological dictionaries of Greek<sup>8</sup> are often inadequate in their treatment of the problem of Oriental borrowings and very often do not represent the current state of research, as well as they do not include many foreign words and the vast majority of glosses. All of this points to the need for a comprehensive lexicon of the Oriental vocabulary that appears in ancient and early Byzantine Greek sources.

<sup>2</sup>E.g. Masson 1967, Bai 2009, Rosól 2013 and Rosól 2018, as well as an important publication on Hebrew and Aramaic words in the Septuagint by Simotas [Σιμωτάς] 1968 .

<sup>3</sup>The most important publication on the subject is the comprehensive book by Brust 2008 ; moreover, see e.g. Schmitt 1967: Hinz 1975: Huyse 1990 .

<sup>4</sup>First of all, the paper by Fournet 1989 ; furthermore, some overviews: Hemmerdinger 1968, McGready 1968, Torallas Tovar 2004 ; cf. also Pierce 1971 .

<sup>5</sup>E.g. Neumann 1961 and an excellent overview by Simon 2018 ; the glosses from Anatolian languages, i.e. from Carian, Lydian, etc., can be found in the digital philological-etymological dictionary *eDiAna*: <https://www.ediana.gwi.uni-muenchen.de/>.

<sup>6</sup>Haas 1960, Sowa 2008: 39–68, Obrador Obrador-Cursach 2020: 412–424.

<sup>7</sup>Cf. old works by Lewy 1895 and Muss-Arnolt 1892 .

<sup>8</sup>Frisk 1960: 72, Chantraine 1999, Beekes 2010 .

Regarding the structure and composition of entries in the lexicon, they will consist of 4 or, if sufficient data is available, 5 sections. The basic section will be a headword with three types of tags. The first indicates the general provenance of a given word: Anatolian, Armenian, Austroasiatic, Cushitic, Dravidian, Egyptian, Hurro-Urartian, Indo-Aryan, Iranian, Kartvelian, Phrygian, Semitic, Sino-Tibetan or, simply, Unspecified origin. The second tag refers to the earliest attestation of a word in Greek texts. Accordingly, the words will be classified into six historical periods: Mycenaean (15th–12th cent. BCE), Archaic (8th–6th cent. BCE), Classical (5th–4th cent. BCE), Hellenistic (3rd–1st cent. BCE), Roman (1st–4th cent. CE), and Early Byzantine (5th–6th cent. CE). The third tag allows a general differentiation of the word type: loanword or gloss. In doubtful cases, more than one tag may appear in a given category, e.g. Iranian?/Semitic?, Classical/Hellenistic period, Loanword/Gloss. The information in this section will be particularly useful when searching the electronic version of the dictionary, as the browsing tools will enable to search for words in two different ways: a) by typing a headword in the search box (in the Greek alphabet); b) by selecting tags. More tags and search tools will be introduced in the future.

The second section (marked ↗) will present word forms (including grammatical features and, where possible, vowel length, together with definitions and sources), as well as variants and *variae lectiones* in manuscripts. An entry will also contain the most important derivatives and compounds. In addition, this section will draw attention to possible difficulties in determining the exact meaning of a word, such as plant identification of phytonyms or measurements for unit names. In the case of rare words (especially, *hapax legomena*), suggestions for conjunctures will also be cited. Furthermore, possible Latin counterparts will be given, if a word (usually being a borrowing from Greek) appears in Latin sources.

The third section (marker: △) is something new in lexicographic practice. It will contain quotations and pieces of information that may directly or indirectly point to possible Oriental roots of a word. Each Greek (or Latin) quotation will be provided with an English translation. For instance, the gloss διβιρουπτ m. ‘chief scribe (an Iranian title)’ (attested in the inscription abbreviated as ŠKZ, v. 57) will be accompanied only by the following comment:

△ ŠKZ is a trilingual, i.e. Middle Persian, Parthian and Greek inscription of Shapur I written on the walls of the Ka’ba-ye Zartosht (Cube of Zoroaster) in Naqsh-e Rostam (ca. 12 km from Persepolis).

In the case of βωσαρή f. ‘a kind of elephant (in India)’, a relevant passage from the *Periplus of the Erythraean Sea* will be quoted in Greek and English:

△ *PME* 62: Ἀπὸ δὲ ταύτης [scil. Μασσαλίας] εἰς αὐτὴν τὴν ἀνατολὴν διαπεράσαντι τὸν παρακείμενον κόλπον ἢ Δησαρηνὴ χώρα, φέρουσα ἐλέφαντα τὸν λεγόμενον βωσαρῆ [...].

If you go due east from it [scil. from the region of Masalia, i.e. Machilipatnam] across the bay that lies alongside, you come to the Desarene region [i.e. probably Odisha], the habitat of a kind of elephant called *bōsarē*. (trans. L. Casson).

Of course, not every entry will have this section, but in some cases it can be quite extensive (even two or three pages long), depending on the material available.

The fourth section (marker: ●) will provide results of etymological investigations along with extensive linguistic data from Oriental languages, e.g. ἀγάλοχον (also ἀγάλλοχον) n. ‘agarwood (a fragrant wood obtained from some tree species of the genus *Aquilaria*, especially *Aquilaria malaccensis* Lam.)’ is explained in this way:

● An Oriental word, originally Indo-Aryan or Dravidian – Indo-Aryan: Sanskrit *agaru-* and *aguru-* ‘agarwood tree (*Aquilaria malaccensis* Lam.)’, Pali *agaru-*, *agalu-*, *agaḷu-*, *akalu-* ‘the fragrant resin or wood of agarwood tree (*Aquilaria malaccensis* Lam.)’ etc.; Dravidian: Tamil *akil* ‘agarwood tree (*Aquilaria malaccensis* Lam.)’; the drug agar obtained from the tree’, Malayalam *akil* ‘agarwood tree (*Aquilaria malaccensis* Lam.)’ etc.; cf. Biblical Hebrew *’āhālōt* (formally pl.) and *’āhālīm* (formally pl.) ‘agarwood (of *Aquilaria malaccensis* Lam.)’ along with Greek ἄλωθ which is a transliteration of *’āhālōt* in the Septuagint, Middle Persian *awalūg* ‘agarwood’ (probably from earlier *\*agalūk*; cf. *w* in Middle Persian *murw* ‘bird’ vs. Young Avestan *mərəya-* ‘id.’), Modern Persian *āluwwa* ‘id.’, Syriac *’wlg*, *’blwg* ‘id.’ (from Iranian). The closest to Greek ἀγάλοχον is the early Middle Persian form *\*agalūk*, but, probably, the Greeks did not adopt this word from Persian (we would expect κ instead of χ). Cf. an unlikely suggestion that ἀγάλοχον comes from Indo-Aryan *\*agalukha-* with the suffix *-kha-* (this view has no confirmation in the available data).

If there are different hypotheses about the etymology of a word, these are presented in such a way that the dictionary user can easily find them out without having to consult the relevant publications. As a rule, the etymologies proposed in the last 150 years are cited, and only in special situations are older theories given. For example, in the case of the Carian gloss γίσσα ‘stone’ six proposals are listed (the hypotheses are so unlikely as to require no further discussion):

Ⓔ A Carian word without etymology. The gloss was juxtaposed with: 1. γείσον n. ‘projecting part of the roof, cornice’; 2. κίσηρις, -εως and -ιδος f. ‘pumice’; 3. Georgian *kviša* ‘sand’, etc.; 4. Basque *gisu*, *khisu* ‘lime’ (note that this word comes from Latin *gypsum* ‘lime’); 5. Akkadian *hiššu(m)* ‘gravel’, Biblical Hebrew *hāšāš* ‘id.’, etc.; 6. Old High German *kisil* ‘pebble’, Old English *ceosel* ‘gravel’, etc.

As regards etymological investigations, methodological approaches developed in the field of comparative-historical linguistics are largely applied. This approach will allow for phonetic and morphological analyses. Of course, instead of sound laws (known e.g. from Indo-European studies), phonetic counterparts between the source and target languages are crucial here.<sup>9</sup> It would be a methodological error to suppose, for example, some irregularities without a determined cause or to postulate a semantic change without sufficient evidence for it.<sup>10</sup>

The last section of an entry (marker: 📖) will deliver bibliographical notes divided into two groups: a) Data – publications (especially dictionaries) that provide linguistic data from Oriental languages; b) References – the most important publications concerning the etymology of a given word. The bibliography (located under a separate tab in the electronic version or at the end of the publication) will be divided into three groups: a) Editions of sources (ancient literary texts, inscriptions, papyri) with abbreviations used throughout the lexicon; b) Book abbreviations (especially dictionaries and reference books); c) Books and articles.

