Stefan Hartmann*, Nils Hansson and Adrian Loerbroks

The language of gratitude: An empirical analysis of acknowledgments in German medical dissertations

https://doi.org/10.1515/gcla-2024-0009

Abstract: Acknowledgment sections are a rich but underused resource for understanding how language is used for social purposes (such as expressing gratitude and communicating social relations networks), and how conventions and patterns emerge in this process. This paper presents a usage-based case study combining qualitative and quantitative methods for analyzing a dataset of >300 acknowledgment sections from medical dissertations written in German. In our quantitative analysis, we gauge keywords and key n-grams and assess the relative position of recurrent words in each text. Our analysis shows that this text type has developed clear conventions, with acknowledgments in the professional domain being followed by a usually smaller set of expressions of gratitude associated with the private domain. In addition, our quantitative analysis suggests recurrent patterns that can be linked to specific socio-pragmatic functions. For instance, an analysis of n-grams attested in text segments associated with the professional vs. the private domain shows some differences with regard to the typical patterns chosen in those segments. Our analysis also raises a number of future research questions, thus showing that acknowledgment sections are a highly interesting object of study that deserve to be investigated in more detail.

Keywords: Acknowledgments, formulaic language, corpus linguistics, medicine

1 Introduction

Acknowledgments are ubiquitous in academia. In many disciplines, academic work is usually a collaborative process and one that involves not only the proverbial

* **Corresponding author: Stefan Hartmann,** Faculty of Arts and Humanities, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, hartmast@hhu.de
Nils Hansson, Department for the History, Philosophy, and Ethics of Medicine Medical Faculty, Heinrich Heine University Düsseldorf, Nils.Hansson@hhu.de and Universitätsklinikum Düsseldorf, Düsseldorf, Germany, Nils.Hansson@uni-duesseldorf.de
Adrian Loerbroks, Institute of Occupational, Social and Environmental Medicine, Medical Faculty and University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, adrian.loerbroks@uni-duesseldorf.de

standing on the shoulder of giants, but also, and perhaps even more crucially, being backed by peers. This is why many publications from papers to monographs feature acknowledgments sections. Previous work has already emphasized that acknowledgments are a rich resource for addressing a number of meta-scientific questions (e.g. Ben-Ari 1987; Hyland 2003; Sanderson 2005; Brown 2009; Caesar 1992; Cronin et al. 1992, 1993, Cronin & Overfelt 1994; Giannoni 2002; Mantai & Downing 2015; Micciche 2017; Desrochers et al. 2017). This holds especially true for doctoral thesis acknowledgments (see e.g. Hyland 2003, 2004 and especially Wesian 2015). After all, a doctoral thesis is the result of a long process that usually significantly shapes this stage of the doctoral candidate's life. As such, acknowledgments can be highly informative about a variety of issues: the social support networks of doctoral candidates, including formal and informal ways of different types of support (e.g. emotional, instrumental, informational); the institutional and personal contexts that have been relevant in the PhD phase; and, perhaps most relevant from a linguistic point of view, the linguistic encoding of expressions of gratitude. As Hyland (2003: 265) puts it, acknowledgments "bridge the personal and the public, the social and the professional, and the academic and the lay."

From a cognitive-linguistic perspective, it is the interaction of these aspects that makes thesis acknowledgments an interesting resource. Cognitive linguistics and its currently most popular subfield, Construction Grammar, have experienced a 'social turn' in recent years, emphasizing that language can only be fully understood if we take sociocultural contexts into account (Croft 2009, Schmid 2016; for recent overviews, see Hoffmann 2022, Ungerer & Hartmann 2023; Morin et al. 2024). At the same time, researchers like Stumpf (2015) or Mellado Blanco et al. (2022) have started to combine phraseology and Construction Grammar, which is also highly relevant for our topic because thesis acknowledgments can be assumed to consist of formulaic patterns to a considerable extent. In a similar vein, many construction grammarians have turned their attention to text linguistics, analyzing textual conventions and patterns from a constructionist point of view (e.g. Della Constanza 2013, Ziem 2013, Merten 2018). The present paper adds to these lines of research.

Drawing on a corpus of contemporary doctoral dissertations, our goal is to gauge the recurrent linguistic patterns that are used in thesis acknowledgments, to examine who is thanked for what in these data, and how these two aspects interact, i.e. whether there are formulaic patterns that are used more often for some types of acknowledgments than for others (e.g., for acknowledgments in the professional vs. in the private domain). For the present pilot study, we decided to work with a small sample of doctoral dissertations from the Faculty of Medicine at Heinrich Heine University Düsseldorf that were published in 2022. This enables an in-depth study of a dataset that can be expected to be fairly homogeneous, as all theses taken into account were defended at the same faculty. We combine a quantitative analysis of

the full sample with a qualitative analysis of a small subsample. This allows for a proof-of-concept study demonstrating that a combination of quantitative computational models and qualitative content analysis using inductive annotation categories is a promising approach for gaining a deeper understanding of the structure and function of acknowledgment sections not only in medical dissertations.

