Universität
zu Köln

Universität zu Köln

HeidelGram

Diachronic Analysis of Grammatical Forms and Functions
in a Corpus of 16th- to 19th-Century English Grammar Books

ICAME43 | Cambridge, UK | Beatrix Busse, Nina Dumrukcic, Sophie Du Bois, Ingo Kleiber | July 2022

1. Headings as Structural Markers – Contemporary

Definition of Headings

"label the dominant topic or theme of the subsequent text" (Lorch, 1989, 210)

Function of Headings

- Modern purpose: Aesthetics, signposting, positioning and retrieval of information (Watson 2017, 2)

- Cognitive perspective: help readers identify major topics and ideas, emphasize structure of the book, serve as labels for sections within the text (Waller 1987)

Author vs. Publisher

- Modern, contemporary grammars (e.g., Quirk et al. 1985, see right): Content is the author's, not the publisher's or printer's decision, but the style is determined by the printers
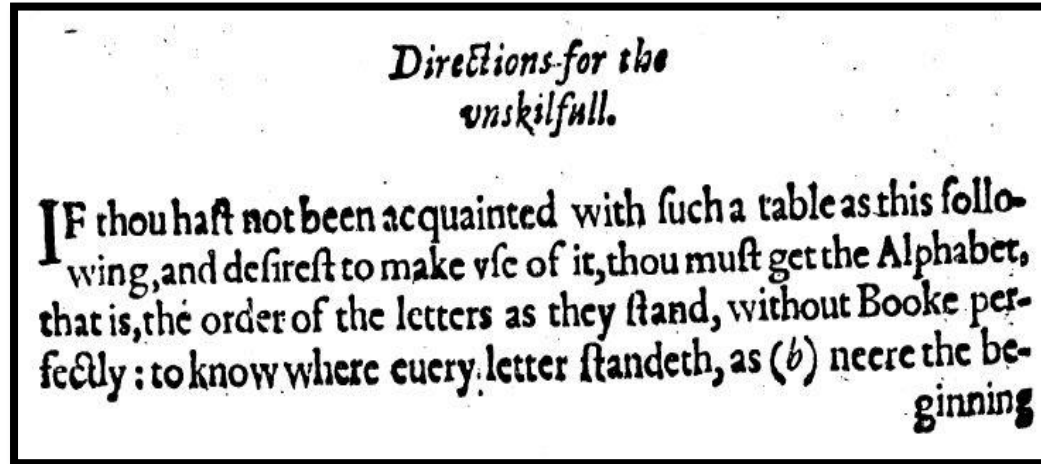
1. Examples of Genre-Specific Headings

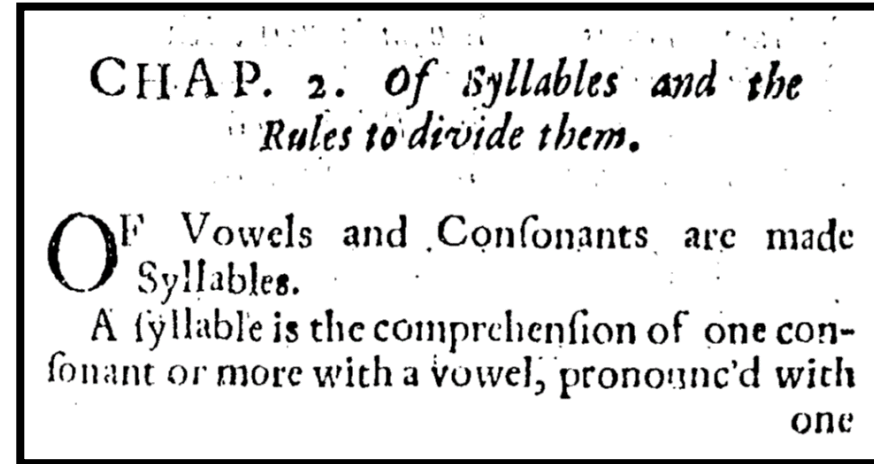**Fig. 1:** Excerpt from Coote's 1596 *English Schoole-Maister*, showing how headings stood out from the main body of text.

Fig. 2: Excerpt from Wharton's 1654 *The English Grammar*, showing the signpositng of a new chapter and its content.

2. Aims of this Paper

1. What is the function of headings in 16th- to 19th-century English grammars?

2. What grammatical phenomena are being signposted within the headings?

3. What form-to-function developments can be observed in grammatical terminology in diachrony?

Disclaimer: The findings shown are performed on a sample of the corpus. *(Pilot Study)*

2. Sub-Heading Corpus as Part of HeidelGram

Overarching Aims

- Compile and analyze a corpus of 16th- to 19th-century historical English grammars

- Innovatively combine corpus-linguistic and network-analytical approaches

- Investigating changing as well as stable discourses of historical English grammar writing

- Compilation and analysis of a 19th-century pilot corpus (Busse et al., 2018; 2019)
- Grammarians' reference categories (based on 19th-century data) (Busse et al., 2020)
- Building a network of 16th-century grammarians and references (Busse et al., 2021a)
- Fine-tuning language models for historical NER tasks (Busse et al., 2021b)