In summary, the *Lexicon of Oriental words in Ancient Greek*, which will contain approximately 1200 entries, aims to be a useful reference work in the fields of etymological, philological, historical and cultural studies. First of all, it seeks to constitute an important addition to the knowledge of Greek etymology in general, as well as the presence of Oriental words in Greek sources. The data collected will also be a significant voice in the discussion on intercultural contacts in antiquity, namely between the West and the East over the course of two thousand years of history. The investigations will provide linguistic evidence for studies concerning the mutual relations and Eastern impact on Greek culture. The lexicon may also be useful for the study of other languages, which have adopted words of Oriental origin, either directly from Greek or via Latin.

<sup>9</sup>Cf. considerations made by Brust 2008: I–LXI. in the introduction to his book on Indo-Iranian vocabulary in Greek.

<sup>10</sup>See Jasanoff & Nussbaum 1996 who levelled criticism at an overly loose approach to phonetic and semantic change of Bernal 1987–2006; cf. Nardelli 2013. For the methodology, see also Hoch 1994: 6–14, Griffith 1997, Ciancaglini 2008: 1–91, as well as, more general, Thomason & Kaufmann 1988.

## Acknowledgments

This research was funded by National Science Centre, Poland (research project no. UMO-2020/39/B/HS2/00934).

## References

- Bai, Gang. 2009. *Semitische Lehnwörter im Altgriechischen* (Schriftenreihe Schriften zur vergleichenden Sprachwissenschaft 4). Hamburg: Kovač.
- Beekes, Robert. 2010. *Etymological dictionary of Greek*. In collab. with Lucien van Beek (Leiden Indo-European etymological dictionary series v. 10/1-2). Leiden: Brill.
- Bernal, Martin. 1987. *Black Athena: The Afroasiatic roots of classical civilisation*, vol. 1-3. New Brunswick: Rutgers University Press.
- Brust, Manfred. 2008. *Die indischen und iranischen Lehnwörter im Griechischen*. 2. verb. und verm. Aufl. (Innsbrucker Beiträge zur Sprachwissenschaft). Innsbruck: Inst. für Sprachen u. Literaturen d. Univ. Innsbruck, Abt. Sprachwiss.
- Chantraine, Pierre. 1999. *Dictionnaire étymologique de la langue grecque: Histoire des mots*. Paris: Klincksieck.
- Ciancaglini, Claudia A. 2008. *Iranian loanwords in Syriac* (Beiträge zur Iranistik). Wiesbaden: Reichert.
- Fournet, Jean-Luc. 1989. Les emprunts du grec à l'égyptien. *Bulletin de la Société de Linguistique de Paris* LXXXIV, 1. 55–80. <https://shs.hal.science/halshs-00001023> (2 April, 2024).
- Frisk, Hjalmar. 1960. *Griechisches etymologisches Wörterbuch*, vol. 1-3. Heidelberg: Winter.
- Griffith, R. Drew. 1997. Criteria for evaluating hypothetical Egyptian loan-words in Greek: The case of αἴγυπτος. *Illinois Classical Studies* 22. 1–6. <https://www.jstor.org/stable/23065323> (2 April, 2024).
- Haas, Otto. 1960. Die phrygische Sprache im Lichte der Glossen und Namen. *Balkansko Ezikoznanie/Linguistique Balkanique* 2. 25–68.
- Hemmerdinger, Bertrand. 1968. Noms communs grecs d'origine égyptienne. *Glotta* 46(3). 238–247. <https://www.jstor.org/stable/40266061> (2 April, 2024).
- Hinz, Walther. 1975. *Altiranisches Sprachgut der Nebenüberlieferungen* (Göttinger Orientforschungen Iranica). Wiesbaden: Harrassowitz.
- Hoch, James E. 1994. *Semitic words in Egyptian texts of the New Kingdom and third intermediate period*. Princeton: Princeton Univ. Press.
- Huysse, Philip. 1990. Persisches Wortgut in Athenaios' "Deipnosophistai". *Glotta* 68(1). 93–104. <https://www.jstor.org/stable/40266852>.

- Jasanoff, Jay H. & Alan Nussbaum. 1996. Word games: The linguistic evidence in Black Athena. In Mary R. Lefkowitz & Guy MacLean Rogers (eds.), *Black Athena revisited*, 177–205. Chapel Hill – London: The University of North Carolina Press.
- Lewy, Heinrich. 1895. *Die semitischen Fremdwörter im Griechischen*. Berlin: Gaertner.
- Masson, Émilie. 1967. *Recherches sur les plus anciens emprunts sémitiques en grec* (Études et commentaires). Paris: Klincksieck.
- McGready, Alfred Glen. 1968. Egyptian words in the Greek vocabulary. *Glotta* 46(3). 247–254. <https://www.jstor.org/stable/40266062> (2 April, 2024).
- Muss-Arnolt, W. 1892. On Semitic words in Greek and Latin. *Transactions of the American Philological Association* 23. 35–156. DOI: 10.2307/2935792. <https://www.jstor.org/stable/2935792> (2 April, 2024).
- Nardelli, Jean-Fabrice. 2013. Black Athena fades away: A consideration of Martin Bernal’s linguistic arguments. *Exemplaria Classica* 17. 281–374. DOI: 10.33776/ec.v17i0.2343.
- Neumann, Günter. 1961. *Untersuchungen zum Weiterleben hethitischen und luwischen Sprachgutes in hellenistischer und römischer Zeit*. Wiesbaden: Harrassowitz.
- Obrador-Cursach, Bartomeu. 2020. *The Phrygian language*. Leiden: Brill.
- Pierce, Richard Holton. 1971. Egyptian loan-words in ancient Greek. *Symbolae Osloenses* 46(1). 96–107. DOI: 10.1080/00397677108590630. (2 April, 2024).
- Rosół, Rafał. 2013. *Frühe semitische Lehnwörter im Griechischen*. Frankfurt am Main: Peter Lang.
- Rosół, Rafał. 2018. Early Semitic loanwords in Greek. In Łukasz Niesiołowski-Spanò & Marek Węcowski (eds.), *Change, continuity, and connectivity. North-Eastern Mediterranean at the turn of the Bronze Age and in the early Iron Age*, 334–344. Wiesbaden: Harrassowitz.
- Schmitt, Rüdiger. 1967. Medisches und persisches Sprachgut bei Herodot. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 117(1). 119–145. <https://www.jstor.org/stable/43372247>.
- Simon, Zsolt. 2018. Anatolian influences on Greek. In Łukasz Niesiołowski-Spanò & Marek Węcowski (eds.), *Change, continuity, and connectivity: North-Eastern Mediterranean at the turn of the Bronze Age and in the early Iron Age*, 376–418. Wiesbaden: Harrassowitz.
- Simotas [Σιμωντάς], Panagiotis N. [Παναγιώτης Ν.] 1968. *Αι αμετάφραστοι λέξεις εν τω κειμένω των ο’* (Επιστημονική επετηρίς Θεολογικής σχολής). Θεσσαλονίκη: Αριστοτέλειον Πανεπιστήμιον Θεσσαλονίκης.
- Sowa, Wojciech. 2008. *Studien zum Phrygischen*. 1st edn. Göttingen: Cuvillier.

- Thomason, Sarah Grey & Terrence Kaufmann. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: Univ. of California Press.
- Tovar, Sofía Torallas. 2004. Egyptian lexical interference in the Greek of Byzantine and early Islamic Egypt. In Petra Sijpesteijn & Lennart Sundelin (eds.), *Papyrology and the history of early Islamic Egypt*, 163–198. Leiden & Boston: Brill. DOI: 10.1163/9789047405474\_015.



# Chapter 21

## The discovery of a Russian-Tajik dictionary

Abdusalom Mamadnazarov

Tajik National University

Bahriddin Navruzshoev

Tajik National University

The purpose of this article is to analyze the lexicographic and lexical peculiarities of the Russian-Tajik Dictionary (RTD) by Sh. Khatymtayev, published in 1899, and recently discovered in the library of the Tajik National Academy of Sciences. The lexicographic and lexical peculiarities of this publication remain almost unknown and unexplored up to now. The present analysis of this dictionary is being carried out based on the single copy, stored in the Academy library and numbered 337973A. Largely unknown by linguists, this Russian-Tajik Dictionary can be considered as the first thematic Russian-Tajik dictionary ever published.