Focusing on medical dissertations can be insightful because they stand out compared to other doctoral dissertations in several ways. For one thing, obtaining a doctoral degree is much more common in human medicine than in other disciplines. Students of medicine account for a share of 6.88 percent of all students in Germany, but no less than 26 percent of all doctoral students in German work in the domain of medicine.[1] For another, the duration of the doctorate is comparatively short – according to calculations of the German Research Foundation, it usually takes less than a year.[2] However, it should be pointed out that in many cases, candidates start the work on their medical dissertation already during their undergraduate studies, which makes it hard to determine the actual duration of this phase. Also, the European Research Council does not recognize the MD, or the equivalent German degrees Dr. med. and Dr. med. dent., as a research doctorate.[3] In general, work on a medical dissertation can be expected to be much more closely tied to a specific working group than is usually the case in other disciplines. In addition, medical faculties in Germany tend to have a fairly hierarchical structure, which is arguably the case to a lesser extent e.g. in arts and humanities faculties. It seems plausible to expect that these contextual factors may have at least some ramifications regarding the structure and content of acknowledgment sections. The characteristics that set apart medical doctorates from PhD studies in other domains may also be one reason why acknowledgment sections have been shown to emerge very early on as a text type in its own right especially in dissertations from the domain of human medicine (Wesian 2015: 212, Hansson et al 2019).

The remainder of this paper is structured as follows: In Section 2, we review previous literature on acknowledgments, focusing on their hypothesized functions as well as on their text type characteristics. Section 3 presents our own case study, with Section 3.1 focusing on the quantitative analysis and Section 3.2 zooming in on a small subsample that is subjected to an in-depth qualitative analysis. Section 4 summarizes the main results and discusses avenues for future research.

1 See https://www.hsi-monitor.de/themen/internationale-studierende/studierende-grunddaten/verteilung-studierende-nach-fach/ (last accessed 11/07/2024)

2 See e.g. https://www.wissenschaftsrat.de/download/2023/1196-23.pdf?__blob=publicationFile&v=16 (last accessed 07/04/2024)

3 See e.g. https://erc.europa.eu/sites/default/files/document/file/ERC_policy_on_PhD_and_equivalent_doctoral_degrees_2016.pdf (last accessed 07/04/2024)

2 State of the art

Acknowledgments in general and dissertation acknowledgments in particular have been studied from a variety of perspectives and in a number of different disciplines. Both their typical form and their function(s) have been subject to scientific scrutiny, although linguistic studies have been relatively rare so far.

In two seminal studies, Hyland (2003, 2004) analyzed 240 MA and PhD dissertations written by non-native speakers of English in Hong Kong. Hyland (2003) defined the persons and groups that are acknowledged in different categories, distinguishing academics, friends, family, and others; the first group, academics, was the largest one across six disciplines he investigated, although there are some differences between them (e.g. in computer science, friends and family together make up almost 50% of the acknowledgments, much more than in the other disciplines). Hyland (2004) showed that acknowledgment sections often follow a similar pattern, consisting of (i) a reflecting move that comments on the writer's research experience, (ii) a thanking move in which the writer gives credit to individuals and institutions, and (iii) an announcing move, in which the author accepts responsibility for flaws or errors (Hyland 2004: 308–309).

Mantai & Downing (2015) analyzed acknowledgments from 79 PhD dissertations from Australian universities, focusing on the types of support provided and on the question of who (e.g., supervisors) and what (e.g., institutions) is acknowledged as providing support. They show that three broad types of support – academic, emotional and instrumental – can be distinguished, and that social support emerges as the main category from the data they analyze.

Mikołajczyk (2013), based on a corpus of German PhD dissertation acknowledgments that is unfortunately not described in detail (i.e. it remains unclear how many texts from which disciplines she took into account) investigates how the (main) supervisor of a thesis is referred to in acknowledgment sections, and distinguishes five types: 1. mentioning the social role of the supervisor, an address form, titles, and first as well as last name, e.g. *meine Doktormutter Frau Prof. Dr. XY* 'my supervisor (lit. doctor-mother) Ms. Prof. Dr. XY', 2. social role + first and last names without address form and titles, e.g. *meine Doktormutter XY*, 3. the function of the supervisor + address form + titles + first and last name, e.g. *die Lehrstuhlinhaberin Frau Prof. XY* 'the chair Ms. Prof. XY', 4. only address form + titles + first and last name, e.g. *Frau Prof. Dr. XY*, 5. only first and last name. The first and the fourth type are the most frequent ones in her data; she interprets the fact that variants in which both address forms and titles are used as an expression of the social hierarchy between the supervisor and the PhD candidate (Mikołajczyk 2013: 87).

Koller (2001) proposes a general framework for analyzing acknowledgments from a stylistic and text-linguistic perspective. Conceding that acknowledgment sections are hardly ever "creative", Koller (2001: 289) still observes considerable

variation on the lexical, syntagmatic, stylistic, and textual level. For instance, there are different alternative formulaic patterns that are typically used, e.g. *jemandem Dank schulden* 'owe thanks to somebody', *Dank gebührt* 'thanks are due to', or *meine Dankbarkeit gilt* 'I am grateful to' (lit. roughly 'my gratitude applies to'; Koller 2001: 291). Also, the expression of gratitude can be qualified in various ways, e.g. by stating the degree of gratitude, e.g. with qualifiers such as *herzlich(st)* '(most) heartfelt'.