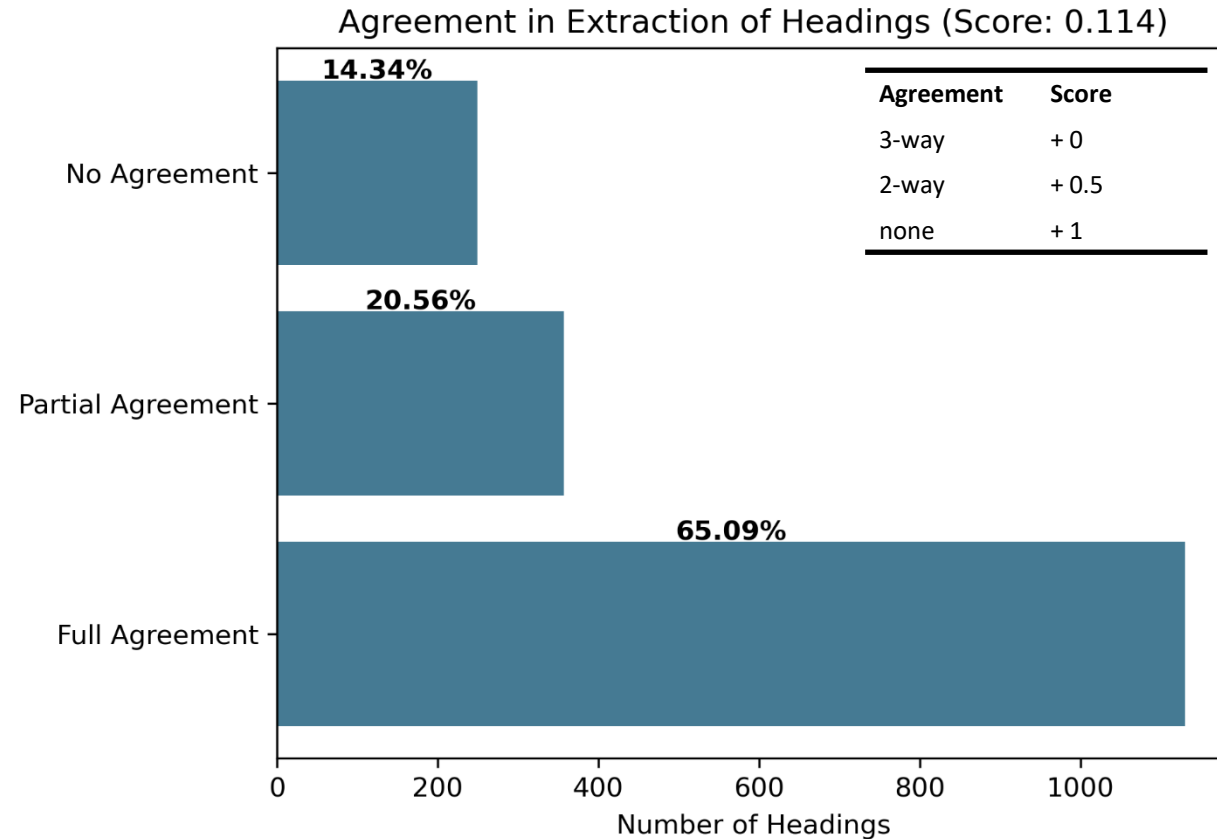| Network 1	Network 2	Network 3
Network of Grammars and Grammarians How grammar writers react to other authors	**Network of Verbal Hygiene Patterns** How grammar writers position themselves w. r. t. language	**Network of Grammatical Phenomena** How lexical inventory of grammar changed in diachrony

4. Data – A Common Understanding of Headings

- Sample of 953 pages taken out of the middle section of approx. 100 grammar books
- Three independent raters
- Fuzzy matching of items that accounts for variability and typos
- **Custom score** to determine IRR; neglecting probability due to a fuzzy set of choices

Agreement in Extraction of Headings (Score: 0.114)

Agreement	Score
3-way	+ 0
2-way	+ 0.5
none	+ 1

- No Agreement: **14.34%**
- Partial Agreement: **20.56%**
- Full Agreement: **65.09%**

Number of Headings (x-axis: 0, 200, 400, 600, 800, 1000)

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Universität zu Köln

4. Data

Author	Title	Year	Tokens	Headings
Sherry	A Treatise of the Figures of Grammar and Rhetorike	1577	215	37
Mulcaster	The First Part of the Elementarie	1582	494	78
Bullokar	Brief Grammar of English	1586	129	18
Meurier	The Conjugations in Englishe and Netherdutche	1586	149	51
Coote	The English Schoole-Maister	1596	371	60
Hume	Orthographie and Congruitie of the Britan Tongue	1617	227	57
Butler	English Grammar	1633	238	67
Wharton	The English Grammar	1654	311	72
Lewis	Plain & Short Rules	1675	43	7
Cooper	The English Teacher	1687	539	205
Mattaire	The English Grammar	1712	729	244
Barker	An English Grammar	1733	247	33
Kirkby	A New English Grammar	1746	356	86
Fenning	A New Grammar of the English Language	1771	702	263
Rhodes	A Concise English Grammar	1795	709	187
Earnshaw	The Grammatical Remembrancer	1817	301	106
Marcet	Mary's Grammar	1835	216	93
Morell	The Analysis of Sentences Explained and Systematised	1852	331	110
Abbott	English Lessons for English People	1871	1032	291
Nesfield	English Grammar: Past and Present	1898	2505	851

- Sample: five books per century $n_{books} = 20$

- For these 20 books, all headings are extracted $n_{headings} = 2916$

- Selection is made based on publication date, so as to sample from the beginning, middle, and end of each century

- Total numbers across all 20 books:
  Tokens for all the headings: 9844
  Types for all the headings: 2035

Universität zu Köln

5. Method

1. Verify definition of *headings* via IRR
2. Select a sample of 5 books per century
3. Manually extract headings from the sample
4. Perform corpus analyses using *Wmatrix* and *AntConc*
5. Qualitative evaluation of quantitative findings
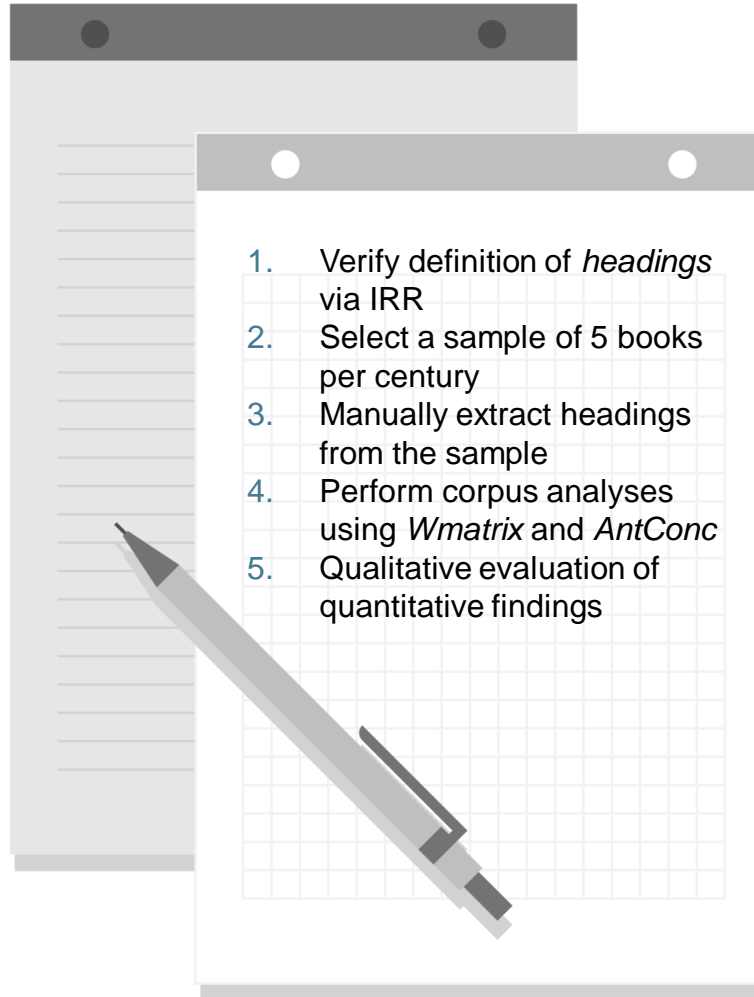
Wmatrix