The dictionary consists of 72 pages, has no original cover, the binding in which it is currently discovered appeared much later. The absence of the flyleaf, as well as the last pages, seriously complicates the work of identifying additional data about this publication. The RTD is a thematic dictionary, contains 1370 entries divided into 49 topics.

The dictionary suffers from a serious deficiency: it fails to arrange either the topics of the dictionary, or the words – entries inside the topics in alphabetical order.

### 1 Introduction

The purpose of this article is to analyze the lexicographic and lexical peculiarities of the recently discovered *Russian-Tajik Dictionary* (Khatymtayev 1899) by



Sh. Khatymtayev, published in 1899, which builds on previous work by (Mamadnazarov 2013, 2016, 2021). The lexicographic and lexical peculiarities of this publication have remained largely unknown and unexplored until now. The present analysis of this dictionary is based on the single copy stored in the library of the National Academy of Sciences of Tajikistan, numbered 337973A. The author, Khatymtayev, is to date poorly documented, and work is underway to compile a biography and locate copies of the text extant in Russia.

Largely unknown by linguists, this *Russian-Tajik Dictionary* can be considered as the second Russian-Tajik dictionary ever published after V.F. Nalivkin's *Russian-Persian Dictionary of the Dialects of the Turkestan Region*, published some ten years earlier. It is, however, the first *thematic* Russian-Tajik dictionary ever published. It was compiled by Shohhaidar Khatymtayev and published at the Y. Baranovskiy Printing House in New Bukhara (Central Asia) and approved by the Russian Imperial Political Agency in Bukhara. The *Russian-Tajik Dictionary* (henceforth RTD) was first published in 1899 in the city of New Bukhara (Kagan), and then in 1913 in Tashkent under the title *Brief Russian-Tajik Dictionary*.

The analysis of this dictionary is being carried out using the only surviving copy, a valuable lexicographic publication of its time, that was preserved in the personal book collection of Professor M. Andreev, acquired by the Library of the Academy of Sciences of the Republic of Tajikistan in 1967. M. Andreev is widely known for his scientific publications on the history, culture and ethnography of Tajiks. Throughout his creative life, M. Andreev collected books and historical sources, forming a unique book collection dedicated to the history and culture of the Central Asian region. The literature review indicates that this *Russian-Tajik Dictionary* has not been described or studied by linguists and lexicographers in the Republic of Tajikistan or beyond.

## 2 Literature review

The introduction to the first *Russian-Tajik Dictionary*, published in two volumes in the Soviet period in 1933–1934 in Leningrad, and compiled by S. Alizade, A. Ismailzade, R. Hashim and M. Yusupov, gives no hint about even the existence of Khatymtayev's dictionary (Alizade et al. 1933: 5–6).

In the introduction to the second *Russian-Tajik Dictionary*, edited by A.P. Dehoti and N.N. Yershov, we again do not observe any information about this work (Dehoti & Yershov 1949: 5–6).

The 1985 *Russian-Tajik Dictionary*, edited by M. Asimov, contains only one sentence in the introduction: "In 1899 and in 1913, two short *Russian-Tajik* dictio-

naries were published in Tashkent (compiled by Shohhaidar Hotamtoev)” (Asimov 1985: 7). As we see, the place of publication of the RTD is incorrect in this dictionary.

A small article consisting of only eight sentences devoted to this dictionary was published in the *Tajik Soviet Encyclopedia* by A. Sayfullaev in Tajik, which provides general information about the dictionary by Hotamtoev. In particular, it is noted that this dictionary contains 1370 words and is of great importance as a lexicographic, dialectological, historical and language publication. The words are arranged by topics, such as Pronouns, Verbs, Creation in General, Airy Phenomena, etc. (Sayfullaev 1983: 72).

To date, that is all the information about this dictionary in linguistic literature in Tajik and Russian. As this survey of the linguistic literature shows, this *Russian-Tajik Dictionary* has not been described and studied by linguists and lexicographers in the Republic of Tajikistan or beyond.

It is impossible to determine when and how this dictionary entered Andreev’s collection, but its presence in the library is of great value for the National Academy of Sciences library, as the dictionary is one of the first Russian-Tajik dictionaries in the history of Russian-Tajik bilingual lexicography.

The dictionary consists of 72 pages, has no original cover, the binding it currently has having been added much later. The absence of the flyleaf, as well as the final pages, seriously complicates the work of identifying additional data about this publication; for example, we cannot estimate the number of published copies of the book. According to the author, there is a word list at the end of the dictionary, but this copy lacks it, and we cannot say in what order this list was compiled.

The Tajik scholars A. Mukhtarov and A. Sayfullaev in their article published in a newspaper entitled – *This was the first dictionary*, dated July 9, 1981 noted “the uniqueness and significance” (Mukhtarov A 1981: 4) of this dictionary, unknown not only to the public, but also to lexicographers. At the same time, it is interesting that they wrote about a copy of the book found in the Russian State Library in Moscow. These scholars clearly did not know that a copy of this publication existed in the Library of the Academy of Sciences of Tajikistan.

The RTD is a thematic dictionary containing 1370 entries divided into 49 topics. From the point of view of macrostructure [Hartman, 2002, 91], the dictionary consists of a *Preface* (in Russian and Tajik), the thematic dictionary itself and the table of contents at the end of the dictionary. The sub-title reads: “A guide to how Russians can learn to speak Tajik and Tajiks can learn to speak Russian without help” (Khatymtayev 1899: 1).

It is interesting to note that the same page in Tajik, in Arabic script, calls this the *Russian-Persian Dictionary* and then states that: “According to the content of this book, a Russian-speaking teacher can speak *Persian*, and *Muslims* can speak Russian” (Khatymtayev 1899: 1). Further, we find below in Uzbek: “Basmakhona Baranovskiy Yangi Bukhoroda”<sup>1</sup>. We have no idea why this sentence is in Uzbek.

The introduction in Russian, which consists of only five sentences, notes that, “For the convenience of finding the right word, the words are divided into groups according to categories of appropriate meaning. The category titles are placed at the end of the dictionary” (Khatymtayev 1899: 3). In addition, on the same page, a small text appears in Tajik under the headline: “From the author”, the content of which differs from the Russian introduction. In particular, it is claimed that: “With this dictionary *Persian* speaking people can quickly and easily learn the Russian language, because Russian and *Muslim* words and expressions are written in Russian and *Muslim* script and anyone who has set a goal can find topics in the table of contents and find them in the dictionary” (Khatymtayev 1899: 3).

Comparing the Russian original of the following: “With this dictionary, Russians can easily get acquainted with the *Tajik* language, and *Tajiks* with Russian, not knowing the script of these languages, because Russian words are written in Russian and *Tajik* letters and vice versa” (Khatymtayev 1899: 3) with the text in Tajik (above), it can be seen that the word *Tajiks* is replaced by *Forisyon* (‘Persians’), and *Tajik language* is translated as ‘Muslim language’. We may thus conclude that, on the title page and the original Tajik text in the “Introduction”, the author avoided using the words “*Tajiks*” and “*Tajik language*” several times and replaced them with “*Persians*” and “*Muslim Language*”. For what reason we do not know. We can only presume that at that period the authors did not distinguish between the Persian and Tajik languages. Tajik and Persian were used interchangeably.

This work can be considered a bilingual thematic translation dictionary, since it is divided into 49 topics. Some topics are divided into several sub-topics, for example: “Man and parts of the human body” [p. 32], “Gender, Age and Family” [p. 41], “Tools, Weapons and Guns” [p. 57]. Taking this division into account, the total number of topics in the RTD increases to 70.

This figure is not accurate, however, since this copy consists of only 72 pages, with some pages missing. Thus, for example, whilst A. Sayfullaev talks about the topic “Phrases for exercise” saying that: “This section consists of 29 sentences with translation into Tajik” (Sayfullaev 1983: 4), this topic is missing from our

---

<sup>1</sup>trans. Printed in Baranovskiy’s Printing House

copy. It follows that in all probability the number of topics and pages of the dictionary were greater than we are aware of.

The entries are arranged in two columns on various topics including: Pronouns, Verbs, Conjunctions, Adverbs and Prepositions, Ordinal Numbers, Spiritual Words, Creation In General, Airy Phenomena, Numbers, etc.

