

Emergencia de nuevos géneros textuales y terminología
en la historia de los lenguajes de especialidad

Linguistica Philologica

Lingua – Usus – Variatio

Edited by
Gerda Haßler and Vladislava Warditz

Volume 2

Editorial Board

Sylvie Archaimbault (Paris-Sorbonne University, France)
Bernard Colombat (Paris Diderot University, France)
Olga Glazunova (Saint Petersburg State University, Russia)
Toon Van Hal (Catholic University of Leuven, Belgium)
Nicola McLelland (University of Nottingham, UK)
Thorsten Roelcke (Technical University of Berlin, Germany)
Alan Timberlake (Columbia University, New York, USA)
Fedor Uspenskij (Vinogradov Russian Language Institute, Moscow, Russia)

Natalia Terrón Vinagre / Jenny Brumme

**Emergencia de nuevos géneros
textuales y terminología en
la historia de los lenguajes de
especialidad**



PETER LANG

Lausanne - Berlin - Bruxelles - Chennai - New York - Oxford

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.



ISSN 0531-7312
ISBN 978-3-631-89760-7
E-BOOK 978-3-631-89761-4
E-PUB 978-3-631-89762-1
DOI 10.3726/b21165

© 2023 Peter Lang Group AG, Lausanne

Verlegt durch:
Peter Lang GmbH, Berlin, Deutschland

info@peterlang.com <http://www.peterlang.com/>

Alle Rechte vorbehalten.

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

A Gerda Hassler, en su 70.º cumpleaños

Tabla de contenido

Lista de Colaboradores	11
------------------------------	----

Presentación	13
--------------------	----

BLOQUE I. EMERGENCIA DE NUEVOS GÉNEROS TEXTUALES

Gerda Hassler

Características textuales de los textos científicos y la fijación de una tradición discursiva en la recepción de los <i>Spectators</i> en Europa Textual characteristics of scientific texts and the fixation of a discursive tradition in the reception of <i>Spectators</i> in Europe	21
--	----

Carsten Sinner

A bifurcação de géneros textuais em português europeu (XVIII-XXI) The Bifurcation of textual genres in European Portuguese (18th to 21st centuries)	39
--	----

Bertha M. Gutiérrez Rodilla

Los repertorios lexicográficos de medicina «modernos»: un género renacido y renovado en los siglos XVIII y XIX The “modern” lexicographical repertoires of medicine: a genre reborn and renewed in the 18th and 19th centuries	61
---	----

Mónica Vidal Díez

La tradición discursiva recogida en las <i>Memòries manuscrites</i> de la Reial Acadèmia de Medicina de Catalunya, 1820–1900 The discourse tradition collected in the <i>Memòries manuscrites</i> of the Reial Academia de Medicina de Catalunya, 1800–1900.	75
--	----

Antoni Nomdedeu-Rull y Xavier Torreadella-Flix

Creación y evolución de la crónica deportiva en la prensa de noticias española. Estilo literario y lenguaje técnico Creation and evolution of the sports chronicle in the Spanish news press. Literary style and technical language	97
--	----

<i>Encarnación Tabares Plasencia y Carsten Sinner</i>	
El testamento vital o instrucciones previas: análisis jurilingüístico y textual de un género en desarrollo	
The Living Will or Advanced Directives: a Jurilinguistic and Textual Analysis of a Developing Text Genre	117

<i>Yvonne Kiegel-Keicher</i>	
Recursos específicos de la difusión y popularización de materia especializada en la crisis de la Covid-19	
Mechanisms of diffusion and popularization of COVID-19-specific terminology during the pandemic	139

<i>Thierry Nallet</i>	
El cuento especializado: género, estilo y recursos lingüísticos	
The specialised short story: genre, style and linguistic resources	157

BLOQUE II. TERMINOLOGÍA

<i>María Teresa Cantillo Nieves</i>	
Terminología metalúrgica y minera de los siglos XVI y XVII: propuesta de organización semántica	
Metallurgical and mining terminology of the 16th and 17th centuries: semantic organization proposal	179

<i>Maria José Bocorny Finatto, Maria Filomena Gonçalves y Rafaela Radünz Lazzari</i>	
Léxico e terminología em um novo gênero textual do século XVIII: o manual para enfermeiros	
Lexicon and terminology in a new textual genre in the 18th century: the nurse's manual	199

<i>Natalia Campos Martín</i>	
Las traducciones científicas de Mariano de Larra y Langelot: <i>Aire</i> (1803), <i>Venenos</i> (1819) y <i>Plantas</i> (1836).	
Mariano de Larra y Langelot as translator of scientific texts: “Air” (1803), “Poisons” (1819) and “Plants” (1836).	233

Sandra Iglesia Martín

El diccionario como texto divulgativo en el siglo XIX: el *Diccionario Nacional* de Domínguez y el descubrimiento del vanadio

The dictionary as an informative text in the 19th century: Domínguez's *Diccionario Nacional* and the discovery of vanadium. 253

Pol Garriga Martínez

Aproximación diacrónica al vocabulario de la salud mental de principios del siglo XX en español y alemán

Diachronic approach to early 20th century mental health vocabulary in Spanish and German. 271

Lista de Colaboradores

Natalia Campos Martín
Institut Interuniversitari López Piñero. Universitat de València.

Maria Filomena Gonçalves
Universidade de Évora

Pol Garriga Martínez
Universitat Pompeu Fabra

Bertha M. Gutiérrez Rodilla
Universidad de Salamanca

Gerda Hassler
Universität Potsdam

Sandra Iglesia Martín
Universitat Rovira i Virgili

Maria José Bocorny Finatto
Universidade Federal do Rio Grande do Sul

Yvonne Kiegel-Keicher
Justus-Liebig-Universität Giessen

Thierry Nallet
Université Grenoble Alpes– ILCEA4

Antoni Nomdedeu-Rull
Universitat Rovira i Virgili

Rafaela Radünz Lazzari
Universidade Federal do Rio Grande do Sul

Carsten Sinner
Universität Leipzig

Encarnación Tabares Plasencia
Universität Leipzig

María Teresa Cantillo Nieves
Universitat Autònoma de Barcelona

Xavier Torrebaddella-Flix
Universitat Autònoma de Barcelona

Mónica Vidal Díez
Universitat de Barcelona

Maria José Bocorny Finatto

Universidade Federal do Rio Grande do Sul

Maria Filomena Gonçalves

Universidade de Évora

Rafaela Radünz Lazzari

Universidade Federal do Rio Grande do Sul

Léxico e terminologia em um novo gênero textual do século XVIII: o manual para enfermeiros

Lexicon and terminology in a new textual genre in the 18th century: the nurse's manual

RESUMO: Em Portugal, obras impressas para formação de enfermeiros surgiram bem mais tardiamente do que as publicações para médicos. A primeira obra semelhante a um «manual de Enfermagem», escrita originalmente em português, intitula-se *Postilla Religiosa, e Arte de Enfermeiros* (1741). Neste trabalho, buscam-se traços marcantes do seu perfil lexical, com vistas a uma possível definição de um gênero textual emergente no século XVIII. Para ajudar a reconhecer esse perfil, conta-se com o apoio de técnicas estatísticas. Assim, são examinadas as suas frequências de palavras, expressões recorrentes e terminologias em cotejo com uma obra médica, *Observações medicas doutrinaes* (1707).

PALAVRAS-CHAVE: lexicometria, terminologia histórica, glossários, linguagem da medicina

ABSTRACT: In Portugal, printed materials for nursing training emerged much later than those intended for physicians. The first work similar to a “nurse’s manual,” originally written in Portuguese, is entitled *Postilla Religiosa, e Arte de Enfermeiros* (1741). In this article, we seek to identify striking features of its lexical profile, aiming at a possible definition of an emerging textual genre in the 18th century. With the support of statistical techniques, the frequencies of words, expressions, and terminologies are examined in comparison with a medical handbook, *Observações medicas doutriniais* (1707).

KEYWORDS: lexicometry, historical terminology, glossaries, language of medicine

1. INTRODUÇÃO

Em Portugal, materiais impressos para formação de enfermeiros surgiram bem mais tardiamente do que os destinados a médicos (Gonçalves 2020). A primeira obra semelhante a um «manual de Enfermagem», escrita originalmente em português, intitula-se *Postilla Religiosa, e Arte de Enfermeiros* (1741). Seu autor, Fr. Diogo de Santiago (f. 1747), padre da ordem de São João de Deus (1495–1550), atuou no hospital militar de Elvas. Nessa cidade, aproveitando suas experiências no atendimento aos enfermos, escreveu a referida *Postilla*, doravante identificada por *Arte*.

Neste trabalho, buscamos traços marcantes do seu perfil lexical, com vistas a uma possível definição de um gênero textual emergente no século XVIII. Para tanto, observaremos frequências de palavras, expressões e terminologias ao longo da *Arte*, apoiando-nos em técnicas da Estatística Linguística (Biderman 1967) e em ferramentas computacionais (Berber Sardinha 2004; Quaresma 2013).

Com tal suporte, contrastaremos o repertório vocabular da *Arte*, no segmento voltado aos noviços-enfermeiros, e o da obra *Observações medicas doutrinaes*, do médico Curvo Semedo (1707). Além disso, intentaremos um brevíssimo contraponto, a ser expandido em trabalhos futuros, com parte das edições das *Gazetas Manuscritas da Biblioteca Pública de Évora* (Menezes 1673), vol. 1, 1729–1731, as quais constituem testemunhos da língua cotidiana em Setecentos.