- Identify salient parts of speech
- Identify salient semantic categories
- Part-of-Speech tagging is difficult considering the text type (headings + historical texts)

AntConc

- Frequency analyses
- Collocation analyses
- Concordances

Implications of Historical Grammar Data

- Grammatical bins highly frequent
- Historical spelling
- Shifts in meanings of words

Funded by

DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Universität zu Köln

6.1. General Frequencies

	16th century	**17th century**	**18th century**	**19th century**
Avg. nr. of words per heading	5.57	3.33	3.37	3.02
Avg. nr. of headings per book	48.8	81.6	162.6	290

Observations:

- There is a tendency for headings to get shorter over time.
- Later grammar books tend to use more headings to structure their text.

Literature:

- 19th-century grammars are allegedly more structured via chapters, headings, and paragraphs (Wolf 2011, 38)
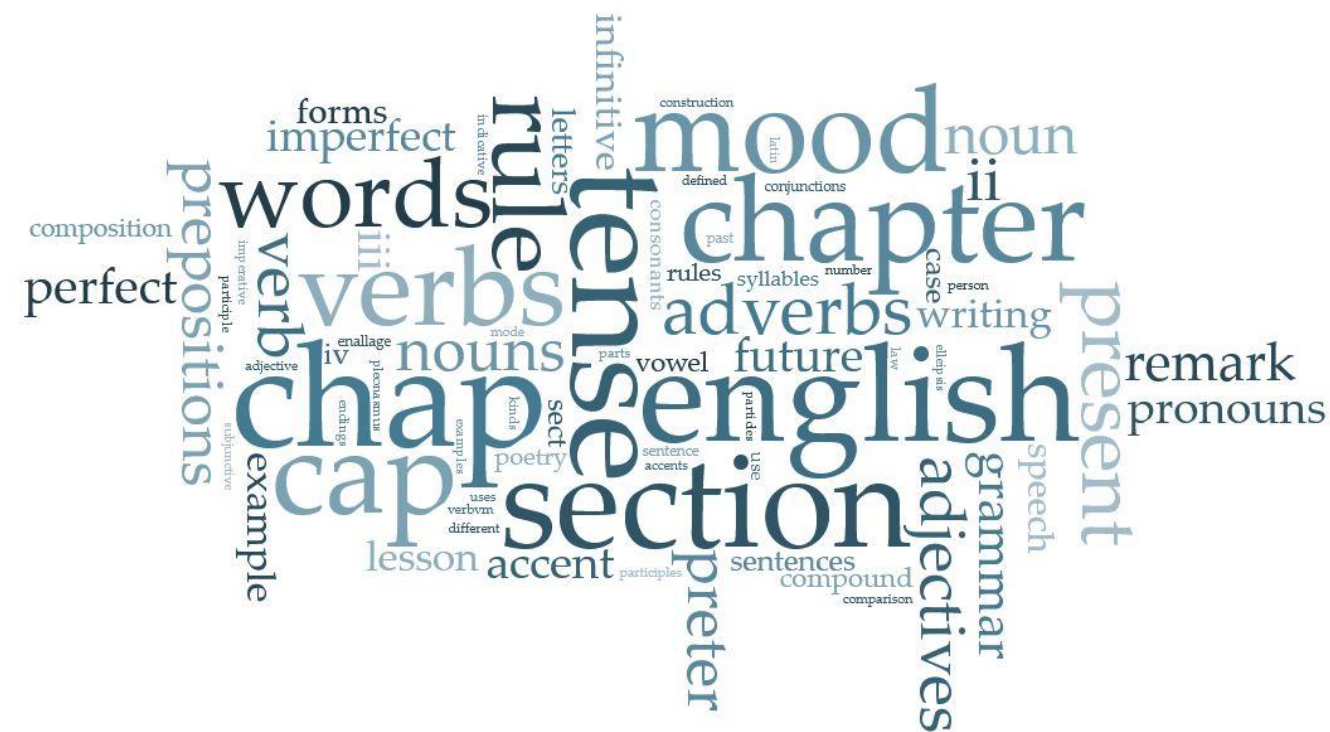
6.1. The Most Frequent Terms

Most frequent words in the corpus
(abs. freq.):

1. chap (102);
2. english (98);
3. tense (96);
4. section (93);
5. chapter (82)

Excluding structural terms
such as chapter, section, etc.
(abs. freq.):

1. english (98);
2. tense (96);
3. mood (70);
4. verbs (70);
5. rule (67)

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Universität zu Köln

6.1. The Most Frequent Terms – Examples

> 45. An Apostrophe is also used to distinguish the Possessives of appellative Names, and such proper Names as end in *s* from Plurals; as the *Tree's Height, Charles's Horse*: of which hereafter.
>
> ## C H A P. IV.
>
> *Of the Division of Syllables.*
>
> 1. WHEN a single Consonant (not accented) comes between two Vowels, it begins the latter Syllable; as in *a-mount, be-queath, na-ked*. But when such a Consonant bears the Accent (in my opinion) it ought to go with the former; as in *Hab-it, Drag-on, Wid-ow*.

Chapter heading, indicating start of new section

Sub-heading, indicating the topic of the next section

Fig. 1: Excerpt from Kirkby's 1746 *A New English Grammar*.