The dictionary suffers from serious deficiencies in that it fails to arrange either the topics of the dictionary or the words within the topics in alphabetical order or in any other order or system of the Russian language. This makes it difficult to use when searching for a particular word. In this regard, is it not possible to call this publication a dictionary at all as, like all linguists, by a dictionary we understand “A book containing a systematic description of the vocabulary of the language” Akhmanova 2009, 421. When 10, 20 or even 30 words are given on a particular topic in a book, finding the right word does not seem to be a problem, but when 321 words are randomly given on a topic (for example, “Verbs” in this dictionary), finding the right word becomes highly problematic. Thus, this publication should more accurately be called a thematic ‘word list’, and not ‘a dictionary’.

The description of Russian words and their translations in the dictionary begins with a capital letter without a paragraph in two ways: in both Cyrillic and in the Arabic script of the Tajik language. For the convenience of readers of this article, instead of the Arabic script, we have used a transliteration in Roman script. For example:

- (1) (RTD, 1899, 6)
  - a. Быть. Шуданъ  
Вить. Shudanъ  
‘to be’
  - b. Родиться. Таваллудъ-ёфтанъ.  
Roditsya. Tavallud yoftan.  
‘to be born’
  - c. Шевелить. Джумониданъ.  
Shevelit Jumonidan  
‘to shake’
  - d. Дышать. Дамъ-гирифтанъ.  
Dishat. Dam giriftan.  
‘to breathe’
  - e. Рассматривать. Дида-гаштанъ.  
Rassmatrivat, dida gashtan.  
‘to consider’

As these examples show, Russian words are not arranged in alphabetical order; they begin with a capital letter and end with a full stop. Transliteration into Russian and translation into Tajik was carried out in Cyrillic and the Arabic script, which we replaced with Roman script.

A common RTD entry consists of a Russian entry with a capital letter without a paragraph and a Tajik equivalent or equivalents arranged in two columns, without any grammatical and stylistic content, special notes or illustrative examples.

The entire dictionary is divided into forty-six topics. The translation of Russian words into Tajik is given in Cyrillic and the Arabic script, so that those who do not know the Tajik script can use the dictionary. It consists of a variety of topics, but the size and content of the topics differ significantly from each other. The largest topic, for example, “Verbs”, is represented by 321 entries, “Man and parts of the human body” is represented by 90, and “Conjunctions, adverbs and prepositions” by 74 words and expressions. The remaining topics have from 4 to 12 entries and from 30 to 45 words and expressions.

The first topic, “Pronouns” (Zamirho), consists of 32 entries (p. 5–6). This includes personal pronouns, as well as interrogative, negative and other types of pronoun and their translation into Tajik, but the pronouns themselves are not divided into types. For example:

- (2) a. Kto. Ki.  
‘who?’
- b. Chto. Chi.  
‘what?’
- c. Kotoriy. Kadom.  
‘which?’
- d. Chey etot. In kist.  
‘Whose is this?’
- e. Nekto. Shakhse.  
‘somebody’

Lexical analysis of the examples shows that there are also some incorrect Russian Tajik translations. For example, the Russian phrase *Chey etot* ‘whose is this’ is translated as *in kist* (‘who is it’), and the correct translation in Tajik is *In az kist*. Some words are translated into colloquial Tajik variants, for example:

- (3) a. Moy etotъ. Ini-manъ.  
‘this is mine’

- b. Tvoy etotъ. Ini-ty.  
'this is yours'
- c. Ego etotъ. Ini-vay.  
'this is his'
- d. Nash etotъ. Ini-mo.  
'this is ours'
- e. Vash etotъ. Ini-shumo.  
'This is yours'

The fact that the dictionary also contains colloquial words suggests that the dictionary was intended to teach the spoken language as well, as indicated in the preface.

The lexical units of the largest topic in this dictionary, namely "Verbs", as well as words and expressions in other topics, are not arranged in the Russian alphabetical order, which makes it very difficult to find the right word. The peculiarity of this section is that the verbs that have different forms in the imperative mood are also listed in the dictionary as independent lexicographic entries or lexicographic articles. For example:

- (4) (RTD, p. 7)
  - a. Begatъ. Davidan.  
'to run'
  - b. Bdzhite. Bezhite.  
'run'
  - c. Vstatъ. Barkhestanъ.  
'to stand up'
  - d. Vstavayte Khizidъ.  
'stand up'
  - e. Slishat. Gush kardanъ.  
'to listen'
  - f. Slushayte. Gush Kunidъ.  
'listen'

The dictionary also contains phrases, expressions and sentences in the imperative mood:

- (5) a. Lozhites spatъ. Khobъ ravidъ.  
'Go to bed' [p. 9].  
b. Imya dayte. Nomъ monidъ.  
'Name it' [p. 13].  
c. Zaprygayte loshadey. Asbhorо bandidъ.  
'Harness the horses.' [p. 14].  
d. Vipraygayte loshadey. Asbhorо yala kunidъ.  
'Unharness the horses.' [p. 14].

There are also examples in which Russian words are translated into Uzbek with the Tajik auxiliary verb *mondan*.

- (6) a. Plavat. Suzmishъ-kardanъ. *Taj.* shino kardan  
'to swim' [p. 15].  
b. Prosverlit. Tishikъ-kardanъ. *Taj.* surokh kardan  
'to drill' [p. 16].  
c. Kleit. Elimъ-kardanъ. *Taj.* shiresh kardan  
'to glue' [p. 16].

We also find that Russian verbs are translated into Tajik incorrectly (we have given the correct English translation in parentheses). For example, the phrase in the imperative mood *Gum shaved!* has a rough sense corresponding to the English *Get out!*

- (7) a. Bit. Shudan.  
'to become (to be)' [p. 16].  
b. Meritъ. Raymudanъ.  
'to understand (to measure)' [p. 16].  
c. Meryayte. Barkashidъ.  
'measure (weigh)' [p. 16].  
d. Spryachtes. Gum shavidъ.  
'get out (hide)' [p. 16].

There are also verbs that are translated correctly into Tajik (with the Arabic script) (the third column), but words in Cyrillic are not accurately translated into Tajik, with errors and with a hard sign at the end of each word (the second column). For example:

- (8) a. Bolit golova. Saramъ-dardъ-mikinatъ.  
 ‘My head aches.’ [c. 17].  
 b. Boznositshe. Migrozh raftanъ.  
 ‘to ascend’ [c. 18].  
 c. Vospreshat. Mangъ kardanъ.  
 ‘to forbid’ [c. 18].  
 d. Naznachat. Taganeъ kardanъ.  
 ‘to appoint’ [c. 19].  
 e. Khudet. Loghar-Gardidanъ.  
 ‘to lose weight’ [c. 21].

In these words, the translation into Tajik with the Arabic script is mostly accurate without any problems, but the translation into Tajik in Cyrillic has errors and inaccuracies, and each word ends with a solid sign (ъ). It seems to us that it was problematic for Russians to understand and pronounce these words correctly.

The second largest topic in this dictionary is *Man and parts of the human body*, represented by 90 words and expressions. Some examples from p. 33 with our English translations are:

- (9) a. Lobъ. Pishona.  
 ‘forehead’  
 b. Glazъ. Chashmъ.  
 ‘eye’  
 c. Brovъ. Abруъ.  
 ‘eyebrow’  
 d. Zrachekъ. Siyohii Chashm.  
 ‘eyeball’

From the examples given, we see that the words of the Russian language were mostly written with a final hard sign ending (ъ), which was preserved in translation and spelling in the Tajik language in Cyrillic. Spelling with a final hard sign (ъ) was used in Russian until the beginning of the 20th century. For example, a famous Russian lexicographer V.I. Dal used the same spelling when compiling his *Explanatory Dictionary of the Live Great Russian Language* in the 19th century in four volumes (Dal 1966). The reason why Sh. Khatymtayev used the hard sign in the RTD is, in our opinion, that the translation into Tajik with Cyrillic

spelling was intended for Russians who could use Tajik without knowing the Arabic script. Apparently, the compiler of the dictionary did not know that in Tajik the sounds do not have a category of hardness and softness. However, in translations into Tajik in the same words with the Arabic script, such a phenomenon is not observed, that is, the firm sign (Ъ) was not written.