The most in-depth study of acknowledgment sections in (German) dissertations to date has been conducted by Wesian (2015). Using a diachronic sample of PhD dissertation acknowledgment sections from four disciplines – German studies, medicine, engineering, and law – from 1897 to 2015, she focuses on the emergence of the acknowledgment section as a text type in its own right. Importantly, Wesian distinguishes two functions of dissertation acknowledgments, a social and a communicative one. Drawing on Brown & Levinson's (1987) seminal work on politeness, she argues that dissertation acknowledgments usually display forms of positive politeness, i.e. strategies targeted at the addressee's positive face, with 'positive face' being defined as "the positive consistent self-image or 'personality' (crucially including the desire that this self-image be appreciated and approved of) claimed by interactants" (Brown & Levinson 1987: 61). This is part of the social function of acknowledgments. Regarding the communicative function of acknowledgments, Wesian (2015: 206) argues that acknowledgments bring together different 'textual acts' (in analogy to 'speech acts'). Working with an open-ended set of textual acts, she conducts a qualitative analysis of her sample to gauge the dominant act in each acknowledgment section. Her results show that expressing gratitude is, unsurprisingly, the dominant act in most but not all of the texts she analyzes. In particular, older texts in her sample tend to combine the expression of gratitude with other elements (which is why the sections are usually not labeled *Danksagung* 'acknowledgments' but rather *Vorwort* 'preface'). As for the function of expressing gratitude, she distinguishes between explicit and implicit acts of expressing gratitude (Wesian 2015: 214). In the former case, an explicit expression of gratitude is used, e.g. *I would like to thank my supervisor*. In the latter case, only the action that is acknowledged is mentioned, without an explicit expression of gratitude, as in (1).

(1) Claudia Müller vom Institut für Pathologie der RWTH Aachen half mir beim Erlernen der Methodik. (quoted from Wesian 2015: 214)

'Claudia Müller from the Institute of Pathology at RWTH Aachen helped me to learn the methodology.'

In a similar vein, Jakobs (1997: 22) points out that in addition to the main function of expressing gratitude and making transparent the contribution of others to the current work, acknowledgment sections can fulfill additional functions – in particular, the author can signal their position in the scientific community. As such,

acknowledgments can also serve the purpose of self-valorization. Koller (2001: 287) also emphasizes this point and cites Genette's (1997: 238) tongue-in-cheek remark that "an author who has so many friends of both sexes cannot be completely bad."

Turning from functional to formal aspects of dissertation acknowledgments, Wesian (2015: 190) shows that they typically display three linguistic features: use of modal verbs, omission of the acting subject, and a tendency to avoid the first person singular personal pronoun at the beginning of a sentence. These features can be seen in examples (2) and (3) from our dataset (introduced in more detail in Section 3).[4] In (2), the verb *danken* 'to thank' is fronted to avoid the pronoun *ich* 'I' at the beginning of the sentence (*Danken möchte ich* instead of *Ich möchte danken* 'I would like to thank'), and while the acting subject is not omitted here, it is backgrounded by the use of the modal verb *mögen*. In the first sentence of (3), by contrast, the acting subject is indeed omitted by using the phrase *Mein Dank gilt*, lit. 'my thank applies/pertains to', roughly '(my) thanks are due to...'. The examples also show that these are of course just tendencies that do not always apply – for instance, the second sentence of (3) has the personal pronoun *ich* in sentence-initial position.

(2) **Danken** möchte **ich** ebenfalls Herrn Prof. Dr. med. [NAME], der die Zweitbetreuung übernommen und mich nicht nur während meiner Doktorarbeit stets unterstützt, sondern mir im Rahmen des Mentoring-Programmes der Uni Düsseldorf zum Studium generell stets Hilfestellung gegeben hat. (229)

'**I** would also like to **thank** Prof. Dr. med. [NAME], who has agreed to be my second supervisor and has not only continuously supported me during the time of my thesis but has also always offered me support more generally through the mentoring programme of the University of Düsseldorf.'

(3) **Mein** größter Dank gilt Herrn Prof. Dr. med. [NAME] für eine Betreuung, die nicht besser hätte sein können. Ich danke Ihnen für die Überlassung des Themas, für jedes bereichernde Gespräch, für all Ihre Hilfestellungen und Denkanstöße und für alles Menschliche, was Sie mir darüber hinaus noch auf meinen Weg mitgegeben haben. (229)

'**My** greatest thank is due to Prof. Dr. med. [NAME] for a supervision that could not have been any better. I would like to thank you for assigning this topic to me, for each rewarding conversation, for all your help and thought-provoking impulses and for all the human/interpersonal things that you have given to me on my way.'

4 We have consecutively numbered the acknowledgments (in random order), the numbers in brackets after each example indicate the file number. Despite the fact that all acknowledgment sections are publicly available, we chose to anonymize the segments for privacy reasons.

On a more abstract and theoretical level, the expression of gratitude has also been discussed in the literature on linguistic politeness. From classics like Bühler (1934) and Jakobson (1960) to present-day approaches in fields such as pragmatics and interactional linguistics, it has been acknowledged that language is an inherently social phenomenon and that, as such, linguistic utterances carry more than just propositional meaning. Watts (2003: 173) adopts Halliday's (1978) distinction between ideational and interpersonal meaning. The former is equivalent to propositional meaning, i.e. it pertains to propositions that can be captured in truth-conditional terms, while the latter pertains to the 'emotive' (Jakobson) or 'expressive' (Bühler) functions of language that usually cannot be expressed in terms of truth conditions. Expressions of thanks convey interpersonal and what Watts (2003: 174) calls procedural meaning, which pertains to "sets of procedures through which propositional meaning can be derived" and includes text-/discourse-deictic functions as well as aspects of meaning that are "pertinent to the relationship between the speakers and/or between the speakers and the context of the utterance." Importantly, all languages have – often formulaic – linguistic expressions that have come to express procedural meaning through processes of pragmaticalization (Watts 2003: 180). Formulaic expressions of gratitude are paradigm examples of such expressions of procedural meaning, or EMPs.