6.1. The Most Frequent Terms – Examples 2. **english** (98)

Numbered heading indicating the topic of the following subsection

Higher level heading, indicating the topic of the following section

125. Influence of Early English Poetry on the Elizabethan Writers.—The introduction of a fourth alliterated letter is a mark of lateness of date in early English poetry. This shows that the taste for alliteration did not vanish with the decay of alliterative poetry. It is true that the introduction of rhyme, supplying a different kind of poetic regularity, diminished the need of alliteration; but alliteration still clung even to rhyming poetry.
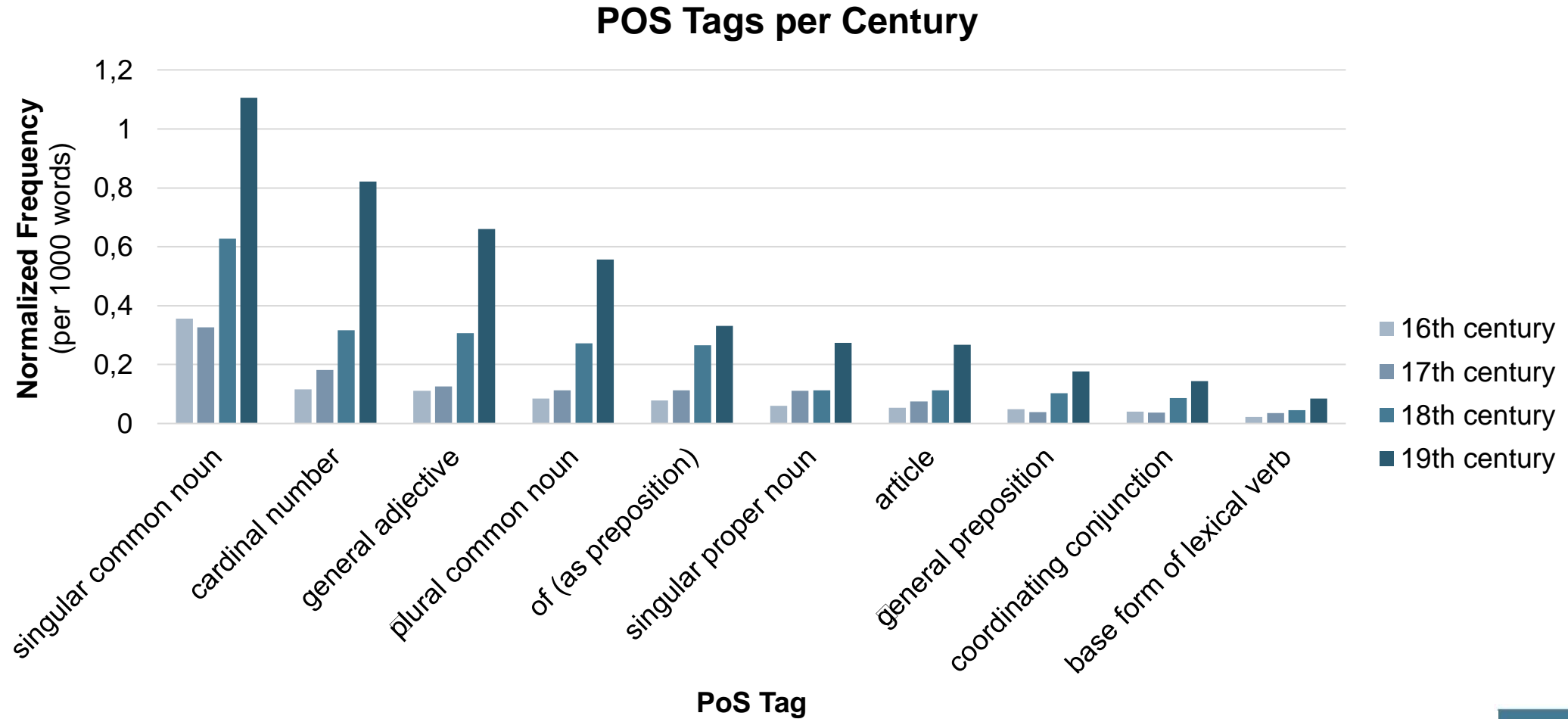
Fig. 1: Excerpt from Abbott's 1871 *English Lessons for English People.*

SECTION 3.—MIDDLE ENGLISH.

(*Approximate dates*, A.D. 1200-1500.)

388. Character of Middle English.—In its Middle period English went through three kinds of change :—
(*a*) The Mercian dialect, or Midland, as we should now call it, became eventually predominant in the place of the Wessex or Southern, which up to the Norman Conquest and for two centuries afterwards had held the first place.

Fig. 2: Excerpt from Nesfield's 1898 *English Grammar Past and Present.*

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Universität zu Köln

6.1. PoS Results per Century

POS Tags per Century

Normalized Frequency (per 1000 words) — vertical axis: 0, 0,2, 0,4, 0,6, 0,8, 1, 1,2

PoS Tag — horizontal axis: singular common noun, cardinal number, general adjective, plural common noun, of (as preposition), singular proper noun, article, general preposition, coordinating conjunction, base form of lexical verb

Legend: 16th century, 17th century, 18th century, 19th century

6.1. The Recurring Feature „Of the"

Clusters with "of"
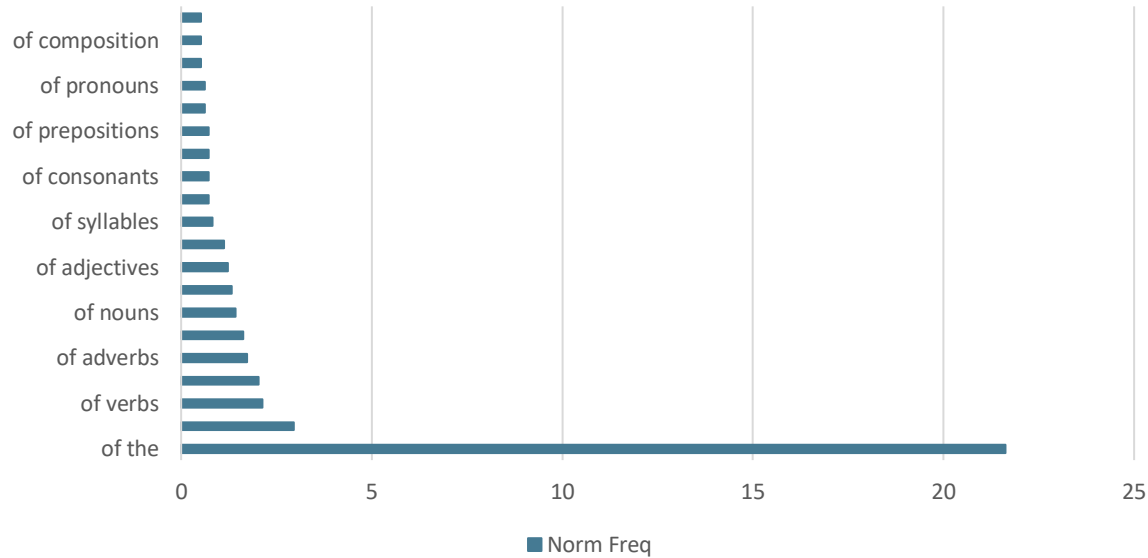
	16th century	17th century	18th century	19th century
OF THE				
Books	4 / 5	4 / 5	5 / 5	4 / 5
Abs. Freq.	25	59	73	63
Norm. Freq. (p. 1000 w.)	18.41	43.45	26.61	14.37

Sect. 7.	Of the	Vowel. (Cooper 1678)
General Rules.	Of the	Accidents of Sentences. (Cooper 1678)
CHAP. XIV.	Of the	Formation of a Verb Active. (Barker 1733)
Preposition. 12.	Of the	Interjection. (Earnshaw 1817)
ADDITIONAL REMARKS,	Of the	ARTICLE. (Fenning 1771)
…		

# 6.1. Function of Headings in Early Modern English Grammars

Observations

- *Of* as a preposition is salient across all centuries
  - Used to introduce a new section; on the topic of....
  - Recurring feature „Of the …"

- Numbers, particularly cardinal numbers were the most frequent
  - unsurprising considering most of the authors use a numeric section for the chapters/sections

- Some headings have a modern approach of conveying grammatical content, others are only used as structural markers (e.g., *Cap., Chap.* etc.)