Ao comparar a *Arte* e um guia médico, circunscrevemos, por um lado, o conjunto de palavras e/ou expressões comum às duas obras e, por outro, seus repertórios exclusivos. Com tal propósito, pressupomos que:

- a) repertórios vocabulares particulares e compartilhados, espelhados em registros escritos, podem ajudar a perceber organizações conceituais e semânticas próprias de um domínio de saber e/ou conhecimento, permitindo a percepção de «visões de mundo»;
- b) repertórios lexicais permitem apontar traços importantes da constituição de um gênero discursivo e textual emergente, caso do primeiro manual de enfermagem escrito em português no século XVIII.

Antes de proceder à análise da *Arte* e ao seu cotejo com outros textos, segue uma breve contextualização da obra em apreço. Depois, nossas referências teóricas e metodológicas, assim como uma síntese dos resultados obtidos.

2. ARTE DE ENFERMEIROS: O AUTOR E A OBRA

Diogo de Santiago faleceu em 1747, no convento-hospital de Elvas. Não há registros de seu nascimento. Atuava na Ordem dos Irmãos Hospitaleiros, cuja fundação inspirou-se na figura e ações de João de Deus (1455–1550), nascido na cidade de Montemor-o-Novo. João de Deus foi beatificado em 1630 e canonizado em 1690, notabilizado pelo socorro aos pobres e criação de hospitais.

Conforme a Câmara Municipal de Elvas¹, o Convento de São João de Deus foi erigido em 1645, o primeiro da sua ordem em Portugal, para ser um hospital militar. Em finais de 1641, D. João IV ordenou a construção de hospitais militares nas principais praças do país, entre as quais Elvas. Seu edifício foi iniciado em 1642, dentro das muralhas da cidade, e em 1653 foi ampliado para abrigar a *Vedoria Geral*. Às *vedorias* cabia a administração financeira e o provimento dos exércitos.

No entanto, teria sido sempre pequeno para a quantidade de feridos e doentes que recebia. Em 1659, ano da Batalha das Linhas de Elvas, contava com 10 religiosos que podiam atender até 350 doentes. Além disso, nele foi estabelecida, no século XVIII, uma *aula* de Cirurgia e Anatomia para estudantes dos Corpos da Província.

Isso significaria que, em 1741, Fr. Santiago, enquanto terminava a sua obra, no edifício junto à muralha de Elvas, estaria convivendo com diferentes práticos e estudantes, em um hospital movimentado, com muitas pessoas e muito trabalho. Assim, estaria exposto a diferentes conhecimentos e saberes. Com tal «moldura de significação», supomos que sua escrita seja também um testemunho de várias situações de emergência.

O que hoje conhecemos como «capacitação do enfermeiro» preenche a maior parte dos 71 capítulos da *Arte*, pois são 59 os consagrados especificamente a esse tema, ainda que a formação religiosa não deixe de estar presente no cuidado físico dos enfermos. Desde 2019, temos uma digitalização da *Arte*, em acesso gratuito, em *Internet Archive* (<https://archive.org/details/b30507340>). A Figura 1 traz a sua portada.

1 Informações do texto «Hospital Militar – Convento de São João de Deus» disponível em: <https://www.cm-elvas.pt/descobrir/patrimonio/militar/hospital-militar-convento-de-sao-joao-de-deus/> [consulta: 29-V-2023].

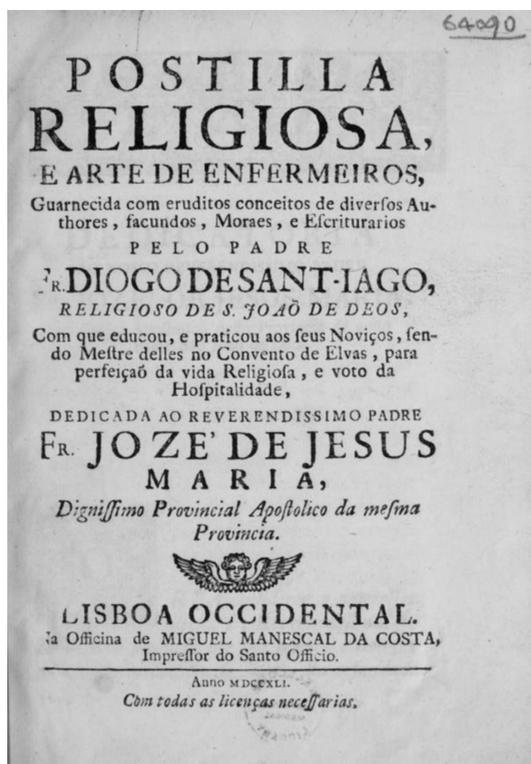


Fig. 1 – Arte – Portada. Fonte: *Internet Archive*

Nesse arquivo, como paratextos, temos a Dedicatória; o Prólogo ao Leitor; as Licenças Censórias; um Índice. Em seguida, o Tratado I, com 5 capítulos – de foco religioso; o Tratado II, com 59 capítulos – sobre Enfermagem; o Tratado III, com 7 capítulos – sobre temas de assistência à morte. Vale observar que todos os parágrafos da obra são numerados, o que buscava facilitar a consulta do leitor a partir de 2 índices remissivos. A seguir, vemos segmentos do Prólogo ao Leitor e do Capítulo I do Tratado II. Vale reiterar que reproduzimos o texto fielmente, com a ortografia conforme o original.

PROLOGO AO LEITOR

He muito vulgar em todos os Mestres pela sciencia, que professaõ, ou arte, que exerci-
taõ, ser preciso attributo das suas facultades ensinar, discorrer, e postillar, conforme a

doutrina, ou methodo, a que Calliope pelo rhetorico a huns ao thalamo conduzio, ou Minerva pela sciencia a outros nas aulas destinou, [...].

Repare pois o discreto Leitor, que nesta Postilla requintey particularmente os conselhos da politica Religiosa, com outros muitos diversos, e mais proxivamente experimentados, não para que delles se aproveitem aquelles, de quem eu posso aprender, mas sim aquelles, a quem pela occupação tenho obrigação de ensinar; e como os principais são os do nosso instituto, lhe ajuntemy os desta Arte de Enfermeiros na praxe moderna, que revista, e corregida por Medicos doutos, e Cirurgiões peritos, ficou capaz de se aproveitarem della os meus Noviços, e Religiosos da minha Sagrada Religião, em que claramente mostro aos de outra, que só para nós foy o meu postillar, e não foy para outros o meu escrever; o que bem se verifica nos das mais Sagradas Religiões, que tendo por instituto o curar almas doentes pelas culpas, a nossa só o tem de curar corpos enfermos pelas queixas, que he de que contém o segundo Tratado deste Livro, e ajudar a bem morrer o terceiro, [...].

TRATADO II.

ARTE DE ENFERMEIROS

para assistir aos enfermos, com as advertencias precisas para a applicação dos remedios.

CAPITULO I.

Advertencias para o Enfermeiro.

104 Se o Prelado vos eleger Enfermeiro, day-lhe logo o agradecimento de formar conceito da vossa capacidades para emprego de tanta importancia, e merecimento, de cuja occupação pende a saude da alma, e corpo do enfermo, credito da nossa Religião, e instituto della. Haveis de advertir, que o Enfermeiro, que he caritativo, considera que o que faz ao enfermo, Deos o recebe, estima, e remunera. Deve haver produção generica de amor entre o enfermo, e Enfermeiro; que se a alma está mais onde ama, que onde anima, como disse Santo Agostinho; e o amor transforma a amante no amado, como disse Seneca: segue-se que mais deve padecer o Enfermeiro a impulsos da caridade, com que o ama, que o enfermo pela actidaõ da queixa, que padece; donde se verefica, que o enfermo, que padece, pôde as queixas sosster; mas o Enfermeiro abraçado em amor de caridade não as pôde tolerar; porèm fazendo o que pôde, satisfaz ao que deve: applica-lhe os remedios: dá-lhe vigilante: não dorme, se o enfermo não soceda; e assim em perpetuo labyrintho não admite tregoaõs ao descanso, até que o enfermo não tenha alivio, em cuja acção cumpre com toda a Ley de Deos. Assim o escreveo Saõ Paulo aos Romanos, dizendo: *Quem ama ao proximo, toda a Ley encheo*. E nestas amorosas assistencias dos enfermos vos nõ haveis de queixar, mas sim louvar a Deos, como fazia o Santo Job nas suas tribulações.

3. HIPÓTESES DE PARTIDA E MATERIAIS EM CONTRASTE

Supondo que *Arte* conecta-se ao *modus dicendi* de manuais médicos coevos, mas com apropriação linguística e conceitual peculiares, buscamos descrever e analisar seus usos vocabulares. Assim, transcrevemos e reunimos os 59 capítulos do seu Tratado II em arquivos para processamento. Esses arquivos encontram-se, para acesso público, devidamente identificados². A seguir, um exemplo dos seus conteúdos, oferecidos em formato «somente texto» (.txt):

CAPITULO III.

Emborcação como se faz.