The analysis of the dictionary entries and their translations into Tajik shows that even in Russian words ending in a soft sign (ь) for some reason a hard sign (Ъ) was used instead in the Tajik translations. For example:

- (10) a. OchenЪ. BisyorЪ.  
‘very’ [p. 29].
- b. ZhiznЪ. UmurЪ.  
‘life’ [p. 29].
- c. SmertЪ. MargЪ.  
‘death’ [p. 29].
- d. OgonЪ. OtashЪ.  
‘fire’ [p. 30].

There are also several words in the dictionary where the soft sign (ь) was preserved in translations into Tajik in Cyrillic. For example:

- (11) a. StepЪ. ChulЪ.  
‘desert’ [p. 30].
- b. BoleznЪ. KasalЪ.  
‘illness’ [p. 39].
- c. InbirЪ. ZandzhabilЪ.  
‘ginger’ [p. 54].

The dictionary also contains the words where there is no hard sign in Russian, but the hard sign has been used in translations into Tajik in Cyrillic. For example:

- (12) a. Boroda. RishЪ.  
‘beard’ [p. 33].
- b. Sheya. GardanЪ.  
‘neck’ [p. 33].
- c. Gorlo. HalqЪ.  
‘throat’ [p. 33].

- d. Plecho. Dushъ.  
'shoulder' [p. 33].

In the section entitled "Flowers" (*gulho*), on page 51, there are only six words, of which two could hardly be related to the topic of flowers: For example:

- (13) a. Podsolnechnikъ. Oftobъ-farastъ.  
'sunflower'  
b. Opiumъ. Afyunъ.  
'opium'

Seventy-five lexical units represent the third biggest topic described in the RTD – *Conjunctions, Adverbs and Prepositions*. As in the previous sections, lexical units are arranged in an irregular manner, not observing the alphabetical order of the Russian language. The section *Conjunctions, Adverbs and Prepositions* is not divided into subgroups. For example:

- (14) a. Esli. Agarъ.  
'if' [p. 21].  
b. Khotya. Agerchi.  
'although' [p. 21].  
c. Dazhe. Hatto.  
'even' [p. 21].  
d. Okolo. Nazdi.  
'by' [p. 22].  
e. Vnezapno. Nogohъ.  
'suddenly' [p. 22].

From the analysis of these words and others in the dictionary, it is clear that *Conjunctions, Adverbs and Prepositions* are given in the dictionary without any system or order, without indicating that they belong to one or another part of speech. Here, among other words, we can find *Interrogative Pronouns* that are not related to this topic. These and some other *Interrogative Pronouns* should have been described in the first topic – *Pronouns*. For example:

- (15) a. Kak. Ch-tariqa.  
'how' [p.22]

- b. Skolko. Chandъ.  
'how much' [p. 21].
- c. Gde. Gdzho.  
'where' [p. 23].
- d. Otkuda. Azgdzho.  
'where from' [p. 23].
- e. Kuda. Bakujo.  
'where to' [p. 23].

Native Tajik letters and sounds, such as: [q], [ʁ], [h], [ə:], [i:], [dʒ] in the translations of Russian words into Tajik, in this dictionary in Cyrillic, are not represented in any way, although in words with the Arabic script of the Tajik language they are mostly correctly transliterated. This indicates that Russians using this dictionary could not accurately and correctly read or use words and expressions from the RTD in their speech. Since it is not known what the circulation of the book was or how many copies had been printed, we cannot say how often the dictionary was used, or whether it was used at all.

- (16) a. Vmeste. Khamro.  
'together' [p. 22].
- b. Otkuda. Azgadzho.  
'where from' [p. 23].
- c. Pozadi. Darъ-Aqibъ.  
'in the back' [p. 23].
- d. Togda. Hamon-baqtъ.  
'at that time' [p.24].

We also observed that in the dictionary the translations of some compound words, phrases and expressions into Tajik in Cyrillic were written with a hyphen and a capital letter, which is not observed in translations into Tajik with Arabic script. For example:

- (17) a. Potomu. Azъ-onъ-sababъ.  
'For that reason' [p. 22].
- b. Vnizu. Darъ taki-poinъ.  
'downstairs' [p. 23].

- c. Nigde. Darъ-khichъ-dzhoъ.  
‘nowhere’ [p. 23].
- d. Togda. Hamonъ-vaqtъ.  
‘at that time’ [p. 24].
- e. Dnyomъ. Vaqtiъ-ruzъ.  
‘time of the day’ [p. 24].
- f. Oreshnikъ. Chahor-mazhzъ.  
‘nut’ [p. 52].

The fourth largest topic in Khatymtayev’s idictionary is called *Arts and Crafts*, which contains 46 Russian language words and expressions with translation into Tajik in two versions – in Cyrillic and the Arabic script (pp. 59–60). As in other sections, lexical analysis of translations shows that the main problem is translation into Tajik with Cyrillic letters. For example:

- (18) a. Kupetsъ. Tuzharъ. Tojir.  
‘trader’
- b. Torgoviy-domъ. Tizharъ-khane. tojirkhona  
‘trade house’
- c. Torgovlya bratevъ. Savdo-barodaronъ. savdoi barodaron  
‘brothers trade’
- d. Torgovlya. Migamila. muomila  
‘– trade’
- e. Kamenotyosъ. Sanъ-tarashъ. sangtarosh  
‘mason’
- f. Shveya. Zani-darizduzъ. zani darzduz  
‘tailor’

As was noted previously, the translation into Tajik written in Cyrillic in this section does not correspond in most cases to the translation represented in the Arabic script. In the words above, for example, *mgamilya* is very far from the word *muomila* in Tajik, in both writing and sound. These considerations explain why we doubt that Russians could benefit from this dictionary effectively. There are also mistakes in the words *tizharkhane* and *savdo-barodaron* and other words.

However, the originality of the translation of some words is of interest for modern Tajik, such as: *bankrotstovo* – *bankruptcy* – *tazhar-shikastagi*, *skvorknyak*

– *furrier* – *pustin-soz*, *drevosek* – *woodcutter* – *ezumbur*, *tokar* – *turner* – *dukchi*, *raznoschik* – *peddler* – *charchinfurysh*, *cyryulnik* – *sartarosh* – barber and other words.

One of the sections located at the end of the dictionary is called *Time in General*, where 20 Russian words, expressions and 4 sentences are translated into Tajik on page 72. The translations of most Russian words and phrases are conveyed adequately and correctly. For example:

- (19) a. День. Ruzъ.  
‘day’  
b. Nochъ. Shabъ.  
‘night’  
c. Minuta. Daqia.  
‘Minute’  
d. Zima. Zimiston.  
‘Winter’  
e. Sneg idyotъ. Barf-borida-istoddst.  
‘It is raining.’  
f. Segodnya ochenъ zharko. Imruzъ-bisyoгъ-garmъ.  
‘It is very hot today.’

However, there are also inaccurate and inadequate translations or errors. For example, the word *Vedro* ‘bucket’, which apparently got into this topic by mistake, is translated as *ruzi khubъ* ‘good day’.

- (20) a. Vedro. Ruzi khubъ.  
‘Good day.’  
b. Nepogoda. Ruzi-badъ.  
‘bad day’ (should be: ‘bad weather’)  
c. Dozhdъ. Buranъ-borida-istodaastъ.  
‘It is raining’ (should be: ‘rain’)

The dictionary suffers from another serious deficiency, failing to arrange either the topics of the dictionary or the words – entries within the topics in alphabetical order or in any recognisable Russian language system. This makes it difficult to use and to find the right word.

In Khatymtayev's dictionary, the scope and content of the topics, in terms of spelling and semantics, differ significantly from each other. The biggest topic, for example, *Verbs*, is represented by 321 words and collocations. The topic *Conjunctions, adverbs and prepositions* is represented by only 74 words and expressions. The remaining topics have fewer, from six to 12 and from 30 to 35 words and expressions.

An analysis of the translation into equivalent Russian words shows that the translations into Tajik in the Arabic alphabet is carried out fairly accurately without much difficulty, but the translations into Tajik in Cyrillic suffer from errors. Russians faced great difficulty understanding and pronouncing the Tajik words and collocations correctly with the spelling provided in Cyrillic.

Analysis of the examples reveals that Russian language words were mostly written with a final hard sign (Ѣ). Russian spelling used this sign (Ѣ) at the end of the words until the beginning of the 20th century, something which this dictionary preserved in translations and spellings in Tajik in Cyrillic. Khatymtayev probably used the hard sign (Ѣ) in the Tajik translations because the translation into Tajik with Cyrillic script was intended for Russians who could use Tajik without knowing the Arabic script. Apparently, Khatymtayev did not know that the Tajik language lexical units do not have the category of *hardness and softness*. Analysis of translations into Tajik also revealed that the author translates most of the Russian language words and expressions correctly. However, he also makes many mistakes and, in some cases, he uses translations into Uzbek.