- Function: introducing a new section and drawing attention to a feature.

- The headings get shorter but increase in number with a higher level of content lexical items = more concrete signposting.

Literature

- Typical page design of technical books: (Tebeaux 1991, 250)
  - Italicized headings
  - Numbered lists of topics and overviews before chapter headings

- ➢ intended to be used as reference manuals rather than for thorough, sustained reading

- Early Modern printers: made changes to spelling not sanctioned by the authors (Howard-Hill 2006, 16)

- Topical headings used for memorizing the contents of longer works (Engel 2022, 4)

6.2. Semantic Tag Results per Century

Semantic Tags per Century
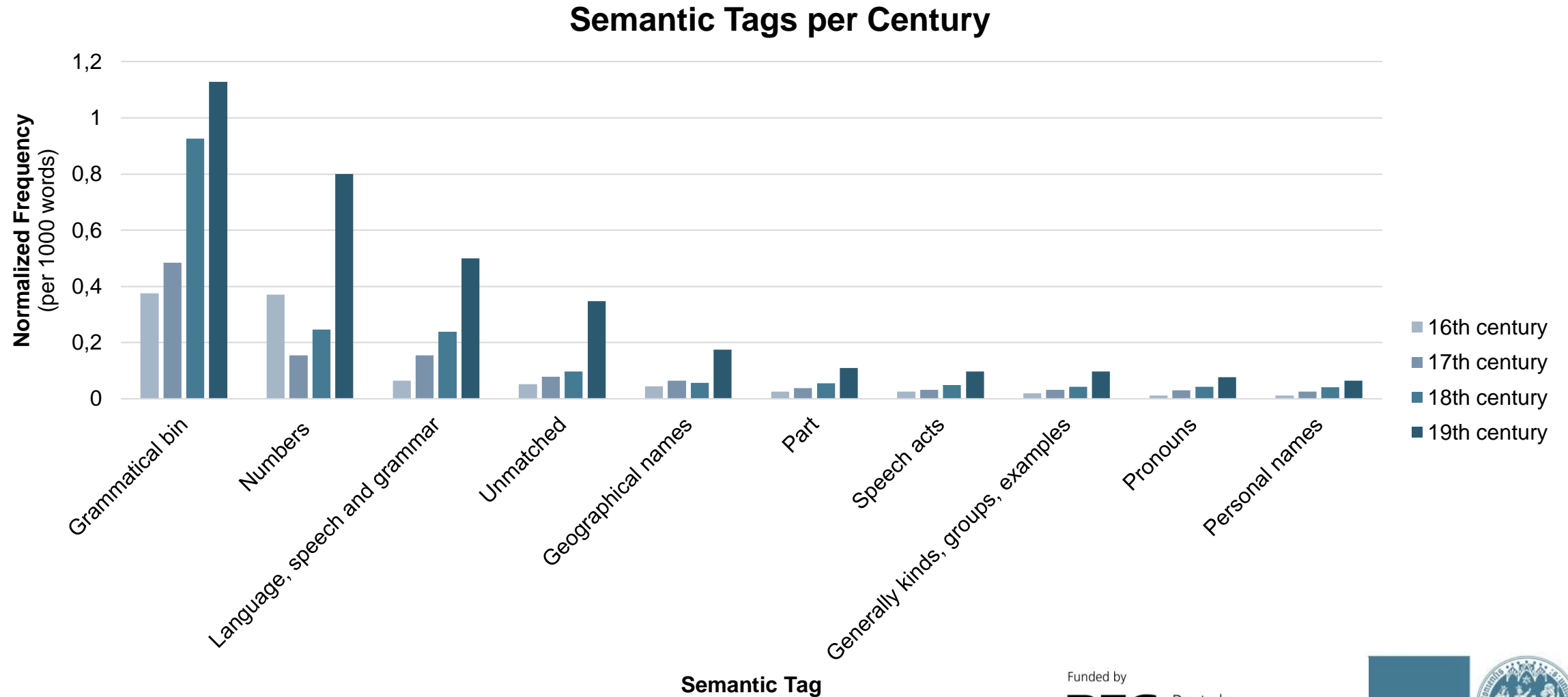
ICAME43 | Cambridge, UK 2022 | Beatrix Busse, Nina Dumrukcic, Sophie Du Bois, Ingo Kleiber

Funded by
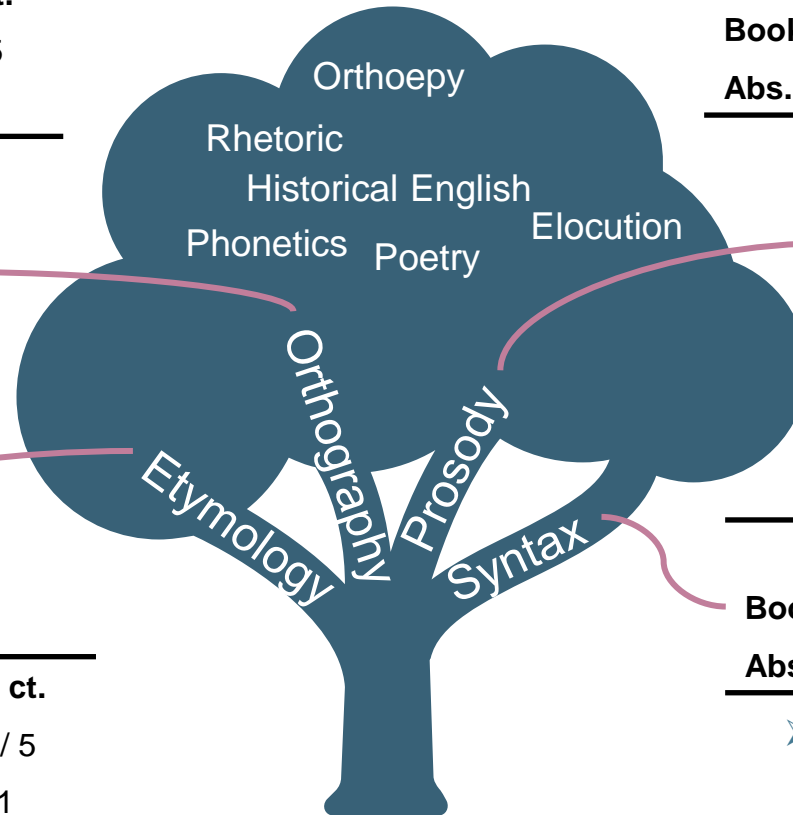DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Universität zu Köln