115

COrtareis ao enfermo todo o cabelo da cabeça à ponta da tisoura , o mais rente , que puder ser. Para melhor lhe applicares o remedio , lhe mudareis a cabeceira para os pés , pondo-lhe ao redor da cabeça hum lenço torcido , e bem apertado, em fôrma de capella, para que o cozimento não corra pelo rosto do enfermo, o qual mandareis pôr de costas, com a cabeça fóra da cama. Estando já nesta fôrma , tendo debaixo huma bacia, lhe hireis lançando o cozimento muy devagar por hum jarro de bico, com moderada quentura , tornando a encher o jarro do mesmo, que cahir na bacia. Durará esta applicação em quanto durar o calor no cozimento , o qual ha de cahir no meyo da cabeça do enfermo, e da altura de dous palmos. Acabada esta applicação , se não esfregue a cabeça do enfermo, nem em quanto se lhe faz , que seria augmentar-lhe a queixa ; só sim se lhe ha de enxugar a cabeça brandamente, e tirando-lhe o lenço, se lhe porá hum toucador.

116

Se o Medico mandar fazer alguma untura, seja tibias, e se fará depois que a cabeça estiver enxuta ; e feita a untura, se lhe porá hum papel pardo , e em cima o toucador ; advertindo , que se o tempo for frio , se fechem as janellas em quanto se fazem estes remedios.

Para contraste, escolhemos um dos muitos tratados médicos de João Curvo Semedo (1635–1719), obra que já foi explorada em outras ocasiões (Finatto *et al.* 2018): *Observações médicas* (1707), doravante denominada *Observações*. Esta obra, com cerca de 600 páginas de texto, traz 101 relatos de atendimentos a variados perfis de pacientes – crianças, mulheres, nobres, «pessoas do povo», escravos e idosos –, destinando-se também para a formação dos médicos. Em

2 Cf. Finatto, Maria José: <https://sites.google.com/view/projeto38597>, via projeto intitulado *Corpus Histórico da Linguagem da Medicina em Português (Século XVIII): Terminologia Diacrônica e Humanidades Digitais*, com apoios: CNPq – PQ e Iniciação Científica da UFRGS e FAPERGS-CAPES– 06/2018– Internacionalização. Proc.19/2551-0000718-3.

formato de transcrição «facilitada», seus arquivos também estão em acesso livre³. Abaixo, um trecho de um deles.

OBSERVAÇAM IV .

Da meſma ſenhora , que eſtando com a conjunção menſal , comeo varias frutas , as quaes , ou por muitas , ou por frias , & flatuoſas , lhe cauſaraõ taõ grande aballo no eſtomago , que brevemente as vomitou com tanta violencia , que ſe lhe originou huma pontada no lado eſquerdo , taõ aguda , & inſoportavel que me obrigou a ſangralla no braço da parte doloroſa , ſem fazer caſo da conjunção ; & foy o ſucceſſo taõ feliz , que dentro de tres dias ficou boa , & livre da morte .

SERia coufa enfadoſa referir aqui as muitas doenças , que as mulheres padecem por cauſa da madre ; baſta dizer em ſumma que a madre he authora de infinitas calamidades ; porque , como diz Hippocrates , (1.) dos meſes reprezados , ou por qualquer cauſa diminuidos , ou demaſiadamente profuſos trazem graviffimas doenças os ſeus principios . Em confirmação deſta verdade me ſeja permitido referir a Obſervação ſeguinte .

Para brevíssimo contraponto com a *Arte*, outra fonte, também já explorada (Finatto *et al.* 2018), temos o volume I da coleção *Gazetas Manuscritas da Biblioteca Pública de Évora*, com «notícias» de 1729 a 1731. Para seu acesso, em formato de arquivo processável, buscamos o *Corpus Histórico do Português Tycho Brahe* (Galves *et al.* 2017). Nesse *corpus*, há um arquivo único relativo ao volume I (Menezes 1673), com 14832 *types* (palavras diferentes) e 85517 *tokens* (palavras). A seguir, uma amostra do conteúdo, formato «somente texto» (.txt), em grafia não modernizada.

Diario de 23 de Agosto de 1729

Pelas cartas de Vasco Fernandez Cezar, se soube a noticia, que aqui todos ignoravão, de que elRey o tinha feito Conde de Sabugoza villa junto a Vizeu de que não sabemos se lhe desse senhorio.

Chegou Rodrigo Cezar, gordo, mas não cheio, mostrou grande desin-teresse; as minas que descobrio tem grande quantidade de ouro, e se achou hum grão de meia arroba, porem he máo o clima, e tão dillatado o caminho, que comeu varias cousas ascarozas, ficou de posse do governo Antonio da Sylva Pimentel.

Elegeu-se abbadessa de Santa Clara Donna Mauricia, de muita capacidade tendo praticado Fr. Antonio da Piedade a renuncia voluntaria da sua antecessora, de que se espera o sucego daquelle convento.

Chegou a náõ de Macáo com João Baptista Rollano, e a vida da India, por donde se sabe que sem embargo da conquista de Mombaça e todas as suas grandes dependencias, não faltava.

3 <https://sites.google.com/view/projeto38597/transcri%C3%A7%C3%B5es-de-observa%C3%A7oes-medicas-doutrinaes-1707>.

Conforme se observa pelos trechos dos arquivos sob exame, para realizar comparações, foi necessário proceder a uma padronização quanto aos estilos de transcrição. Há divergências, em particular o uso de *f*, *ſ* (*s* curto e *s* longo) e *S*, os espaços originais antes dos sinais de pontuação e, especialmente, a separação de palavras com hífen, quebras de linhas; o que é feito em um arquivo não há em outro. Assim, realizamos uma padronização básica, somente para viabilizar o processamento: ignoramos apenas a separação silábica translinear, mantivemos as junturas originais e preservamos as variações e traços da escrita original de cada fonte, tal como em *agoa* e *agua*, *mão* e *mau*, *tisoura* e *tesoura*, *abſceſſos* ou *dailhe* etc.

4. FUNDAMENTOS TEÓRICO-METODOLÓGICOS

Para a análise do repertório vocabular da *Arte*, seguimos orientações teórico-metodológicas da *Terminologia Diacrônica* (Dury / Picton 2009) e da *Léxico-estatística* de base computacional (Quaresma 2013). Buscamos os itens lexicais mais e menos frequentes na *Arte* e *Observações* e quisemos identificar os itens comuns às duas obras e exclusivos de cada uma.

Aos elementos levantados, agregamos elementos conceituais e semânticos colhidos em outros *corpora* históricos (representados aqui, *a priori*, pelas *Gazetas Manuscritas*) e em dicionários antigos e atuais, de maneira a apontar elementos potencialmente úteis para caracterizar um gênero textual e discursivo emergente no século XVIII. Com eles, temos um estudo de *Lexicologia Histórica* e *Diacrônica*, ao mesmo tempo, estrutural e quantitativo.

Antes de mais nada, frisamos que a noção de palavra presente nesse tipo de estudo, apoiado em ferramentas computacionais, tende a ser a de *palavra gráfica*, tratando-se de *escrita* em um enfoque bastante básico ou, até mesmo, rudimentar. Assim, *grosso modo*, uma palavra, colhida nas nossas fontes-arquivo, será apenas um conjunto de letras/caracteres sequência entre dois espaços em branco. Não obstante, convém alertar que, a despeito da simplificação, tal tipo de *palavra*, nesse enfoque, corresponde a toda unidade linguística mínima que possa constituir significado, delimitado na escrita por dois espaços em branco e/ou sinal de pontuação.

Tanto o conceito de palavra como o de língua variam bastante, consoante o tipo de Linguística com que se trabalha e conforme o ponto de vista, já que este estabelece o objeto. Assim, a título de exemplo, *pé-de-moleque* tanto pode ser uma só palavra só, como, pelo contrário, pode ser tratada como uma palavra composta por outras três. No nosso estudo, que se dedica apenas à língua escrita de um modo relativamente superficial, vamos considerar a distribuição

e frequência de nossas palavras (gráficas) em dois grupos «clássicos», categorização que não está isenta de controvérsia: as palavras gramaticais, que têm um valor mais instrumental, com o seu sentido dependente da combinatória em que se encontram (artigos, determinantes, pronomes, preposições e partículas); e as lexicais, cujo conteúdo e significado plenos se associam a fatos, fenômenos, processos (substantivos, adjetivos, verbos).

No Brasil, desde o ano 2000, conhecemos um novo tipo de Linguística: a *Linguística de Corpus* (LC), apresentada, pioneiramente, por Berber Sardinha (2004). Essa é uma Linguística que aposta nas descrições, com apoio computacional, dos usos das palavras conforme realizados em vastas coleções de textos reunidas em formato digital.

A LC lida com amostras de usos de uma língua que são armazenadas em formato de arquivos e que servem para representar uma língua no seu todo. Seu enfoque é marcadamente quantitativo, do qual se parte para as considerações qualitativas. Em LC, coleções de textos sob exame são chamadas de *corpus* ou *corpora*, compostos e analisados com auxílio computacional. Ferramentas apoiam a percepção de modos e de frequências de emprego, «regras» ou padrões mais ou menos recorrentes da combinação de determinadas palavras entre si, como também permitem identificar agrupamentos recorrentes, chamados *clusters*.