The analysis of the spelling of the RTD shows that it fails to represent or convey important native Tajik sounds, such as: [q], [ʁ], [h], [ə:], [i:], [dʒ] in the translations into Tajik in this dictionary in Cyrillic. However, it renders them mostly correctly in words with the Tajik/Persian script. This fact indicates that Russians using this dictionary could not accurately and correctly read or use words from this dictionary in their speech.

### 3 Conclusion

Lexicographic and lexical analysis of the Khatymtayev's RTD identified the features and peculiarities of this dictionary both from the point of view of translation and transliteration. As a result, it can be stated that this work, despite its lexicographic shortcomings from the point of view of modern principles of lexicography, is an important late 19th century lexicographic and lexical source of these languages. It reflects the vocabulary and spelling of the Russian and Tajik languages of its time and is of interest as a historical, lexical and dialectological

monument of these languages in general and Bukhara region colloquial speech in particular. It seems to us that the compiler of the dictionary was neither a professional lexicographer nor a linguist. We cannot judge how widely the dictionary could have been used, since its circulation is not known. In conclusion, it should be emphasized that a thorough and detailed study of this source is necessary, since this cannot be done within the framework of an article.

## References

- Alizade, S., A. Ismailzade, R. Hashim & M. Yusupov (eds.). 1933. *Russian-Tajik dictionary*, 2 vols. Leningrad: Comintern.
- Asimov, M. (ed.). 1985. *Russian-Tajik dictionary*. Moscow: Russian Language.
- Dal, Vladimir Ivanovich. 1966. *Explanatory dictionary of the living great Russian language*, 4 vols. Moscow: Russian Language Media.
- Dehoti, A. P. & N. N. Yershov (eds.). 1949. *Russian-Tajik dictionary*. Stalinabad: State Publisher of Foreign & National Dictionaries.
- Khatymtayev, S. H. (ed.). 1899. *Russian-Tajik dictionary*. New Bukhara: Baranovskiy.
- Mamadnazarov, Abdusalom. 2013. *Comparative analysis of the main types of dictionaries in the Tajik lexicography*. Dushanbe: R-Graph.
- Mamadnazarov, Abdusalom. 2016. *Formation and development of the Tajik translation lexicography of the XX and early XXI centuries*. Revised and expanded edition. Dushanbe: R-Graph.
- Mamadnazarov, Abdusalom. 2021. A select history of modern Russian Tajik special lexicography. In Hans Van de Velde & Fredric T. Dolezal (eds.), *Broadening perspectives in the history of dictionaries and word studies*, 302–316. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Mukhtorov A, Sayfullaev A. 1981. This is how the first dictionary looked like. *Communist Tajikistan*. 4.
- Sayfullaev, A. (ed.). 1983. *Russian-Tajik dictionary*. Dushanbe: Tajik Soviet Encyclopedia.

# Name index

- Abel-Rémusat, Jean-Pierre, 311  
Achinger, Krzysztof, 275, 292, 293  
Ackermann, Tanja, 142  
Adamska-Sałaciak, Arleta, 333  
Ahačič, Kozma, 144, 147, 155, 156, 160  
Aksan, Doğan, 247  
Aksoy, Ömer Asım, 249  
Alizade, S., 428  
Allen, Colin, 35  
Alston, R. C., 60  
Anon, 337  
Antonowicz, Julian, 323, 337  
Aracı, Emre, 246, 254, 261  
Arcodia, Giorgio, 298  
Arnaud, Pierre., 200  
Arnauld, Antoine, 80, 81  
Ashdowne, Richard, 45  
Asimov, M., 429  
Aspaas, Per Pippin, 351, 354  
Autelli, Erica, 400
- Babos, Kálmán, 367, 370, 377, 379, 381  
Bai, Gang, 419  
Bailey, Nathan, 328  
Baker, Mona, 342  
Baker, Thomas M, 106  
Baraga, Frederic, 103, 104  
Barberi, Giuseppe Filippo, 402  
Bárdosi, Vilmos, 363, 372, 390, 396, 400  
Basciano, Bianca, 298  
Bately, Janet, 60
- Beekes, Robert, 419  
Behar, Cem, 246, 263  
Béjoint, Henri, 101, 223  
Belleflamme, Clémence, 375  
Bergantini, Gian Pietro, 404  
Berglund, Lisa, 325  
Bernal, Martin, 422  
Bianchi, Thomas Xavier, 248, 250  
Biffi, Marco, 217  
Bilić, Josip, 174  
Bloomfield, Leonard, 264  
Bocast, Alexander K., 79, 82, 84, 92  
Boisrobert, André, 204  
Bonavilla, Aquilino, 404  
Bottéro, Françoise, 298  
Boutard, Jean-Baptiste-Bon, 213, 214, 216, 218–221, 223, 225, 226  
Boyer, Abel, 321, 328  
Brandes, George, 323  
Bray, Laurent, 203, 368  
Brewer, Charlotte, 12, 23, 24  
Brlek, Mijo Ivan, 185–188  
Brockey, Liam M., 299  
Brust, Manfred, 419, 422  
Buczek, Katarzyna, 324  
Budin, Gerhard, 248  
Buffier, Claude, 80, 81  
Bulut, Mustafa Hilmi, 245  
Bussotti, Michela, 313, 315  
Butterwick, Richard, 324
- Calderino, Cesare, 304, 306, 308

*Name index*

- Calepino, Ambrogio, 305, 306  
Casanova, Emili, 142  
Castorina, Miriam, 303, 315  
Catach, Nina, 196  
Catto, Michela, 300, 301, 309  
Cesar-Nedzbala, Polonca, 107  
Çetin, Kamile, 245  
Cetro, Rosa, 6, 212, 213, 223  
Chalmers, John, 274  
Chambers, Ephraim, 80, 84, 92  
Chantraine, Pierre, 419  
Chartier, Roger, 213, 214, 221  
Chen Zhili (陈至立), 275, 292, 293  
Cheung Kam-Siu, 291  
Ciancaglini, Claudia A., 422  
Cieśla, Michał, 323, 324  
Civezza, Marcellino, 311, 312  
Coblin, Weldon South, 299  
Coleridge, Herbert, 23  
Coles, Elisha, 330  
Collignon, Lucien, 164, 165  
Collinot, André, 222  
Condillac, Étienne Bonnot de, 79, 80,  
93  
Cooper, H. R., 144  
Cormier, Monique C., 328, 343  
Corneille, Thomas, 204  
Cortelazzo, Manlio, 404  
Coseriu, Eugenio, 203  
Coste, Alain, 78, 82, 89  
Cowie, A. P., 60  
Crook, Stephen, 17  
  
d'Alembert, Jean Le Rond, 78, 81, 82,  
92  
Dąbkowski, Grzegorz, 247, 249, 262  
Dahl, Ove Christian, 359  
Dal, Vladimir Ivanovich, 435  
Dalzell, Tom, 34  
  
Danet, Pierre, 331, 335, 337  
Danielson, Virginia, 263  
Danilewicz-Zielińska, Maria, 324  
de Colonia, Dominique, 84  
De Guignes, Joseph, 313  
de La Touche, Pierre, 204  
Deanović, Mirko, 184, 185  
Dehoti, A. P., 428  
Di Fiori, Giacomo, 314  
Diderot, Denis, 82, 92  
Dilçin, Dehri, 249  
Djäv, Ulrika, 356  
Dolci, Sebastian, 185  
Dolezal, Fredric T., 343  
Doroszewski, Witold, 249, 264  
Drnovšek, Marjan, 102, 105, 121  
Dubois, Jean, 163, 164, 168, 188  
Dubois, Jérémie, 402  
Durkin, Philip, 44, 52  
Durman, George, 101  
  
Eberlein, Christian, 335, 337  
Eckhardt, Sándor, 363, 371, 372, 377,  
387, 389, 396  
Elligers, Anne, 378  
Elman, Benjamin, 299  
Estienne, Robert, 2  
Eusebi, Cristina, 217  
  