6.2. The Classical Branches of Grammar

Orthography, Orthographie, Orthographicis

	16th ct.	17th ct.	18th ct.	19th ct.
Books	1 / 5	1 / 5	1 / 5	1 / 5
Abs. Freq.	1	5	1	1

Prosody, Prosodia

	16th ct.	17th ct.	18th ct.	19th ct.
Books	1 / 5	0 / 5	2 / 5	2 / 5
Abs. Freq.	1	0	2	3

Orthoepy

Rhetoric

Historical English

Phonetics   Poetry   Elocution

Etymology   Orthography   Prosody   Syntax

Etymology, Etymological

	16th ct.	17th ct.	18th ct.	19th ct.
Books	1 / 5	0 / 5	2 / 5	1 / 5
Abs. Freq.	1	0	3	1

Syntax, Syntaxis, Syntactical

	16th ct.	17th ct.	18th ct.	19th ct.
Books	0 / 5	0 / 5	5 / 5	3 / 5
Abs. Freq.	0	0	9	5

➤ Categorization into „simple", „complex", and „compound" sentences: only in 19th ct.

Frequency data extracted using AntConc (Anthony 2022)
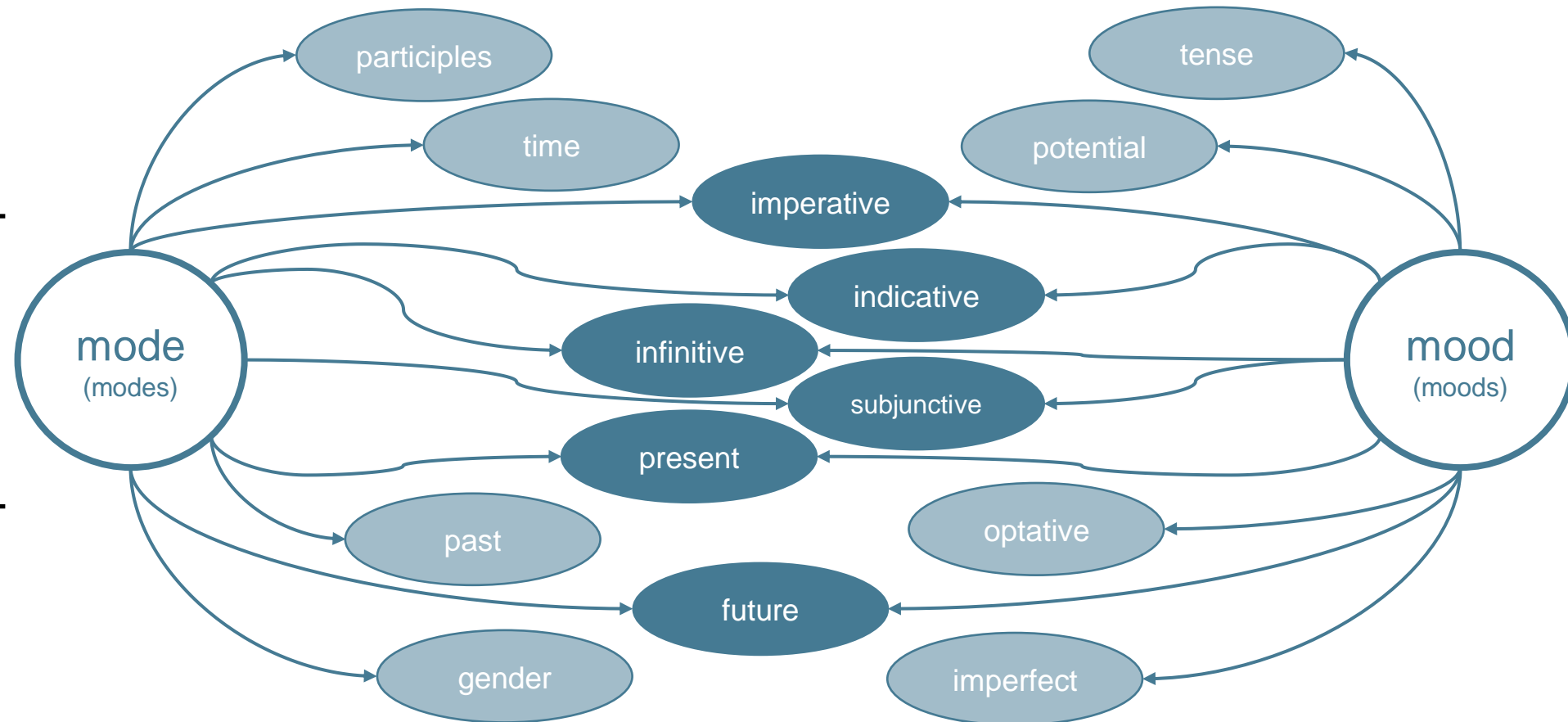
6.2. Grammatical Content from 1550 to 1900

Observations

- The frequency analyses of the main branches of grammar imply that the field of Syntax became more salient and more refined in the 18th and 19th century

- The other three main branches already occurred in earlier grammars

- Other fields of grammar (Orthoepy, Phonetics, etc.) occur sporadically in some grammars

- The chapters and sections were often numbered, particularly in the 16th and 19th century.