Aijmer (1996: 37) distinguishes thanking strategies on two levels: firstly, implicit and explicit expressions of gratitude can be distinguished; both implicit and explicit expressions of gratitude can be framed in an emotional or non-emotional way. For example, acknowledging a debt would be a non-emotional way of explicitly expressing gratitude; expressing appreciation of the addressee and/or their action, e.g. by saying *that's lovely!*, is an example of an emotional, but explicit way of expressing gratitude. Importantly, the explicit, non-emotional strategy of "acknowledging a debt of gratitude" is, according to Aijmer (1996: 38), "restricted to writing and to certain situations such as thanking one's teacher or family in the preface of academic works."

From a cognitive-linguistic perspective, variation within and between acknowledgment sections in the ways gratitude is expressed is particularly interesting as it touches upon the key question of construal (see e.g. Langacker 2008: 55–89), i.e. how the conceptual content conveyed by a linguistic expression is framed and perspectivized by choosing a particular way of phrasing it over potential alternative ones. Importantly, differences in construal do not just pertain to aspects such as which entities are foregrounded or backgrounded, or from which vantage point an entity or event is conceptualized (see e.g. Ungerer & Schmid 2006 and Evans 2019 for comprehensive discussions of the concept of construal) but also to the social domain (see e.g. Hart 2014). Cognitive linguistics and Construction Grammar have recently started to incorporate ideas from third-wave sociolinguistics, which sees language users "not as passive and stable carriers of dialect, but as stylistic agents,

tailoring linguistic styles in ongoing and lifelong projects of self-construction and differentiation." (Eckert 2012: 96–97) In other words, language users can – and do – change their linguistic performance to express different sociolinguistic identities (Hoffmann 2022: 242). In a similar vein, Schmid (2020: 40) points out that self-presentation and positioning can be an important source of grammatical structure (also see Hyland 2003). This leads to the prediction that expressions of gratitude relating to the professional vs. the private domain in acknowledgment sections may differ considerably, as the candidates essentially perform their different identities – as academics, employees, family members, friends – and bring them together in a short, condensed text. This is another aspect that makes acknowledgments so interesting: they can be seen "not simply as reflections of the doctoral experience, but as performances of the writers' desired identity" (Mantai & Dowling 2015: 108).

The formulaic nature of expressions of gratitude in general has been emphasized e.g. in Jautz' (2013) book-length investigation of thanking formulae in English. Based on corpus data from British and New Zealand English, she finds 887 expressions of gratitude that can be allocated to 19 more general patterns (Jautz 2013: 84) and that are connected with certain typical syntactic realizations, which in turn contain three optional elements: (i) naming of a benefactor, (ii) use of intensifying particles, (iii) naming of a reason along with the expression of gratitude, e.g. *thanks for your advice* (Jautz 2013: 285). A certain degree of formulaicity is characteristic not only for expressions of gratitude but also for many other polite speech acts such as apologies. This is in line with Terkourafi's (2015) suggestion that politeness itself can be understood in terms of conventionalization: "when we learn, through socialization in a community or group, that 'this is the way to do some thing' [...], what we are implicitly learning is that this is the right way of doing this thing (apologizing, requesting goods) in this type of context." (Terkourafi 2015: 16) In other words, language users have a considerable amount of meta-knowledge about the most common ways of doing specific things with language – e.g. performing requests, apologizing, or, of course, expressing gratitude. This does not mean that these illocutionary acts can only be performed using formulaic expressions, or that all expressions of e.g. gratitude are conventionalized to the same degree. Instead, politeness can arise as a generalized conversational implicature in a minimal context in the case of fairly fixed, "ritualized", patterns, or as a particularized conversational implicature in the case of novel expressions (Terkourafi 2015: 17). For the analysis of acknowledgment sections, this means that we can expect to find considerable variation in the degree to which the authors draw on 'pre-fabricated' expressions in the sense of Bybee's (2007: 16) prefabs, i.e. word sequences that are semantically transparent but used conventionally.

Summing up, then, previous research has shown that dissertation acknowledgments serve a number of social and professional functions, and that they typically

follow fairly fixed patterns. But there are also indications of slight and nuanced differences across disciplines, which makes it interesting to 'zoom in' on acknowledgments in one particular domain, combining the manual analysis methods used in most previous studies with large-scale automatic analyses using state-of-the-art natural language processing and data science tools.

3 The present study

Our study draws on a sample of 316 acknowledgment sections of doctoral dissertations. All were submitted to the Medical Faculty of the University of Düsseldorf and published in 2022 via the open access server of the University and State Library Düsseldorf.[5] We deliberately decided to restrict the focus of the current study to this very homogeneous sample. Unlike Wesian's (2015) fairly small but balanced dataset it does not allow for comparisons across time, disciplines, or universities, but it does allow for an in-depth analysis that can serve as a springboard for follow-up studies that take a broader variety of data into account. The advantage of such a homogeneous sample is that it provides us with the possibility to take a close look at acknowledgements conventions that apply in one very specific context and at potential patterns of variation, while at the same time ruling out some potential sources of variation. To some extent, a comparison with the previous literature can help us decide to what degree the detected patterns can be generalized to the text type of dissertation acknowledgments per se; regarding some aspects, this will have to be checked in subsequent studies.

3.1 Methods

As our goal is to inductively detect recurrent patterns in the data, as well as linking them to our theoretical assumptions on the functions of acknowledgment sections elaborated on above, we combine multiple quantitative methods that allow for a) detecting key words and phrases, b) detecting named entities such as person and institution names, and c) assessing the emotional valence of the words used in the texts.