Farina, Donna M. T. Cr., 101  
Farmer, Henry George, 245, 250, 256  
Feldman, Walter, 246, 263  
Félibien, André, 213, 214, 218–225  
Fennis, Jan, 201  
Fernandez, Heberto, 328  
Ferrie, Joseph P., 104  
Fontenelle, Thierry, 101  
Ford, Thomas, 19

- Formey, Jean-Henri-Samuel, 78, 86,  
87, 90, 91
- Fournet, Jean-Luc, 419
- Frączek, Agnieszka, 333
- Franceschi, Maurizio, 312
- Franchi, Saverio, 312
- Francoeur, Aline, 343
- Franić, Ivana, 163, 170, 174, 178–181,  
184, 186
- Franzen, Christine, 60
- Frey, Claude, 165
- Frisk, Hjalmar, 419
- Furetiere, Antoine, 193, 211, 222, 224,  
230
- Furetière, Antoine, 194, 205, 207
- Furno, Martine, 167
- Gabrić Bagarić, Darija, 165, 170, 172,  
173, 177
- Galinski, Christian, 248
- Galisson, Robert, 165
- Gerber, David A., 104, 106
- Gergonne, Joseph Diaz, 78
- Gilliver, Peter, 12
- Glatigny, Michel, 164, 165
- Gobetz, Edward, 99, 100, 102, 104–  
106
- Gordon, Octavian, 142
- Grafenauer, Ivan, 107
- Graham, John, 19
- Graveleau, Sara, 201, 236
- Greve, Martin, 246, 263
- Griffith, R. Drew, 422
- Grosse, Annelie, 85
- Haas, Otto, 419
- Halton, Timothy J., 104
- Hammer, Eric M., 35
- Harras, Gisèle, 248
- Hartmann, R. R. K., 333
- Hausman, Franz Josef, 249
- Hausssner, Friedrich W., 339
- Hemmerdinger, Bertrand, 419
- Hermans, Theo, 142, 143
- Hernández Jaramillo, Alejandro, 108
- Hinz, Walther, 419
- Hoch, James E., 422
- Horvat, Vladimir, 170, 172, 173
- Hoyt, Joanna Micham, 106
- Huang Chu-Ren, 290, 291, 293
- Hubac, Jean, 85
- Hüllen, Werner, 60
- Hunt, Tony, 44, 55
- Huxley, Thomas, 16
- Huyse, Philip, 419
- Ingham, Richard, 43, 49
- Jacobsen, Tove, 378
- Jajdelski, Wojciech, 322
- Jakopin, Franc, 145, 148, 150
- Jakubczyk, Marcin, 333
- James, Gregory, 333
- Jasanoff, Jay H., 422
- Javor Briški, Marija, 156
- Jezerniczky, Margit, 373
- Jefferson, Judith, 43
- Jelovšek, Alenka, 155
- Jeram, Peter Jos, 107, 109–114, 122,  
136
- Kadioğlu, İdris, 245
- Kafker, Frank A., 83, 85, 94
- Kafker, Serena L., 83, 85
- Kalender, Ruhi, 245, 263
- Kaliszewski, Adam Piotr, 335
- Karpinska, Laura, 375
- Kasimi, Ali M. al-, 375

*Name index*

- Katičić, Radoslav, 166  
Kaufmann, Terrence, 422  
Kelemen, Ivett, 353, 354, 356  
Kern, Frank J., 130, 133, 135  
Khatymtayev, S. H., 427, 429, 430  
Kieffer, Jean Daniel, 248, 250  
Kilgarriff, Adam, 13, 14  
Kiss, Mihály, 365, 377–379  
Kjølaas, Per Oskar, 352  
Klausing, Anton E., 328  
Klemenčič, Matjaž, 105, 106  
Klöter, Henning, 298–300  
Kocyba-Grych, Katarzyna, 274, 275,  
277, 292, 293  
Kola, Dymitr F., 331, 335, 337  
Kolecka, Zuzanna, 274, 275, 277, 292,  
293  
Komsuoğlu, Ayşegül, 246, 261  
Konopczyński, Władysław, 323  
Kontler, László, 354  
Korabinsky, Johann Matthias, 248,  
250  
Košutnik, Silvester, 129–131, 134, 135  
Kowalski, Franciszek, 344  
Krašić, Stjepan, 166  
Krašovec, Jože, 142, 157, 158  
Krauss, Werner, 85  
Kubelka, Victor J., 115–128, 134, 136  
Kuryłowicz, Beata, 335  
  
Lach-Szyrma, Krystyn, 322  
Lach-Szyrma, Krystyn, 336  
Lackner, Michael, 299  
Lacombe, Jacques, 213, 214, 218–221,  
224, 225  
Lancelot, Claude, 200, 202  
Landau, Sidney I., 163–165, 170, 331  
Landry-Deron, Isabelle, 315  
Langendonck, Willy van, 146  
  
Lanza, Nicolas, 8, 399, 414  
Laplanche, François, 230, 237, 238  
Le Brun, Jacques Le, 238  
Le Mollé, Roland, 217  
Lea, Henry A., 35  
Leca-Tsiomis, Marie, 230  
Leem, Knud, 353, 358  
Legan Ravnikar, Andreja, 159, 160  
Lehmann, Alise, 200  
Leung, Cecile, 310  
Lewis, Geoffrey, 247, 248, 253  
Lewis, Robert E., 45, 48  
Lewy, Heinrich, 419  
Li, Hui, 301–304, 309, 312, 315  
Libiszowska, Zofia, 325  
Ligara, Bronisława, 249, 264  
Lighter, Jonathan E., 34  
Lillis, Theresa, 18  
Lillo, Jacqueline, 400, 401, 403, 407,  
411, 414  
Linde, Samuel B., 339  
Lipoński, Wojciech, 323  
Locke, John, 80, 81, 89  
Lovrić Jović, Ivana, 172  
Lučić, Josip, 185  
Łukaszewicz, Józef, 324  
Lukszyn, Jerzy, 247  
Luna, Paul, 19  
Luó Chángpéi 羅常培, 300  
Luzac, Elie, 87  
Lynch, Jack, 12  
  
Macaro, Ernesto, 334  
Maës, Gaëtane, 214, 220  
Magay, Tamás, 373  
Magga, Ole Henrik, 352  
Magocsi, Paul R., 144  
Mallet, Edme-François, 78  
Mamadnazarov, Abdusalom, 428

- Marcu, Eva Dorothea, 90  
Marinelli, Bruno, 312  
Marmontel, Jean François, 78, 83, 84  
Marsy, François-Marie de, 213, 214,  
218–221, 223–225  
Martin, Henri-Jean, 213, 214, 221  
Martinussen, Bente, 353  
Mártonffy, Frigyes, 368, 377, 379, 382  
Masini, Federico, 300  
Masson, Émilía, 419  
Matasović, Ranko, 167  
Maticsák, Sándor, 354  
Matthews, Geoffrey J., 144  
Mazi-Leskovar, Darja, 102  
Mazière, Francine, 222  
McDermott, Anne, 59  
McGready, Alfred Glen, 419  
Ménage, Gilles, 200, 204  
Meniński, Franciscus à Mesgnien,  
248, 250  
Messaoudi, Leila, 214, 218  
Meštrović, Zrnka, 165, 170  
Miege, Guy, 328  
Mikhailov, Nikolai, 145  
Milharčić-Hladnik, Mirjam, 106  
Minerva, Nadia, 400, 401, 403  
Miodunka, Władysław, 248  
Mirović, Pavo, 172, 181  
Moon, Rosamund, 59  
Motolese, Marco, 217  
Moureau, François, 94  
Muggleston, Lynda, 12, 16, 23  
Mukhtorov A, Sayfullaev A, 429  
Muljačić, Žarko, 172  
Mungello, David, 301  
Murano, Michela, 165, 400–404, 407,  
411, 413  
Murray, James A. H., 11, 18–20  
Murray, K. M. Elizabeth, 12  
Muss-Arnolt, W., 419  
Nardelli, Jean-Fabrice, 422  
Neumann, Günter, 419  
Nicolaisen, Wilhelm F. H., 142, 157  
Nicole, Pierre, 80, 81  
Nicoll, Josephine, 324  
Nielsen, Sandro, 375  
Niu, Xiaochun, 110  
Noyes, Gertrude Elizabeth, 60  
Nussbaum, Alan, 422  
Obrador-Cursach, Bartomeu, 419  
Oláh, Tibor, 363, 372, 389, 396  
Ölmez, Mehmet, 247  
Orazi, Carlo, 302–305, 310  
Osiński, Alojzy, 337  
Özalp, Mehmet Nazmi, 246, 253  
Özcan, Nuri, 258  
Öztuna, Yılmaz, 253, 263  
Pálffy, Miklós, 363, 372, 377, 389, 396  
Parlatir, İsmail, 249  
Paternicò, Luisa, 298  
Pawlina, Agata, 246, 261, 262  
Pellandra, Carla, 400, 401  
Pellin, Tommaso, 299  
Peng, Jing, 2  
Pernety, Antoine-Joseph, 213–215,  
218, 219, 221, 223, 224, 226  
Perrot, Jean, 366  
Peverelli, Peter, 299  
Peyraube, Alain, 298  
Peyton, V. J., 339  
Piechota, Marek, 324  
Pierce, Richard Holton, 419  
Piotrowski, Tadeusz, 375  
Piotrowski, Wojciech, 324  
Podhajecka, Mirosława, 322, 328