A LC, assim, estabelece que o sentido de uma palavra dependerá, de diferentes modos, das companhias que ela costume ter em diferentes *corpora* ou em um dado *corpus*. Um exemplo disso seria a verificação dos empregos do verbo «causar» em diferentes *corpora* do português brasileiro atual. Geralmente, esse verbo é mais frequentemente seguido de uma palavra de valor negativo como «doença» do que de uma palavra de valor positivo como «alegria». Para tal evidência, consideram-se frequências e colocações mais ou menos frequentes, observadas ao longo de muitos textos, reunidos em *corpora* de estudo e de contraste, especialmente organizados, que representam um dado estado de língua.

A seguir, está a parte descritivo-analítica do nosso trabalho. Nela, filiando-nos à LC, detalhamos os diferentes usos de palavras ao longo de *Arte frente a Observações*. Assim, itens como *Cypriano, póde-se* ou *herva-de-bicho* serão considerados como três palavras únicas. Também observaremos os agrupamentos de palavras recorrentes ao longo dos textos, destacando os grupos repetidos de três elementos (denominados trigramas), como *gema de ovo*.

5. DESCRIÇÃO E ANÁLISES

Nos arquivos para processamento, ignoramos quebras de linhas e hifenação translinear, e não respeitamos um padrão regular de espaçamento antes de sinais de pontuação. Afinal, a ferramenta de processamento que utilizamos, o *AntConc* (Anthony 2022), despreza sinais de pontuação, hífens, números e espaços. Nessa ferramenta, como em LC, vale frisar, cada *palavra* é apenas uma sequência de caracteres separada por espaço.

A partir dos 59 textos da *Arte*, em formato .txt, com o *AntConc*, geramos uma lista de todas as suas palavras. Além de poder produzir uma lista isolada de cada capítulo, a ferramenta também aponta valores de *types* (palavras diferentes: 2395) e *tokens* (palavras: 15837). Uma parte da listagem global de itens reconhecidos pela ferramenta está a seguir.

	<i>Type</i>	Rank	Freq	Range
1	de	1	726	1
2	se	2	705	1
3	e	3	647	1
4	o	4	619	1
5	que	5	567	1
6	a	6	456	1
7	enfermo	7	276	1
8	lhe	8	249	1
9	com	9	231	1
10	naõ	10	221	1
11	para	11	208	1
12	em	12	175	1
13	ou	13	158	1
14	ao	14	148	1
15	do	14	148	1
16	as	16	138	1
17	os	17	129	1
18	he	18	124	1
19	da	19	110	1
20	hum	20	94	1

Vale ressaltar que o *AntConc* não filtra flexões, conjugações, nem ortografias diferentes, apresentando-as todos como *palavras diferentes*. Isto é, itens como *doente* e *doentes* ou *agoa* e *agua* são considerados *types* diferentes. Apesar das variações ortográficas e ausência de lematização dos itens listados,

essa ferramenta permite uma análise inicial. Desse modo, nessa lista, temos as palavras lexicais mais frequentemente empregadas do Tratado II, conforme a tabela 1.

Tab. 1. Itens lexicais mais frequentes da *Arte*

<i>Type</i>	Rank	Freq
enfermo	7	276
medico	23	89
agoa	29	73
enfermos	33	62
tempo	34	58
enfermeiro	40	49
hora	40	49
ajuda	58	34
fraqueza	197	11

O primeiro item mais frequente é *enfermo*, com o plural, *enfermos*, também entre os mais frequentes. Com esse item, vê-se que a *Arte*, em tese, teria como tópico principal – ou «protagonista da história» –, a pessoa doente. Em seguida, vem a figura do médico. Esse papel temático diferenciado, associado à frequência do item, tem por base a ideia segundo a qual a palavra lexical mais frequentemente usada em um texto corresponda ao seu assunto predominante. Essa é lógica embutida, por ex., em uma representação do tipo «nuvem de palavras», gerada por ferramentas computacionais e buscadores na internet.

Dos itens com frequência secundária, há aqueles relacionados ao tempo (*hora* e *tempo*). Por analogia, pensaríamos em «personagens coadjuvantes da história». Nessa posição, em constatação que parece óbvia, temos *agoa*, pois é um elemento imprescindível nas receitas de remédios, tratamentos e cuidados. Enfim, com essas frequências, podemos estimar «personagens principais» e «coadjuvantes», entre os quais estariam os *enfermeiros*.

Com efeito, na *Arte*, a figura do *enfermeiro* parece ter menos destaque do que a figura do *medico*. Isto é, em termos estatísticos, possivelmente as *personas* dos monges e/ou padres, ou mesmo os noviços-enfermeiros, teriam menos proeminência do que os médicos, naquele cenário. No «Prologo ao leitor» (parte que não integrou o conjunto dos 59 capítulos sob exame), o autor enfermeiro declara submeter-se ao saber de médicos e cirurgiões em relação ao conteúdo apresentado: «[...] e como os principais são os do nosso instituto, lhe ajuntee

os desta Arte de Enfermeiros na praxe moderna, que revista, e corregida por Medicos doutos, e Cirurgiões peritos [...]» (Santiago 1741: f. 12).

Quanto a itens relacionados ao assunto *doenças*, como suas denominações ou menção a sintomas, o primeiro elemento mais frequente na lista da *Arte é fraqueza*, mas apenas na posição 197 do nosso *ranking*. Quanto a tipos de tratamentos/remédios, há outro termo: *ajuda*. Por Gonçalves (2020), sabemos que *ajuda(s)* corresponde à denominação atual *laxante e/ou laxativo*.

Em linhas gerais, considerando esses *scores* brutos de frequência mais elevados e os contextos de suas ocorrências, poderíamos pensar também que o foco temático da *Arte* tenderia a privilegiar os sintomas de doenças/males ou *achaques* mais conhecidos àquela época. A tabela 2 sintetiza alguns casos de frequência relativa, em percentuais quanto ao universo de palavras sob exame.

Tab. 2. Porcentagem da frequência das palavras de *Arte*

Total dos 59 capítulos	15.837 palavras/tokens
Palavra/type	% do total de palavras
enfermo	1,74 %
medico	0,56 %
agoa	0,46 %
enfermos	0,39 %
tempo	0,37 %
enfermeiro	0,31 %
hora	0,31 %
ajuda	0,21 %
fraqueza	0,07 %

Fator importante para a observação quantitativa do vocabulário de um texto é a chamada «riqueza lexical» ou *type-token ratio* (TTR). Calculada por meio da Equação 1, essa razão apresenta, de forma preliminar, o quão diversificado se mostraria o vocabulário de um dado texto: uma TTR de 100 % significa que nenhuma palavra é repetida, enquanto 0 %, que todas as palavras se repetem pelo menos uma vez.

Equação 1 – *Token-type ratio*

$$TTR = (\text{types} * 100) / \text{tokens}$$

Na Tabela 3, temos esses valores para o Tratado II.

Tab. 3. Valores de *tokens*, *types* e TTR da *Arte*

<i>Tokens</i> (número total de palavras)	15837
<i>Types</i> (número de palavras diferentes)	2395
TTR (riqueza lexical)	15,12 %

A TTR é útil também para estimar o provável grau de dificuldade de leitura de um texto, pois «o uso de palavras desconhecidas para o leitor [...] interfere negativamente na legibilidade do texto» (Liberato / Fulgêncio 2007: 114). Assim, quanto mais variado for o vocabulário de um texto, maior é a hipótese de o leitor encontrar um termo que não conheça. No caso da *Arte*, embora haja bastante variação ortográfica, por se tratar de um texto didático, podemos alvitrar que, em Setecentos, era importante que o texto fosse de fácil compreensão. Para situar essa TTR, cotejamo-la com outros *corpora*, textos e materiais escritos de épocas e gêneros diferentes, incluindo o volume 1 das *Gazetas Manuscritas*. A tabela 4 compara os valores observados.

Tab. 4. Comparativo de valores de TTR entre *Arte* e outras fontes

Corpus/material	Século(s)	TTR
<i>Arte</i> de Diogo de Santiago	XVIII	15,12 %
<i>Gazetas Manuscritas</i> de Évora 1729–1731 – <i>Corpus</i> Tycho Brahe	XVIII	6,89 %
Machado de Assis – Primeira fase: romance <i>Helena</i> (Pellegrini e Petkoff 2009)	XIX	15,23 %
Machado de Assis – Segunda fase: romances <i>Memórias Póstumas de Brás Cubas</i> , <i>Dom Casmurro</i> e <i>Esau e Jacó</i> (Pellegrini e Petkoff 2009)	XIX e XX	9,62 %
Machado de Assis – contos (Pellegrini e Petkoff 2009)	XIX	24,23 %
<i>A casa das sete mulheres</i> , romance de cunho histórico da autora brasileira Letícia Wierzchowski	XXI	9,07 %
Textos jornalísticos de divulgação científica para leigos sobre biodiversidade (Belmonte 2007)	XXI	40 %
Textos científicos sobre biodiversidade (Belmonte 2007)	XXI	29 %

A variedade lexical da *Arte* aproxima-se mais dos romances dos séculos XIX-XXI, textos cujos autores, em tese, não têm preocupação em não repetir palavras. Por outro lado, vemos uma tendência para uma maior variabilidade de vocabulário nos textos jornalísticos e artigos científicos (textos especializados) atuais.