- There is a gradual increase in mentions of specific language/grammatical phenomena

Literature

- Fundamental components – with little variation – of English grammars between 1600 and 1900: Orthography, Etymology, Syntax and Prosody (Wolf 2011, 38)

- Expected shift of focus due to move from word-based to clause-based framework (Linn 2006, 76)

- Emerging disciplines may have started out „as subsidiary features in grammars" (Michael 1970, 197)

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Universität zu Köln

6.3. Collocates for „Mood" and „Mode" Respectively

16th ct.	0 / 5
17th ct.	0 / 5
18th ct.	1 / 5
19th ct.	2 / 5
Abs. Freq.	17

mode (modes)

participles
time
imperative
indicative
infinitive
subjunctive
present
past
future
gender
tense
potential
optative
imperfect

mood (moods)

16th ct.	1 / 5
17th ct.	3 / 5
18th ct.	4 / 5
19th ct.	2 / 5
Abs. Freq.	79

All data extracted using AntConc (Anthony 2022)

6.3. Collocates for „Tense" and „Time" Respectively
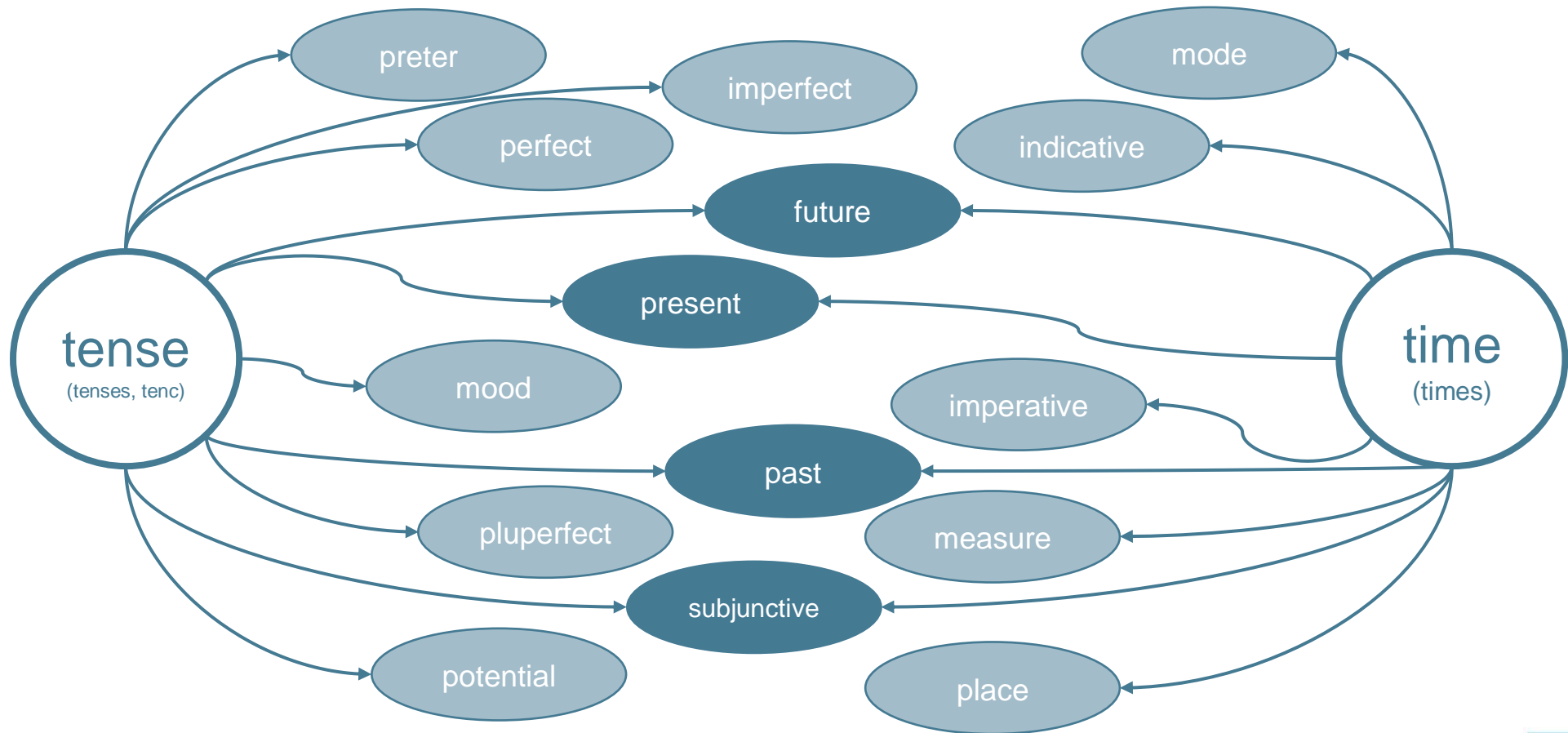
16th ct.	1 / 5
17th ct.	1 / 5
18th ct.	3 / 5
19th ct.	3 / 5
Abs. Freq.	109

tense
(tenses, tenc)

preter
imperfect
perfect
future
present
mood
past
imperative
pluperfect
measure
subjunctive
potential
place
mode
indicative

time
(times)

16th ct.	0 / 5
17th ct.	0 / 5
18th ct.	2 / 5
19th ct.	3 / 5
Abs. Freq.	15

All data extracted using AntConc (Anthony 2022)

Funded by

DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Universität zu Köln

# 6.3. Grammatical Forms and Functions in Diachrony

Observations

- The collocates imply that tense and mode are frequently discussed in close proximity of each other within the books

- The collocates imply that tense-time and mood-mode are different terms used for the same grammatical function

Literature:

- Early 18th century: „collective desire to describe and teach English on its own terms and not through Latin" (McCarthy 2020, 25)

  - Expectation: more ‚nativist' terminology, e.g,. helping verbs, rather than auxiliary verbs

- There is „the impression of a terminological flux" (Sundby et al. 1991, 6)

- Terminology is used in parallel or evolved over time, e.g., for the 18th century (Michael 1970, 185)

Funded by
DFG Deutsche Forschungsgemeinschaft
German Research Foundation

Universität zu Köln

7. Conclusion

- When determining boundaries between headings, we distinguished between headings and sub-headings due to their different functions.

- This pilot study provides insight into how the contents of the grammar books are signposted through the use of headings. We expect the full corpus will steady the tendencies we observed here.

- Once the entire sub-corpus of headings is compiled, we can gain a more holistic view of heading structure and content per century.

- We will also look at what particular grammatical phenomena 'deserve' a section heading and why, while others are mentioned within a paragraph and how this changes diachronically.

- Early observation of form-to-function change in grammatical terminology as demonstrated by the mood-mode and time-tense examples.