*Name index*

- Pogorelec, Breda, 145  
Pokorny, Jenő Ármin, 369, 377, 379, 383  
Polesi, Giacomo, 402  
Polmonari, Simonetta, 303  
Pomey, 204  
Prisland, Marie, 102, 105  
Proffitt, Michael, 330, 334  
Prokosch, Erich, 247  
Pruvost, Jean, 164, 211, 218, 221  
Putanec, Valentin, 170, 172  
Putter, Ad, 43
- Quemada, Bernard, 201–203, 411  
Quémada, Bernard, 195
- Radziwiłł, Antoni, 275, 293  
Raini, Emanuele, 298, 299, 304, 315  
Raukko, Jarno, 142–144, 157  
Redhouse, James W., 248, 250, 254  
Rees, Emma, 15  
Reich, Vicky, 38  
Reichborn-Kjennerud, Finn, 378  
Reinhard, Kurt, 246, 263  
Reinhard, Ursula, 246, 263  
Rey, Alain, 163–165, 168, 169, 202  
Rey, Christophe, 198  
Richelet, Pierre, 194, 195, 197, 198, 200, 202–206, 211, 219  
Rocchi, Luciano, 248  
Rosenthal, David S. H., 38  
Rosier-Catach, Irène, 196  
Rosół, Rafał, 419  
Rothwell, William, 43, 45  
Rundell, Michael, 13, 14  
Rutten, Stuart, 51  
Rykaczewski, Erazm, 322
- Sager, Juan C., 248
- Sági, István, 373  
Sajnovics, János, 355–358  
Samardžija, Marko, 163–166, 170, 171, 188  
Sāmī, Şemseddin (Ch. Samy-Bey Fraschery), 248, 250  
Samuels, Michael Louis, 52  
Sanal, Haydar, 261  
Sartori, Orietta, 312  
Sauer, Hans, 60  
Sauvageot, Aurélien, 371, 377, 386, 396  
Savile, George, 86–88  
Sayfullaev, A., 429, 430  
Schendl, Herbert, 44, 50  
Schlücker, Barbara, 142  
Schmitt, Rüdiger, 419  
Schweickard, Wolfgang, 248  
Sclafani, Marie-Denise, 400  
Sedmak, Jay, 107  
Šekli, Matej, 156  
Sgard, Jean, 213  
Sieber, Patricia, 310  
Siekierska, Krystyna, 322, 329  
Siemieniec-Gołaś, Ewa, 247, 248, 250, 264  
Simon, Richard, 237, 239  
Simon, Zsolt, 419  
Simotas [Σιμωντάς], Panagiotis N. [Παναγιώτης Ν.], 419  
Simpson, John, 16, 33, 330, 334  
Sinapian, Krikor, 248, 250, 262  
Sironić Bonefačić, Nives, 170  
Słowikowski, Adam, 324  
Söderblom Saarela, Mårten, 304  
Solbakk, Aage, 353  
Soulard, Delphine, 89  
Sowa, Wojciech, 419  
Spencer, Philip, 322

- Sroka-Grądział, Anna, 271, 272, 292  
Stachowski, Marek, 249  
Stachowski, Stanisław, 248, 250  
Stafutti, Stefania, 301  
Starnes, DeWitt Talmage, 60  
Stedall, Ellie, 15  
Stein, Gabriele, 60, 61  
Stein, Heidi, 247  
Stolzenberg, Daniel, 233  
Striedter-Temps, Hildegard, 160  
Strzyżyńska, Weronika, 18  
Sullivan, Hannah, 18  
Susel, Rudolph M, 106  
Svensén, Bo, 374  
Sylvester, Louise, 52, 54  
Szabó, Dávid, 363, 366, 372, 390, 396
- Tallarico, Giovanni, 165  
Tambrun, Brigitte, 238  
Teich, Elke, 339  
Thébaud-Sorger, Marie, 239  
Theisz, Gyula, 368, 370, 377, 379, 385  
Thomason, Sarah Grey, 422  
Thomson, Ann, 85  
Tietze, Andreas, 249  
Tinghir, Anton B., 248, 250, 262  
Tohumcu, Ahmed, 245  
Tóth, Anikó, 351  
Tovar, Sofía Torallas, 419  
Townend, Matthew, 44  
Trench, Richard Chenevix, 21  
Trotter, David, 43–45, 49, 50, 53–55  
Trotz, Michał Abraham, 331, 335, 337  
Trotz, Michel A., 335  
Trucchi, Giuseppe, 8, 399, 403, 406  
Turan, Namık Sinan, 246, 261  
Turville-Petre, Thorlac, 51
- Uz, İsmail Kâzim, 248, 250, 262
- Vajs, Nada, 170  
Vaxelaire, Jean Louis, 165  
Vedal, Nathan, 299  
Vermes, Albert Péter, 143, 146  
Victor, Bret, 33  
Victor, Terry, 34  
Von Collani, Claudia, 301  
Vončina, Josip, 170
- Walczak, Bogdan, 322  
Weiner, Edmund, 16  
Wells, Ronald A., 60  
Wheatley, Henry B., 60  
Wieger, Léon, 274  
Wieluch, Doman, 271–276, 278–293  
Wierzbicka, Anna, 165  
Wionet, Chantal, 230  
Wogan-Browne, Jocelyn, 43  
Wolff, Christian, 91, 93  
Wollman, Howard, 322  
Wright, Laura, 44, 49, 50  
Wright, Owen, 246  
Wu Yuemei (吴月梅), 275, 277  
Wüster, Eugen, 217
- Xia Zhengnong (夏征农), 275, 292, 293  
Xu Lin (许琳), 274, 275, 277, 292, 293  
Xu Minglong 许明龙, 310
- Yalçın, Gökhan, 245  
Yang, Huiling, 298  
Yang, Paul Fu-mien, 299  
Yao Xishuang (姚喜双), 274, 275, 277, 292, 293  
Yershov, N. N., 428  
Yin Xiangfeng (殷相峰), 275, 277  
Yong, Heming, 2  
Yöre, Seyit, 246, 261

*Name index*

Yu, Dong Clara, 301

Zalta, Edward N., 35

Zgusta, Ladislav, 164, 333

Zhou, Gang, 110

Zielinski, Lech, 35

Zmarzer, Wanda, 247

Zolli, Paolo, 404

Žugel, Franc, 102, 105, 106

Zwartjes, Otto, 299, 303



# West meets East

Lexicography, in its many forms, is a very old, practical discipline solving practical problems concerning word usage. The term “word” seems more appropriate than “language” in this context, as lexicography addresses more questions relating to what we now call lexicology. As with all areas of human endeavour, what developed gradually through trial and error has eventually been subjected to a theoretical framework. The role of historical lexicography is to look back on the development of these highly varied word lists to understand how we arrived at the tremendous variety that characterises practice throughout the world.

This volume is both a selection of expanded papers from one conference on historical lexicography and lexicology, held under the aegis of the International Society for Historical Lexicography and Lexicology (ISHLL) in Lorient, France, in May 2022, and also the first in a new book series dedicated to the field. The new series represents a collaboration between two sister associations, ISHLL and the Helsinki Society for Historical Lexicography (HSHL). The volume contains texts in both English and French that provide insights into dictionaries, their compilers and users using evidence from numerous languages across the globe. It is also diachronic, moving from topics on medieval usage to contemporary issues concerning open access and digital publishing in historical lexicography. The title reflects the global scope of its authors and content, encompassing Japan to the United States, Eastern Europe to the United Kingdom, and Portugal.