A maior repetitividade de palavras, conforme mencionado, tenderia a tornar a leitura de um texto menos penosa e, conseqüentemente, seu conteúdo seria mais acessível ou facilitado em termos de compreensão da mensagem. Uma exceção a esse padrão, podemos verificar no caso das *Gazetas*, textos comparáveis, *grosso modo*, a jornais «pré-históricos». Nesses textos, há um valor máximo de repetitividade de palavras em relação aos outros, algo a ser investigado. No Brasil, na atualidade, fugir, ao máximo, da repetição de palavras tornou-se princípio de qualidade da redação jornalística. Portanto, há esforços bastante evidentes para não repetir palavras ao longo de um relato, notícia ou reportagem (Belmonte 2007).

Apesar de sua utilidade para a estatística lexical, o valor de TTR na *Arte* é preliminar. Afinal, não fizemos a lematização ou normalização dos itens do vocabulário, tampouco é ponderada a quantidade de *hapax legomena*, vale dizer, as palavras que ocorrem apenas uma vez ao longo dos 59 textos. Então, para complementar a quantificação inicial da TTR, buscamos também a Medida (ou Estatística) de Honoré (MH). MH integra a contagem de *hapax legomena* com o número de *types* e *tokens* de um texto, considerando-se que a «fatia» de palavras únicas em um texto qualquer seja, normalmente, bastante volumosa. Isso, em estatística de vocabulário, chama-se de *Princípio da cauda longa*.

Para obter essa MH, utilizamos a Equação 2, em que «N é o número total de *tokens* (total de palavras no texto, contando as repetições), V_1 é o número de *hapax legomena*, e V é o número de *types* (quantidade de palavras sem considerar suas repetições)» (NILC-Metrix 2021).

Equação 2 – Medida de Honoré (MH)

$$\text{Medida de Honoré} = (100 * \log N)/(1 - (V_1/V))$$

O resultado obtido não tem um valor superior máximo de referência, mas, quanto maior o valor mais rico lexicalmente tende a ser o texto. Nosso cálculo foi feito no *Excel*. O valor de *hapax legomena* foi obtido partir da lista geral de palavras da *Arte*, gerada pelo *AntConc*. O resultado está na tabela 5.

Tab. 5. Valores da MH de *Arte* – em 59 capítulos

N (<i>tokens</i>)	15837
V (<i>types</i>)	2395
V_1 (<i>hapax legomena</i>)	1264
% de <i>hapax legomena</i> com relação ao nº de <i>tokens</i>	7,98 %
% de <i>hapax legomena</i> com relação ao nº de <i>types</i>	52,78 %
Medida de Honoré (MH)	889,32

Calculamos também a TTR e a MH para três capítulos, em isolado, tomados ao acaso da *Arte*, cujos resultados se encontram na tabela 6. Isso foi feito para comparar os valores obtidos com os do Tratado II inteiro e para estimar variações quantitativas pontuais.

Tab. 6. TTR e MH de três capítulos de *Arte*

	Capítulo 1	Capítulo 20	Capítulo 40
<i>Tokens</i>	941	69	105
<i>Types</i>	398	50	59
<i>Hapax legomena</i>	265	38	35
Medida de Honoré (MH)	889,84	766,19	496,88
TTR	42,30 %	72,46 %	56,19 %

Os valores de TTR para os três capítulos da *Arte* variam bastante entre si. O capítulo 40, que tem a menor quantidade de *tokens*/palavras, possui a maior riqueza lexical. A TTR, por sua vez, diminui de acordo com o aumento do número de *tokens*. Do mesmo modo, a MH também varia bastante dependendo do capítulo: aquele com a maior quantidade de *tokens* (capítulo 1) tem um valor quase igual ao do Tratado II completo, ao passo que o capítulo com o menor número de *tokens* (capítulo 20) tem uma MH menor, mas ainda próxima, se comparada com a do capítulo 40. Assim, as TTR dos capítulos são todas maiores do que o valor para o Tratado II como um todo, enquanto os valores da MH, por sua vez, ou quase são iguais ou são mais baixos.

Quanto aos grupos de três elementos (trigramas) mais frequentemente repetidos, observamos, na tabela 7 que, na *Arte*, focam-se principalmente no doente, o tema principal do Tratado II, e na figura do médico, aquele cujas instruções deveriam ser seguidas pelo enfermeiro.

Tab. 7. Trigramas mais frequentes de *Arte*

Trigrama	Rank	Freq
que o enfermo	1	35
se ha de	2	31
que o medico	3	29
se o enfermo	3	29

(continúa)

Tab. 8. Continued

Trigrama	Rank	Freq
se haõ de	5	25
de fõrma que	6	23
se o medico	7	21
ha de dar	8	19
o enfermo naõ	9	16
e se for	10	14

Analisando somente trigramas sem itens gramaticais em meio ao trio (tabela 8) ou com apenas um, observamos que, enquanto a *Arte* apresenta trigramas referentes a alimentos e ingredientes (como *caldo de galinha*), seu foco está, em tese, no corpo do doente, os membros a serem tratados, seu *fluxo de sangue*, a *saude do enfermo*.

Tab. 8. Trigramas mais frequentes de *Arte*

Trigrama	Rank	Freq
barrigas das pernas	25	9
caldo de galinha	32	8
cabeça do enfermo	53	6
fluxo de sangue	53	6
estomago do enfermo	85	5
gadelha de lâ	85	5
gema de ovo	85	5
cama do enfermo	130	4
saude do enfermo	130	4
achaque no peito	203	3

Foram ainda calculadas as porcentagens das frequências dos trigramas com relação ao total de palavras dos 59 capítulos da *Arte*. Os resultados estão na tabela 9.

Tab. 9. Porcentagem da frequência dos trigramas de *Arte*

Total	15837 palavras
Trigrama	% do total de palavras
barrigas das pernas	0,057 %
caldo de galinha	0,051 %
cabeça do enfermo	0,038 %
fluxo de sangue	0,038 %
estomago do enfermo	0,032 %
gadelha de lâ	0,032 %
gema de ovo	0,032 %
cama do enfermo	0,025 %
saude do enfermo	0,025 %
achaque no peito	0,019 %

Depois de caracterizado, desse modo, o vocabulário da *Arte*, cotejamos os itens lexicais mais frequentes das duas obras, reunidos na tabela 10.

Tab. 10. Itens lexicais mais frequentes de *Observações e Arte*

Obra	Type	Rank	Freq
<i>Observações</i>	remedios	31	538
	agua	33	505
	remedio	36	456
	dias	41	390
	doente	44	363
	dores	46	351
	medicos	53	287
	febre	75	205
	curar	70	215
<i>Arte</i>	enfermo	7	276
	medico	23	89
	agoa	29	73
	enfermos	33	62
	tempo	34	58
	enfermeiro	40	49
	hora	40	49
	ajuda	58	34
	fraqueza	197	11

O primeiro item lexical das *Observações* é *remédios*, com o singular *remedio*. Isso indicaria que Semedo se dedicasse mais à descrição do preparo de *remédios* do que a outras questões do tratamento. Com efeito, diversas vezes Semedo descreve a receita de mais de um remédio em um mesmo relato de caso.

Os itens lexicais secundários de *Observações* ligam-se a temáticas similares às da *Arte*: tempo (*dias*), água e médicos (escrito como *agua* e *medicos*). Tempo é uma consideração importante no preparo e aplicação de remédios, assim como nas descrições dos sintomas de doentes, o que parece ser o caso com *dias*: há quanto tempo o paciente apresenta os sintomas. Água, por sua vez, parece ser ingrediente comum nas receitas. Aqui também temos quem cuidava dos doentes: médicos. Semedo muitas vezes escreve, por exemplo, sobre como outros médicos tentaram, antes dele, tratar seus pacientes ou como tratavam certas enfermidades. Além disso, *doente* e *dores* também estão na lista, a pessoa que deveria ser curada e um sintoma provavelmente comum nos casos descritos. Quanto a itens lexicais relacionados a doenças, além de *dores*, o primeiro em *Observações* é *febre*, na posição 75, bem antes do que *fraqueza* na *Arte*.

Também calculamos a porcentagem da frequência, com relação ao total de *tokens*, dos itens lexicais de *Observações* listados na tabela 10, da mesma forma como fizemos na *Arte*. A tabela 11 traz os valores para ambas.

Tab. 11. Porcentagem da frequência de *Observações* e *Arte*

Obra	Palavra/type	% do total de tokens
<i>Observações</i> (147335 tokens)	remédios	0,36 %
	agua	0,34 %
	remedio	0,31 %
	dias	0,26 %
	doente	0,25 %
	dores	0,24 %
	medicos	0,19 %
	febre	0,14 %
	curar	0,14 %
<i>Arte</i> (15837 tokens)	enfermo	1,74 %
	medico	0,56 %
	agoa	0,46 %
	enfermos	0,39 %
	tempo	0,37 %
	enfermeiro	0,31 %
	hora	0,31 %
	ajuda	0,21 %
fraqueza	0,07 %	

Como se depreende, *Arte* não se centra em descrever *doenças*; antes, parece ocupar-se mais com os tratamentos daquelas já conhecidas. Já *Observações*, visando educar futuros médicos, traz mais menções descritivas de doenças que se buscava *curar*.