To prepare the data for further processing, we first added part-of-speech and lemma tags to the data. More specifically, we used *spacy* (Honnibal et al. 2020) with the pre-trained *de_core_news_sm* model for POS-tagging and lemmatizing

5 Thanks to Giacomo Padrini for his help in the data collection process.

the data, as well as for the automatic recognition of named entities. This allows for a rough quantitative investigation of the corpus in terms of text length, lemma and part-of-speech distribution, as well as for approximating the number of names per text.

For a more in-depth qualitative analysis (see Section 3.3), we used a randomly drawn subsample of 30 acknowledgments. While this sample is certainly too small for a truly in-depth study, it seems large enough to draw some interesting first conclusions. We used MAXQDA (VERBI Software 2024) for the inductive analysis. We annotated the acknowledged persons as well as the reasons for acknowledging them, distinguishing, on the most coarse-grained level, between professional and private contacts, and between professional/practical and emotional support. Further subcategories were gauged inductively from the data. For example, a sentence like (4) would be coded as 'co-supervisor > practical support > correction of the thesis' and 'co-supervisor > emotional support > commitment'.[6] The shaded area shows the text segment coded in MAXQDA.

(4) Mein besonderer Dank gilt Herrn PD Dr. [NAME] für die ausgezeichnete Betreuung und engagierte Unterstützung bei der klinischen Durchführung meiner Untersuchungen, sowie seine fachlichen und persönlichen Anregungen während der Durchsicht dieser Arbeit- auch über den Atlantik und mehrere Zeitzonen hinweg. (101)

'My special thanks go to PD Dr. [NAME] for the excellent supervision and committed support in the clinical implementation of my studies, as well as his professional and personal suggestions while correcting this thesis, even across the Atlantic and multiple time zones.'

The qualitatively annotated subsample was also used as input for follow-up quantitative analyses. Firstly, we used distinctive-collexeme analysis (Gries & Stefanowitsch 2004) to check which words occur with above-chance frequency in the segments coded as professional and private, respectively. Distinctive collexeme analysis is typically used to compare two linguistic constructions in terms of the filler words that occur in their open slots, e.g. *will*-future vs. *going to*-future. Here we use it to compare the words that occur in the segments coded as 'professional' to those that occur in the segments coded as 'private'. Again, we used the *spacy* lemma annotation, and we limited the dataset to words tagged as nouns, verbs, or adjectives.

6 The fact that the person mentioned in (4) was the author's co-supervisor, rather than the main supervisor, becomes clear from the broader context of the acknowledgment section in question.

Secondly, using explorative methods that have proven insightful in discourse analysis, e.g. in analyzing political language (see e.g. Bubenhofer 2017), we also analyzed n-grams to detect recurrent multi-word units. Recurrent word combinations can give helpful clues to formulaic patterns that are frequently used in a given discourse and/or a given genre, which in turn allows for a more thorough understanding of the topics and the conventions of the discourse or genre. We broke down the texts into bigrams (i.e., a sentence like *ich danke meinem Betreuer* 'I thank my advisor' would be broken down into two-word units: *ich danke, danke meinem, meinem Betreuer*), trigrams (3-word units: *ich danke meinem, danke meinem Betreuer*) etc. up to 6-grams. All n-grams attested less than 10 times were removed; as some of the remaining n-grams were very similar to each other, we calculated the squared Euclidean distance (d^2) between each n-gram pair in the data. More specifically, we followed Bubenhofer (2017: 76) in creating a matrix containing all lemmas (taken from the *spacy*-tagged data) attested in the n-grams in the columns. For each n-gram, values of 0 or 1 indicate whether the respective value occurs in it, which was used as input for calculating the Euclidean distances between the n-gram pairs. For all n-gram pairs with $d^2<2$, we removed the n-gram for which the quotient of frequency and word count was lower from the dataset. The frequency of the remaining n-grams was compared to their frequency in a reference corpus.

Selecting a reference corpus entails the challenge that it has to be a dataset that can be seen as representative of contemporary German but that is also readily available in a format that allows for extracting n-gram frequency lists. Most reference corpora of German, such as the German Reference Corpus or the corpora of the Digital German Dictionary (DWDS), are not available in a full-text format and only allow for extracting unigram frequency lists. We therefore used the TIGER corpus (Brants et al. 2004), a 700,000-word corpus of German texts, as our baseline to detect n-grams that occur with above-chance frequency in the dataset, using the log-likelihood ratio G^2 as association measure. A limitation of TIGER is that it only contains newspaper texts – as such, it can be considered less representative than other corpora of present-day German. However, it has the advantage that its full text is readily available, which is a prerequisite for compiling n-gram lists that are comparable to the ones we compiled for the corpus of acknowledgment sections. As newspaper texts are written in a similar register as academic ones, and as they can be expected to cover a broad spectrum of topics, we would still argue that TIGER offers a good point of comparison and is therefore well-suited as a reference corpus for our purposes. Flach's (2021) R package *collostructions* was used to calculate the results. Fig. 2 shows the n-grams that occur with above-chance frequency in the data. The size of the n-grams reflects their association strength, i.e. n-grams that occur with highly significantly above-chance frequency in our dataset compared to the reference corpus are displayed in a larger font. Note that we are working

with lemma n-grams, i.e. all inflected forms are reduced to their base form, e.g. *bin, ist, war* 'am, is, was' to *sein* 'to be'. Also, the distinction between uppercase and lowercase letters was removed, and, following the conventions of the taggers used both by the TIGER corpus and in the *spacy* NLP toolchain, the definite articles *der, die, das* as well as their inflected forms are all given as *der*. This is why, e.g., *für die Bereitstellung des Materials* 'for the provision of the material' becomes *für der bereitstellung der material* in the lemma n-grams.