Universität zu Köln
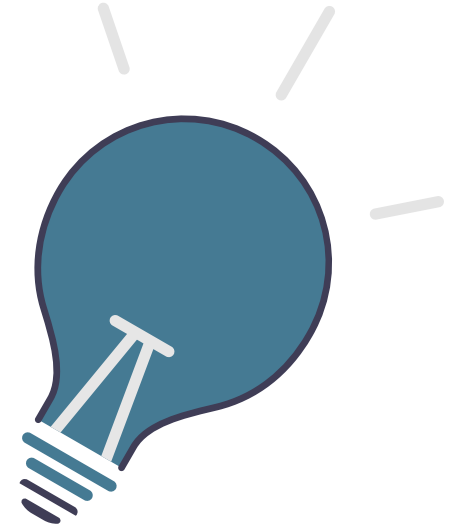
Thank you for your attention!

Bibliography I

Anthony, L. (2022). AntConc (Version 4.1.0) [Computer Software]. Tokyo, Japan: Waseda University. Available from https://www.laurenceanthony.net/software.

Bird, S., Loper, E, and Klein, E. (2009), *Natural Language Processing with Python*. O'Reilly Media Inc.

Busse, B., Gather, K., & Kleiber, I. (2018). Assessing the connections between English grammarians of the nineteenth century: A corpus-based network analysis. In E. Fuß, M. Konopka, B. Trawiński, & U. H. Waßner (Eds.), *Grammar and corpora 2016* (pp. 435-442). Heidelberg University Publishing.

Busse, B., Gather, K., & Kleiber, I. (2019). Paradigm shifts in 19th-century British grammar writing: A network of texts and authors. In B. Bös & C. Claridge (Eds.), *Norms and conventions in the history of English* (pp. 49-71). John Benjamins.

Busse, B., Gather, K., & Kleiber, I. (2020). A corpus-based analysis of grammarians' references in 19th-century British grammars. In A. Cermakova & M. Malá (Eds.), Diskursmuster - Discourse Patterns: Vol. 20. *Variation in time and space: Observing the world through corpora* (pp. 133-172). Mouton De Gruyter.

Busse, B., Kleiber, I., Dumrukcic, N., & Du Bois, S. (2021, July 13). *A corpus-based network analysis of 16th-century British grammar writing*. [Conference session]. Corpus Linguistics International Conference 2021, University of Limerick and Mary Immaculate College, Ireland.

Busse, B., Kleiber, I., Du Bois, S., & Dumrukcic, N. (2021, August 19). *Crossing the Boundary of Time: Fine-Tuning Modern NLP Models for Specialized Historical Corpus Data*. [Conference session]. ICAME42 - Crossing Boundaries Through Corpora, Technical University of Dortmund, Germany.

Engel, W. I. (2022). *The Printer as Author in Early Modern English Book History: John Day and the Fabrication of a Protestant Memory Art*. New York: Routledge.

Howard-Hill, T. H. (2006). *Early Modern Printers and the Standardization of English Spelling*. The Modern Language Review , 101(1), pp. 16-29.

Linn, A. (2006). English Grammar Writing. In B. Aarts & A. M. S. McMahon (Eds.), *Blackwell Handbooks in Linguistics. The Handbook of English Linguistics* (pp. 72–92). Malden: Blackwell.

Bibliography II

McCarthy, M. (2020). *Innovations and Challenges in Grammar*. *Innovations and Challenges in Applied Linguistics*. New York: Routledge.

Michael, I. (1970): *English Grammatical Categories: And the Tradition to 1800.* Cambridge: Cambridge University Press.

Rayson, P. (2009) *Wmatrix: A web-based corpus processing environment.* [Computer Software]. Computing Department, Lancaster University. Available from: http://ucrel.lancs.ac.uk/wmatrix/.

Rodríguez-Álvarez, A. (2016). "'Tis not so easy a matter to read well": Directions for Reading Aloud in Eighteenth-Century English Grammars. *Moderna Sprak*, *110*(2), 106–132.

Sinclair, S., and Rockwell, G. (2016): Voyant Tools. [Computer Software].

Sundby, B., Bjørge, A. K., and Haugland, K. E. (1991): *A Dictionary of English Normative Grammar 1700–1800.* Amsterdam: John Benjamins.

Tebeaux, E. (1991). Visual Language: The Development of Format and Page Design in English Renaissance Technical Writing. *Journal of Business and Technical Communication*, *5*(3), 246–274.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Pearson

Waller, R.H.W. (1987): *The Typographic Contribution to Language: Towards a Model of Typographic Genres and Their Underlying Structures*. Unpublished PhD thesis. University of Reading, Department of Typography and Graphic Communication.

Watson, R. (2017). A 'Heads Up' on Headings. *Nurse Author & Editor*, *27*(1), 1–7.

Wolf, G. (2011). *Englische Grammatikschreibung 1600-1900: Der Wandel einer Diskurstradition*. *Arbeiten zur Sprachanalyse: Vol. 54*. Frankfurt, M., Berlin, Bern, Bruxelles, New York, NY, Oxford, Wien: Lang.

Bibliography III (Grammar Books)

Abbott, E. A. (1871): *English Lessons for English People*. London: Seeley, Jackson, and Halliday.

Barker, I. (1733): *An English Grammar*. York: Tho. Gent

Bullokar, W. (1586): *Brief Grammar for English*. London: Edmund Bollifant.

Butler, C. (1633): *English Grammar*. Oxford: William Turner.

Cooper, C. (1687): *The English Teacher*. London: John Richardson.

Coote, E. (1596): *The English Schoole-Maister*. London: Widow Orwin.

Earnshaw, C. (1817): *The Grammatical Remembrances*. London: Thomas Smart.

Fenning, D. (1771): *A New Grammar of the English Language*. London: Pater-Noster-Row

Hume, A. (1617): *Orthographie and Congruitie of the Britan Tongue*. London: Trübner & Co.

Kirkby, J. (1746): *A New English Grammar*. London: Ludgate-bill.

Lewis, M. (1675): *Plain & Short Rules*.

Maittaire, M. (1712): *The English Grammar*. London: W. B.

Marcet, J. (1835): *Mary's Grammar*. London.

Meurier, G. (1586): *The Coniugations in Englishe and Netherdutche*. Leyden: Thomas Basson.

Morell, J. D. (1852): *The Analysis of Sentences Explained and Systematised*. London: Robert Theobald.

Mulcaster, R. (1582): *The First Part of the Elementarie Which Entreateth Chefelie of the Right Writing of our English Tung*. London: Thomas Vautroullier.

Nesfield, J. C. (1898): *English Grammar: Past and Present*. London: MacMillan & Co.

Rhodes, B. (1795): *A Concise English Grammar*. Birmingham: J. Belcher.

Sherry, R. (1577): *A Treatise of the Figures of Grammer and Rhetorike*. London: Ricardi Totteli.

Wharton, J. (1654): *The English Grammar*. London: John Walker.