Ao comparar itens lexicais com valores de frequência ponderada próximos, vemos mais dos mesmos temas: alguns itens relacionados ao tempo (*tempo* propriamente dito e *dia*) nas *Observações*, além da já mencionada *febre*, na mesma obra. Na *Arte*, temos *boca*, provavelmente associado aos cuidados, e na temática de tratamentos, *untura* e *remédios*, além de *cozimento*, relacionado ao preparo de remédios e alimentos para o doente. O item *curar* há apenas uma vez em *Arte*, nos 59 capítulos em exame (tabela 12).

Tab. 12. Itens lexicais das *Observações* e da *Arte* com frequência ponderada similar

Obra	Type	% do total de tokens
<i>Observações</i>	remédios	0,36 %
	agua	0,34 %
	remedio	0,31 %
	dias	0,26 %
	doente	0,25 %
	dores	0,24 %
	medicos	0,19 %
	dia	0,18 %
	febre	0,14 %
	tempo	0,14 %
<i>Arte</i>	tempo	0,37 %
	enfermeiro	0,31 %
	hora	0,31 %
	ajuda	0,21 %
	remedio	0,23 %
	ajuda	0,21 %
	untura	0,21 %
	remédios	0,20 %
	boca	0,17 %
	cozimento	0,17 %

Além disso, *Arte* tem um vocabulário proporcionalmente mais variado do que *Observações*, como mostra a tabela 13.

Tab. 13. Valores de TTR de *Observações* e *Arte*

	<i>Observações</i>	<i>Arte</i>
<i>Tokens</i>	147335	15837
<i>Types</i>	17447	2395
TTR	11,84 %	15,12 %

A MH das *Observações*, entretanto, mostra-se mais alta do que a de *Arte*, como aponta a tabela 14. Então, apesar de *Arte* ter mais *types* frente à quantidade de *tokens*, *Observações*, quando é considerada a quantidade de *hapax legomena* no texto, teria uma maior variedade vocabular.

Tab. 14. Valores da MH para *Observações* e *Arte*

	<i>Observações</i>	<i>Arte</i>
N (<i>tokens</i>)	147335	15837
V (<i>types</i>)	17447	2395
V ₁ (<i>hapax legomena</i>)	9994	1264
% de <i>hapax legomena</i> com relação ao nº de <i>tokens</i>	6,78 %	7,98 %
% de <i>hapax legomena</i> com relação ao nº de <i>types</i>	57,28 %	52,78 %
Medida de Honoré (MH)	1209,87	889,32

Quanto aos valores de TTR de outros *corpora*, ambas as obras têm uma taxa aproximada de repetição de palavras. E, como a *Arte*, calculamos a TTR e a MH para três capítulos, em isolado, das *Observações*. Os valores estão a seguir (tabela 15).

Tab. 15. TTR e MH de três capítulos das *Observações*

	Observação 1	Observação 30	Observação 60
<i>Tokens</i>	3119	1849	1879
<i>Types</i>	1204	741	750
<i>Hapax legomena</i>	820	526	524
Medida de Honoré (MH)	1095,52	1125,95	1086,48
TTR	38,6 %	40,08 %	39,91 %

Tal como ocorreu na comparação entre os valores para a *Arte*, aqui a obra toda tem uma variedade vocabular maior quando incluímos *hapax legomena* no cálculo, apesar de os capítulos individuais terem proporção maior de *types* com relação a seu número de *tokens*.

Quanto aos trigramas, os das *Observações* são, em sua maioria, relacionados com o preparo de remédios, tema recorrente, mencionando medidas em *onça* (hoje: 28,7g) e ingredientes (*oleo de...*, *agua de...*). A *Arte*, por sua vez, destaca *enfermo* e *medico*. Ambas as listas, a seguir, apontam, em algum nível, o provável foco temático de cada manual.

Tab. 16. Dez trigramas mais frequentes: *Observações* e *Arte*

Obra	Type	Rank	Freq
<i>Observações</i>	tres onças de	1	86
	huma onça de	2	85
	onças de agua	3	68
	de folhas de	4	58
	duas onças de	5	56
	de cada cousa	6	55
	cada cousa destas	7	54
	de agua de	8	49
	de oleo de	9	47
	de maneira que	10	45
<i>Arte</i>	que o enfermo	1	35
	se ha de	2	31
	que o medico	3	29
	se o enfermo	3	29
	se haõ de	5	25
	de fórma que	6	23
	se o medico	7	21
	ha de dar	8	19
	o enfermo naõ	9	16
	e se for	10	14

Observações menciona principalmente ingredientes para as receitas de remédios e alimentos para o doente, além de um objeto usado na preparação (*panela de barro*), depois há um sintoma (*dores de cabeça*). Lembrando que *Arte* tem a maioria de seus trigramas focados no *doente*, aqui também as listas se

aproximam do tema principal de cada obra. Além dos dez trios de palavras de cada tabela, há também o plural de *gema de ovo* em *Observações*, com 16 ocorrências (tabela 17).

Tab. 17. Trigramas mais frequentes de *Observações* e *Arte*

Obra	Type	Rank	Freq
<i>Observações</i>	panela de barro	21	35
	dores de cabeça	37	26
	agua benedicta vigorada	43	24
	agua da fonte	48	22
	gema de ovo	70	19
	agua de tanchagem	76	18
	leite de burra	84	17
	oleo de amendoas	84	17
	caldo de gallinha	97	16
	pós de quintilio	97	16
<i>Arte</i>	barrigas das pernas	25	9
	caldo de galinha	32	8
	cabeça do enfermo	53	6
	fluxo de sangue	53	6
	estomago do enfermo	85	5
	gadelha de lâ	85	5
	gema de ovo	85	5
	cama do enfermo	130	4
	saude do enfermo	130	4
	achaque no peito	203	3

Por fim, comparamos as frequências relativas desses trios na tabela 18. Os trios da *Arte* estão bem mais presentes nas *Observações*.

Tab. 18. Porcentagem da frequência dos trigramas das *Observações* e da *Arte*

Obra	Trigrama	% do total de tokens
<i>Observações</i> (147335 tokens)	panela de barro	0,024 %
	dores de cabeça	0,018 %
	agua benedicta vigorada	0,016 %
	agua da fonte	0,015 %
	gema de ovo + gemas de ovo	0,024 %
	agua de tanchagem	0,012 %
	leite de burra	0,012 %
	oleo de amendoas	0,012 %
	caldo de gallinha	0,011 %
pós de quintilio	0,011 %	
<i>Arte</i> (15837 tokens)	barrigas das pernas	0,057 %
	caldo de galinha	0,051 %
	cabeça do enfermo	0,038 %
	fluxo de sangue	0,038 %
	estomago do enfermo	0,032 %
	gadelha de lâ	0,032 %
	gema de ovo	0,032 %
	cama do enfermo	0,025 %
	saude do enfermo	0,025 %
achaque no peito	0,019 %	

6. TENTATIVA DE SÍNTESE DOS DADOS

Para identificar as palavras presentes na *Arte* e quais lhes são exclusivas frente a *Observações*, utilizamos *AntConc* e *Excel*. Com as listas de palavras da *Arte* e das *Observações* foi possível, no *Excel*, utilizar a função *Remove Duplicadas*, em Ferramentas de Dados no menu *Dados*, para obter os itens exclusivos e em comum. Essa função suprime, em uma mesma coluna, células repetidas. Assim, com as duas listas em uma mesma coluna e usando essa função, as células com itens em comum são excluídas da segunda lista. Ficam apenas os itens exclusivos.

A seguir, um trecho das listas da nossa planilha *Excel*, tanto em ordem de frequência como alfabética.

Types exclusivos		Types exclusivos (ordem alfabética)		Types em comum
<i>OBSER- VAÇÕES</i>	<i>ARTE</i>	<i>OBSER- VAÇÕES</i>	<i>ARTE</i>	AMBOS
15844 palavras	834 palavras	15844 palavras	834 palavras	1561 palavras
je	agoa	á	abanando	a
me	capitulo	aaa	abrigado	à
não	enfermeiro	aaquellas	abrigallo	ab
tres	sima	aarea	abrigará	abaixo
doente	untura	abafadiças	abrirá	aberta
tinha	pano	abafado	abstiver	abertos
pois	determinar	abafamos	abuta	abobora
fe	situação	aballa	acaso	abraza
fem	trez	aballados	acezas	abrazado
annos	fareis	aballar	acode	abrigada
meu	houver	aballo	actuação	abrir
efta	lançará	abanar	actuar	abundancia
oleo	veraão	abas	acudirá	acabada
jaão	conservar	abatida	adiante	acabado
avia	cordial	abbadeça	admitte	acabando
caufa	estará	abbadessa	admittir	acabar
doenças	applicareis	abJceJfos	adstringente	acção
Q	camiza	abdomem	advertência	accidental
fangué	mandareis	abdomen	advertências	accidente
Doentes	dano	abdómen	advertido	accidentes
Mandei	defensivo	abel	advertirá	acertado

Nas listas obtidas, ao final, às vezes, as palavras tidas como exclusivas são variações ortográficas ou flexões e conjugações, pois, como já frisado, nem *Ant-Conc* nem o *Excel* filtram variações. Por exemplo, *agua*, um dos itens mais frequentes de *Observações* é comum aos dois, apesar de ter sido usada apenas uma vez na *Arte*, no capítulo 6. Isso acontece porque o autor usa *agoa* – e essa forma não é usada nenhuma vez em *Observações*. O quantitativo dessas variações está na tabela 19.