Again to test the hypothesis that different addressees and different topics conventionally tend to be mentioned at different positions within an acknowledgment section, we also assessed the relative position of (content) words in the text. Focussing on lexemes that occur at least 20 times in our full dataset, we divided the position of each attestation of the word by the total number of words in the text. For example, if *durchführen* 'conduct, perform' is the 7-th word in a text with 450 words, its position would be 7/450=0.016. For this step, we used the full set of 316 acknowledgment sections again. In addition, to test the hypothesis that more emotional terms tend to cluster towards the end of the text, we used the emotional valence ratings from the Berlin Affective Word List (BAWL-R, Võ et al. 2009), a crowdsourced dataset for which participants were asked to assess the emotional valence of words on a scale from -3 (very negative) to 3 (very positive).

3.2 Quantitative analysis

We first assessed the lengths of the acknowledgments, which varies between 27 and 621 words (mean = 211, median = 198, standard deviation = 93.4). Although there is quite some variation in the lengths of acknowledgments, as the left panel of Fig. 1 shows, there are only few outliers with more than 400 words, which indicates that

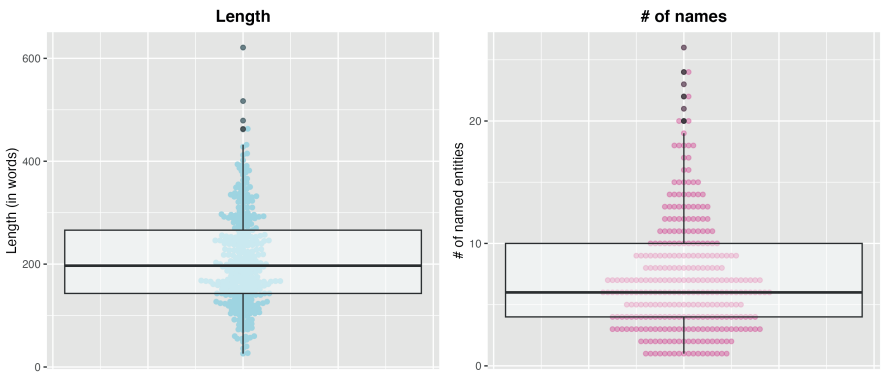

Figure 1: Length of the individual acknowledgment sections (left) as well as the number of automatically recognized named entities in each of the acknowledgment sections (right).

acknowledgments are a fairly standardized text type, at least in terms of length. The right panel of Fig. 1 shows the number of person references or references to institutions detected by the named-entity tagging procedure in each text. As the data were not manually corrected due to the large number of texts that entered the analysis, these numbers have to be taken with some caution, but overall, the data indicate that we typically find 5 to 10 persons or institutions that are acknowledged.

The top 5 n-grams detected in the data reflect both the professional and the personal side of acknowledgments, as well as their largely formulaic nature: *mein arbeit* 'my work', *ermöglichen haben* 'have made possible', *meinen eltern* '(to) my parents', *für die bereitstellung* 'for the provision', and *mein schwester* 'my sister'. Many of the other items among the top 100 n-grams (listed in Table 2 in the Appendix) also

Figure 2: Word cloud of the lemma n-grams that appeared significantly more often in the dataset than in a reference corpus.

underline the formulaic nature of the acknowledgments, pointing to patterns like *an dieser Stelle möchte ich mich bei XY bedanken* 'at this point I would like to thank XY', *mein besonderer Dank gilt* 'my special thanks are due to', or *von ganzem Herzen danken* 'thank from the bottom of my heart'.

The strongly associated n-grams also show once again how acknowledgments pertain to both the professional and the private domain, with both family terms such as *Eltern* 'parents' or *Schwester* 'sister' represented in the associated n-grams, but also titles like *Herr Univ.-Prof* (lit. 'Mr university [= full] professor').

It can also be interesting to take a closer look at the distribution of individual recurrent words across the texts – more specifically, at their relative position in the text. If the assumption is true that acknowledgment sections are a text type with a strongly conventionalized structure and highly formulaic phrasing, we expect that many lexemes occur across different acknowledgments, in similar positions in each text. Also, given that our qualitative analysis has shown that all acknowledgment sections in our dataset start out with acknowledgments in the professional domain and then move on to private acknowledgments (see Section 3.3 below), we expect that more emotional terms tend to occur towards the end of the text, in the sections containing the private acknowledgments.

To check this hypothesis, we calculated the relative position of words in the text as described in Section 3.1. Fig. 3 shows the mean relative position of each word, and the error bars show the standard error of the mean. To test the hypothesis that more emotional terms tend to cluster towards the end of the text in a data-driven way, we used the mean emotional valence ratings from the Berlin Affective Word List (BAWL-R, Võ et al. 2009). As mentioned in Section 3.1, the ratings range from -3 to +3. A visual inspection of Fig. 3 already suggests that terms with a high emotional valence tend to cluster towards the end of a text; indeed, a Spearman's rho test indicates that there is a fairly small but highly significant correlation between the mean position of a word and its absolute[7] emotional valence (ρ=0.17, S=6.63, $p<0.001$).

Our approach has a number of limitations that have to be kept in mind: Firstly, crowdsourced lists like BAWL-R can be problematic as the emotional valence ratings were obtained for words out of context, and of course the emotional valence of any word can differ depending on the context in which it occurs. Secondly, for each word, we only take the mean valence rating into account, although the ratings differ in their degree of dispersion. And thirdly, many frequent words in our dataset do not occur in BAWL-R at all. Nevertheless, the result lend tentative support to our hypothesis, which is further supported by the qualitative analysis in Section 3.3.

7 As we are interested in the emotional valence in either direction, positive or negative, we worked with absolute values for the correlation test, i.e. all negative values were turned into their positive counterparts, with -1 becoming 1, -2 becoming 2, etc.

But the results in Fig. 3 are also interesting beyond the emotional valence measures. As a reviewer correctly points out, items designating more interpersonal relationships tend to occur later in the texts. This is in line with the general tendency, discussed in more detail in Section 3.3 below, that acknowledgments in the professional domain tend to precede those in the private domain. An interesting additional finding relating to the way acknowledgments are typically phrased is that the word form *Danke* 'thanks' tends to occur fairly late in the acknowledgments as well, indicating that it tends to occur when thanks are extended to private contacts. In isolation, *Danke*, like English *thank you* or *thanks*, can be considered "highly ritualised" (Watts 2003: 19; also see Terkourafi 2003: 177), a response to "conventionally polite behaviour" (Jautz 2013: 14). But if accompanied by an explanation, as in the examples in (5), the simplest way of expressing gratitude may be perceived as among the most effective ones.

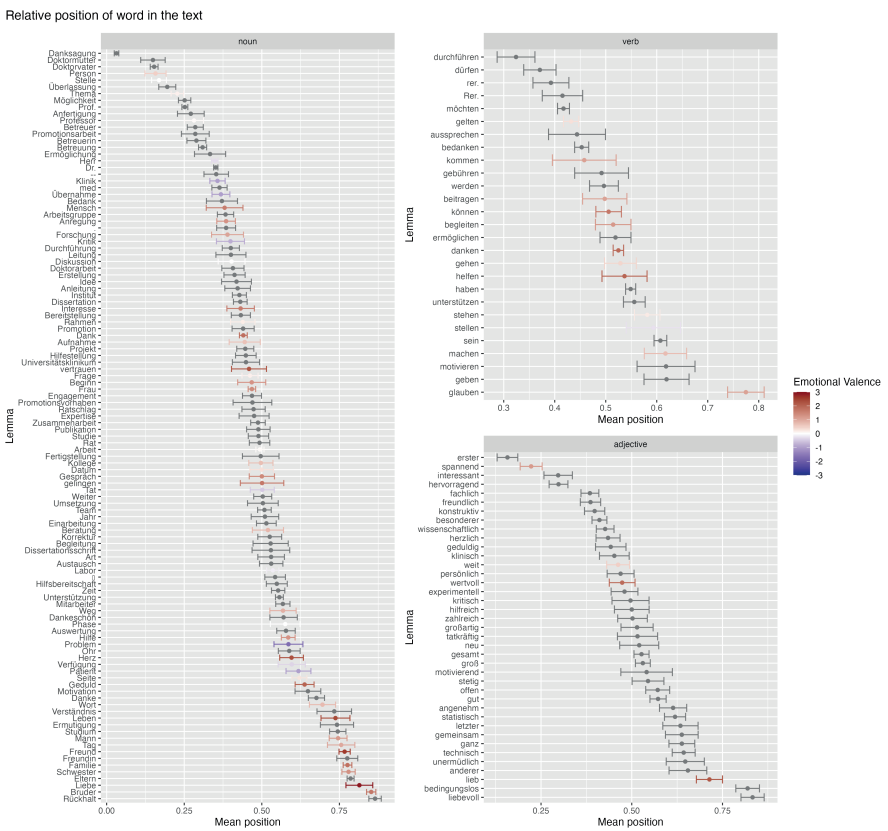

Figure 3: Mean relative positon of a word in the text. The colors show the mean emotional valence rating according to the Berlin Affective Word List (BAWL-R, Võ et al. 2009). Items displayed in grey do not have a valence rating in BAWL-R. The error bars show the standard error of the mean.

(5) a. Danke für die viele Geduld, die vielen Arschtritte und deine bedingungslose Unterstützung in allen Lebenssituationen. ‘Thanks for the large amount of patience, the many kicks in the arse and your unconditional support in all situations of life.’ (234)

b. Danke für ihre zahlreichen Korrekturlesungen, ihre vielen liebevollen Ermutigungen und dafür, dass sie immer an meiner Seite ist. ‘Thanks for her [= my wife’s] many rounds of proofreading, her affectionate encouragement and for always being by my side.’ (314)

c. Danke, dass ihr immer für mich da seid! Ihr habt mich zu der Person aufwachsen lassen, die ich heute bin und euch verdanke ich unendlich viel. ‘Thanks that you are always there for me! You let me grow up to the person that I am today, and I owe you infinitely much.’ (153)

Importantly, *Danke* is used to address the recipient(s) directly, as also becomes clear from the use of second-person forms in (5a) and (5c). Thus, it seems plausible to assume that *Danke*, if not used as a ritualized response to a simple everyday favor (where it would count as merely politic behavior in the sense of Watts 2003: 19, i.e. non-salient behavior that participants construct as being appropriate to the ongoing situation), is perceived as a very direct way of expressing gratitude, and that other formulaic expressions, which often go in tandem with intensification, also serve a hedging function to some extent, reflecting the more distant interpersonal relationship between the thanker and the addressee. As Culpeper et al. (2021: 208) point out, “thanking threatens the speaker’s face more than the hearer’s.” As such, it seems straightforward to assume that when expressing gratitude, language users tend to make use of “hedging” devices to ensure that their relationship to the addressee is not misconstrued. This, however, is currently a working hypothesis that should be tested in more detail in subsequent research.