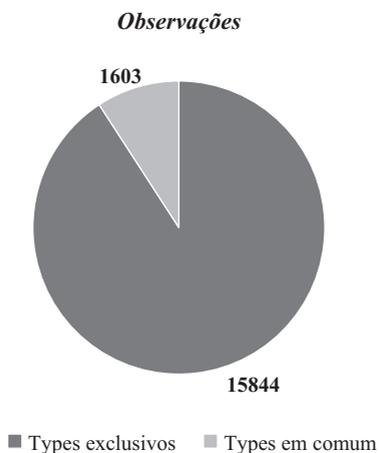
Tab. 19. Itens *agua* e *agoa* em *Observações* e *Arte*

Obra	Item	Rank	Freq	% do total de <i>tokens</i>
<i>Observações</i>	agua	33	505	0,34 %
<i>Arte</i>	agoa	29	73	0,46 %
	agua	1.132	1	0,006 %

Apesar disso, podemos ponderar algo sobre o repertório vocabular das duas obras. Por exemplo, 90,81 % dos 17447 *types* das *Observações* são itens exclusivos. Por outro lado, apenas 34,82 % dos 2395 *types* são próprios da *Arte*. Isso é o que tentam resumir a tabela 20, os gráfico 1 e gráfico 2 e a figura 2, que ilustra proporções. Mesmo com a margem de erro fruto de variação ortográfica, flexões e conjugações, ainda é uma diferença de valor alto. Isso mostra que a maioria dos assuntos tratados na *Arte* aparece de alguma forma também nas *Observações*, mas, frise-se, o contrário acontece pouco.

Tab. 20. *Tokens*, *types* e itens exclusivos de *Observações* e *Arte*

Obra	<i>Tokens</i>	Total de <i>types</i>	<i>Types</i> exclusivos	% de exclusivos com relação ao total
<i>Observações</i>	147335	17447	15844	90,81 %
<i>Arte</i>	15837	2395	834	34,82 %

**Gráfico 1.** Número de *types* exclusivos de *Observações* e *types* em comum.

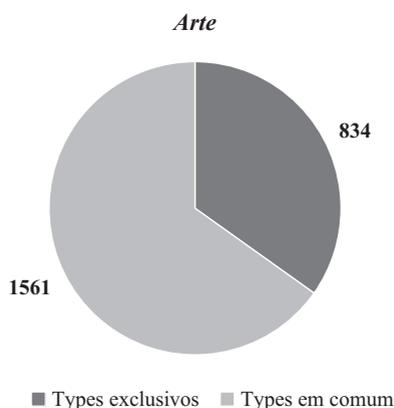


Gráfico 2. Proporção da quantidade de *types* exclusivos de *Arte*. Fonte: elaboração própria

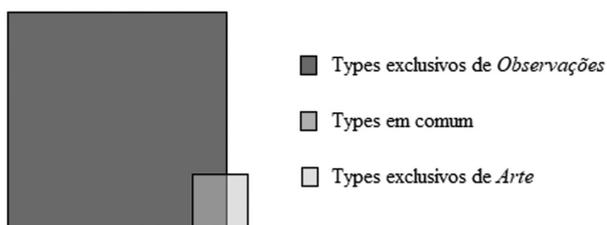


Fig. 2. Proporção de *types* exclusivos e comuns entre *Observações* e *Arte*.

Vários exemplos de itens que aparecem nas duas obras estão entre os mais frequentes em ambas, tais como *medico*, *enfermo* e *remedio*, conforme a tabela 21. Também encontramos, em ambas, referências a Ibn Sina (980–1037), conhecido como *Avicena*, médico e filósofo persa (Colgan 2009). Ele é citado apenas uma vez no Tratado II da *Arte*: «A muitos Medicos tenho ouvido dizer, que para melhor se conhecer a declinação da febre, he quando o calor desce aos pés, e ficaõ mais quentes, que as mãos, e que he doutrina de Avicena» (Santiago 1741: 137). Já em *Observações*, ele é mencionado 12 vezes: «Bem encomendaõ Galeno, e Avicenna que se não sangrem os meninos antes de terem quatorze annos» (Semedo 1707: 84).

Isso apontaria que ambos os autores tinham, provavelmente, apreço pelos trabalhos desse sábio da Idade Média. Entretanto, usam grafias diferentes para *Avicena*.

Tab. 21. Frequências de itens em comum

Obra	Item	Rank	Freq	% do total de tokens
<i>Observações</i>	medico	96	162	0,11 %
	medicos	53	287	0,19 %
<i>Arte</i>	medico	23	89	0,56 %
	medicos	113	21	0,13 %
<i>Observações</i>	enfermo	179	98	0,07 %
	enfermos	463	37	0,03 %
<i>Arte</i>	enfermo	7	276	1,74 %
	enfermos	33	62	0,39 %
<i>Observações</i>	remedio	36	456	0,31 %
	remedios	31	358	0,24 %
<i>Arte</i>	remedio	51	37	0,23 %
	remedios	64	31	0,2 %
<i>Observações</i>	avicenna	1.280	12	0,008 %
<i>Arte</i>	avicena	1.132	1	0,006 %

Por outro lado, há itens exclusivos de alta frequência. É o caso de *enfermeiro*, que apenas aparece na *Arte* e não ocorre nas *Observações*, apesar de o termo já ser atestado desde o século XIII (tabela 22) (Gonçalves 2020).

Tab. 22. enfermeiro(s) na Arte

Obra	Item	Rank	Frequência	% do total de tokens
<i>Arte</i>	enfermeiro	40	49	0,31 %
	enfermeiros	260	8	0,05 %

Outros itens exclusivos da *Arte* são *hospital*, *hospitaes* e *planxetas*. *Planxeta*, conforme Gonçalves (2020), é variante de *prancheta*, uma espécie de gaze usada em feridas (Houaiss 2001). Era material produzido, nesse caso, com fios de linho (*estopas finas*), com atestação pelo menos desde 1601, data da *Recopilação de Cirurgia*, de António da Cruz.

Quanto a palavras exclusivas das *Observações*, temos *morte* e itens relacionados: *morto*, *morta* e seus plurais. Isso pode ser consequência de apenas ter sido examinado o Tratado II da *Arte*. Cabe lembrar que somente a parte III trata de cuidados paliativos, incluindo tarefas burocráticas, como cuidar do testamento do enfermo. Além desses, encontramos em *Observações* menções a Galeno (129–216), médico e filósofo grego que seguia a teoria dos humores (Nutton 2005): «não achei outro mais eficaz (conforme diz Galeno,) que revelar o humor para a parte contrária» (Semedo 1707: 435).

Os dados relacionados a esses itens exclusivos estão na tabela 23.

Tab. 23. Itens exclusivos de *Arte* e *Observações*

Obra	Item	Rank	Freq	% do total de tokens
<i>Arte</i>	hospital	563	3	0,02 %
	hospitais	1132	1	0,006 %
	planxetas	563	3	0,02 %
<i>Observações</i>	morte	137	126	0,09 %
	mortes	3787	3	0,002 %
	morto	1494	10	0,006 %
	mortos	1280	12	0,008 %
	morta	3090	4	0,003 %
	mortas	1365	11	0,007 %
	galeno	198	89	0,06 %

Haver um mesmo item nas duas obras não significa que ele seja tratado da mesma forma ou com a mesma frequência. Um exemplo é *mulher*, que, escrito dessa forma, ocorre 123 vezes nas *Observações*, mas só duas vezes na parte em foco da *Arte* (na obra completa, são encontradas outras nove ocorrências do singular). O plural *mulheres* não parece no Tratado II da *Arte* (apesar de haver seis ocorrências fora dele), mas aparece 33 vezes nas *Observações*, conforme ilustra a tabela 24.

Quanto ao uso de *mulher(es)*, vale não esquecer que Santiago está em um hospital militar (Gonçalves 2020: 68), atendendo combatentes (homens). Por isso, talvez nunca mencione *mulheres* nessa parte, sem contar as restrições aos contatos com *mulheres*, postas no regramento da sua ordem religiosa.

No Tratado I, nesse sentido, ao mencionar *mulheres*, Santiago alerta do seu perigo aos noviços: «Fugi de comunicação com mulheres, ainda que sejaõ

velhas, que a paixão, e sensualidade da carne he fogo, e este achando matéria mais seca, mais facilmente se ateya» (Santiago 1741: 40). Nessa mesma parte, há outro indicativo sobre o papel da *mulher*: «Disse hum sabio, que a mulher em casa havia de estar cozendo, e fóra de casa na Igreja orando» (Santiago 1741: 29).

No Tratado II, nas apenas duas menções a *mulher*, ambas estão ligadas ao leite materno apresentado como remédio. Assim, temos *leite de mulher*, indicado como o melhor ingrediente para alguns tratamentos, como debelar dores de ouvido. Nesse contexto, o *leite de mulher preta* seria remédio mais proveitoso. E, em um *ranking* de qualidades, «esse leite humano seria seguido pelo *leite de burras*, depois o *leite de cabras negras* ou *de cabras ruivas*, o *leite de vacas* e, não havendo outro, recomenda-se o *leite de ovelhas*» (Santiago 1741: 156).

A ausência do «feminino» repercute também no fato de que, apesar de que *homem* tenha apenas uma menção e o plural, nenhuma vez há *enferma*. Nas *Observações*, *enferma* ocorre 72 vezes nos vários casos em que o médico descreve os tratamentos que prestou a várias mulheres. O item *doente*, por sua vez, é um dos mais frequentes das *Observações*, sendo usado 363 vezes no singular e 147 vezes no plural. Na *Arte*, ele não aparece, nem no plural. *Enfermo* e *enfermos* também aparecem nas *Observações*, mas com frequência um tanto menor, conforme resume a tabela 24.

Tab. 24. Frequências de *mulher(es)*, *homem(ns)*, *doente* e *enferma*

Obra	Item	Rank	Freq	% do total de tokens
<i>Observações</i>	mulher	139	123	0,08 %
	mulheres	519	33	0,02 %
<i>Arte</i>	mulher	750	2	0,01 %
<i>Observações</i>	homem	190	92	0,06 %
	homens	190	92	0,06 %
<i>Arte</i>	homem	1132	1	0,006 %
<i>Observações</i>	enferma	243	72	0,05 %
	doente	44	363	0,25 %
	doentes	113	147	0,1 %

Esses itens podem ser verificados também no nosso *corpus* pré-jornalístico, as *Gazetas Manuscritas* (tabela 25).

Tab. 25. *Gazetas Manuscritas: usos de mulher(es), enfermo, doente e medico*

<i>Type</i>	Freq	% do total de palavras
mulher	30	0,058 %
mulheres	2	0,004 %
enfermos	1	0,002 %
doentes	13	0,025 %
doente	10	0,019 %
medico	13	0,025 %
medicos	7	0,013 %

7. ALGUMAS NOTAS FINAIS

Há inúmeras considerações a serem feitas sobre o léxico e o vocabulário, incluindo os modos de dizer, de uma obra como a *Arte*, mesmo em uma mirada pouco exaustiva como esta, apoiada apenas em listagens, contagens e contrastes. Pelo que ficou demonstrado, um significativo conjunto de traços lexicais e de usos de palavras pode contribuir para que se consiga caracterizar a *Arte* como um gênero textual emergente que, além de ter um repertório vocabular em comum com uma obra médica, exhibe condições de uso de palavras, como a repetitividade, e escolhas que lhe são próprias.

Com os itens comuns e exclusivos à *Arte* e às *Observações*, além de um rol de obras de apoio para a interpretação do texto, vislumbram-se várias perspectivas. Poderíamos, por exemplo, selecionar elementos para encabeçar verbetes de um futuro glossário ou hiperdicionário de Medicina do século XVIII, contemplando tanto terminologias quanto palavras e, ainda, dados de outras fontes – antigas e modernas – com temática afim. A tabela 26 exemplifica itens partilhados e, entre esses, possíveis «candidatos a verbete». Como no dicionário histórico de Maroneze (2022), além de informações gramaticais e etimológicas, definições e notas, haveria os contextos de *Arte* e *Observações*, por exemplo, para «accidente», «achaque» ou «zaragatoa», espécie botânica usada como um «cotonete primitivo».

Tab. 26. Amostra de itens partilhados por *Arte e Observações*

Itens gramaticais	Léxico comum	Léxico especializado
ainda	abrazado	achaque
com	abundancia	ajuda
em	acabar	almorreimas
entaõ	acção	enfermidade
hum	accidente	fleuma
mas	affligir	lambedor
naõ	alivio	purga
ou	fonte	sangria
para	homem	suor
por	livrar	tibia
qual	noite	tosse
que	opiniaõ	untar
se	reflexaõ	vomitorio
sobre	setembro	xarope
tal	terra	zaragatoa

O léxico especializado tende a ser uma dificuldade para quem lida com obras como *Arte* ou *Observações*, supondo-se leitores sem formação na área da Saúde ou na área de Linguística Histórica, que hoje acessam textos antigos digitalizados. Assim, um recurso de apoio à leitura, seja dicionário, glossário ou vocabulário, parece ser algo importante. Todavia, apenas o estudo linguístico detalhado, realizado em diferentes frentes, do quantitativo ao qualitativo, poderia embasar uma tal empreitada.

BIBLIOGRAFIA

- Anthony, Laurence (2022): AntConc. Versão 4.0. Tóquio: <<https://www.lauranceanthony.net/software>> [consulta: 29-V-2023].
- Biderman, Maria Tereza Camargo (1967): Estatística linguística. Alfa, São Paulo, v. 11, p. 117– 128.
- Belmonte, Roberto Villar (2007): «A coesão textual frente à regra jornalística da não-repetição de palavras», *Blog do Villar*. Online: <<https://villar21.wordpress.com/2011/03/04/a-coesao-textual-frente-a-regra-jornalistica-da-nao-repeticao-de-palavras/>> [consulta: 29-V-2023].
- Berber Sardinha, Tony (2004): *Linguística de Corpus*. Barueri: Manole.

AQ: The following references are not listed in alphabetical order. Please check the reference order.
Biderman, Maria Tereza Camargo (1967), Belmonte, Roberto Villar (2007), Berber Sardinha, Tony (2004), Senedo, Joao Carvo (1707), Santiago, Fr. Diogo (1741)

- Colgan, Richard (2009): *Advice to the Young Physician: On the Art of Medicine*. New York: Springer Science & Business Media.
- Dury, Pascaline / Picton, Aurélie (2009): «Terminologie et diachronie: vers une réconciliation théorique et méthodologique?», *Revue française de linguistique appliquée*, 14, 31–41.
- Finatto, Maria José Bocorny / Quaresma, Paulo / Gonçalves, Maria Filomena (2018): «Portuguese Corpora of the 18th century: old Medicine texts for teaching and research activities». En: Fiser, Darja / Pancur, Andrej (eds.): *Proceedings of the Conference on Language Technologies & Digital Humanities (20. september-21. september, Faculty of Engineering, Ljubljana, University of Ljubljana, Slovenia)*, Ljubljana: University of Ljubljana, pp. 114–120. <https://www.sdjt.si/wp/wp-content/uploads/2018/09/JTDH-2018-Jose-et-al_Portuguese-Corpora-of-the-18th-century-old-Medicine-texts-for-teaching-and-research-activities.pdf> [consulta: 29-V-2023].
- Galves, Charlotte / Andrade, Aroldo Leal de / Faria, Pablo (2017): *Tycho Brahe Parsed Corpus of Historical Portuguese*. Online: <<https://www.tycho.iel.unicamp.br/~tycho/corpus>>. [consulta: 29-V-2023].
- Gonçalves, Maria Filomena (2020): «A Arte de Enfermeiros (1741): aspetos do léxico relativo a doenças e remédios no século XVIII», *Panace@*, 21 (52): 68–85.
- Houaiss, Antônio (2001): *Grande Dicionario Houaiss da Lengua Portuguesa*. Rio de Janeiro: Instituto Antônio Houaiss.
- Liberato, Yara / Fulgêncio, Lúcia (2007): *É possível facilitar a leitura: um guia para escrever bem*. São Paulo: Contexto.
- Maroneze, Bruno (coord.) (2022): *Dicionário Histórico de Termos da Biologia*. Online: <<https://dicbio.flch.usp.br/>> [consulta: 29-V-2023].
- Menezes, Francisco Xavier de (1673): *Gazetas manuscritas da Biblioteca de Évora*. Vol. I (1729–1731).
- NILC-Matrix (2021): <<http://fw.nilc.icmc.usp.br:23380/metrixdoc#honore>>. [consulta: 29-V-2023].
- Nutton, Vivian (1973): «The Chronology of Galen's Early Career», *Classical Quarterly*, 23 (1): 158–171.
- Nutton, Vivian (2005): «The fatal embrace: Galen and the history of ancient medicine». *Sci Context*, 18 (1): 111–121. doi: 10.1017/s0269889705000384. PMID: 16075496.
- Pellegrini, Jaqueline Rhod / Petkoff, Juliana Viegas (2009): «Estatística lexical em obras de Machado de Assis: um exercício exploratório». En: *Salão de Iniciação Científica UFRGS (21.: 2009 out. 19–23: Porto Alegre, RS)*. Livro de resumos. Porto Alegre: UFRGS.

- Quaresma, Paulo (2013): «Análise linguística de documentos da Biblioteca Pública de Évora Uma abordagem informática». En: Gonçalves, Maria Filomena / Banza, Ana Paula (coord.): *Património Textual e Humanidades Digitais. Da antiga à nova Filologia*. Évora: CIDEHUS, 139–155.
- Semedo, João Curvo (1707): *Observações medicas doutrinaes de cem casos gravissimos*. Lisboa: Na officina de Antonio Pedrozo Galram. <<https://sites.google.com/view/projeto38597/transcri%C3%A7%C3%B5es-de-observa%C3%A7oes-medicas-doutrinaes-1707>> [consulta: 25-V-2023].
- Santiago, Fr. Diogo (1741): *Postilla religiosa, e arte de enfermeiros, Guarneçada com eruditos conceitos de diversos Authores. facundos, Moraes, e Escriturarios*. Lisboa Occidental: Na officina de Miguel Manescal da Costa, Impressor do Santo Officio.