3.3 Qualitative analysis (and quantitative follow-up analyses) of a subsample

In addition to the quantitative analysis, we conducted a qualitative analysis to assess who the authors are thanking and why. While we offer a more in-depth analysis of the qualitative results elsewhere, we will briefly summarize the main results here and add some follow-up quantitative analyses drawing on the manually coded data.

As mentioned in Section 3.1, a sample of 30 acknowledgment sections was annotated manually using an inductive set of categories. To gauge recurrent patterns in the acknowledgment sections, we analyzed the codes that were assigned at least 10 times in more detail. As mentioned in Section 3.1, they were grouped by the macro-categories of private vs. professional contexts and practical, emotional, and other kinds of support, and by groups of persons that are acknowledged. Among the private contacts, parents (32 segments) and partners (42 segments[8]) are explicitly mentioned in most acknowledgments; other groups are mentioned much less often, e.g. siblings (mentioned 10 times overall), external consultants (6), other PhD students or colleagues (5), children (3) or mothers-in-law (1). In the professional domain, the broadest variety of recurrent items can be found in the acknowledgments pertaining to main and co-supervisors.[9]

The manual coding of the sample also allows for assessing the distribution of the coded segments over the individual texts. Fig. 4 shows how the professional and the private acknowledgments are distributed over the texts. The left panel shows how the segments coded for 'professional' or 'private' are distributed over the text; the right panel shows the beginning of each segment. With very few exceptions, the professional and private domains are clearly separated; and without exception, the acknowledgments start in the professional domain and then, in most cases, move to the private domain; two acknowledgments, however, do not contain private acknowledgments at all. As the right panel of Fig. 4 shows, the private acknowledgments never start before the second quarter of the text, and usually even considerably later. This is in line with the findings from our quantitative analysis.

As mentioned in Section 3.1, we also conducted a distinctive collexeme analysis to compare the segments coded as professional to those coded as private. Table 1 shows the results. Titles such as *Prof., Dr., med.* (part of the title *Dr. med.*) unsurprisingly occur in the professional domain; apart from that, the professional segments are characterized by work-related words like *Betreuung* 'supervision', *Durchführung* 'implementation', *Arbeitsgruppe* 'working group' or *wissenschaftlich* 'scientific'. In the private domain, we find many family and relationship terms as well as emotionally loaded adjectives like *liebevoll* 'loving, affectionate', *bedingungslos* 'unconditional' as well as positively connotated nouns like *Ermutigung*

8 Some individuals, especially parents and partners, are mentioned in multiple segments in the same text, which is why the numbers for these two categories are higher than the overall number of texts in the subsample.

9 In the case of the latter, however, we cannot always clearly distinguish between the second examiner of a PhD thesis and other co-supervisors, which, in the context of medical dissertations, often collaborate more closely with the PhD candidate and can therefore potentially play a more important role than the main supervisors.

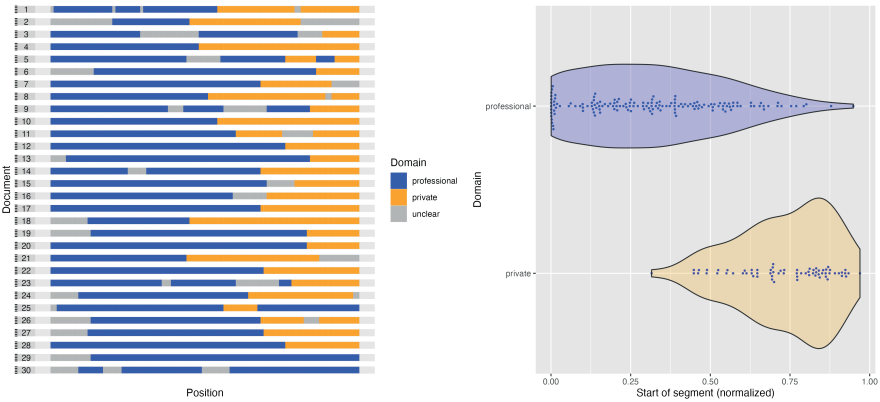

Figure 4: Left panel: Relative position of professional and private acknowledgments in the text. Right panel: Start of each coded segment allocated to either the professional or the private domain in the text.

'encouragement' and *Unterstützung* 'support', which often do not just pertain to the doctoral phase but to the entire 'path of life' (*Lebensweg*). This is in line with the quantitative findings above: As Fig. 3 shows, terms related to interpersonal relationships tend to occur later in the texts. These results lend further support to the hypothesis that acknowledgment sections in German medical dissertations (and probably also in other disciplines and other languages) usually show a fairly clear division between professional acknowledgments that are written in a relatively distanced, professional style and private acknowledgments that are written in a more emotional style.

4 Discussion and conclusion

Combining quantitative and qualitative methods, we have shown that acknowledgment sections of medical dissertations show a fairly clear and predictable structure, which indicates that this is a text type with clear conventions. This also becomes clear from the fact that some medical faculties offer guidelines on how to structure and phrase the acknowledgment section in medical dissertations.[10] Our n-gram analysis points to a high proportion of formulaic language, which in turn

10 See e.g. https://www.medizinische-fakultaet-hd.uni-heidelberg.de/fileadmin/medizinische_fakultaet/Dekanat/Promotion/Dr_sc_hum/2023_08_01_Abfassungsrichtlinien_Version_2_3.pdf (last checked 07/04/2024)

Table 1: Top 25 verbs, nouns, and adjectives identified as distinctive collexemes in segments coded as professional or private, respectively.

	professional			private		
Rank	Lemma	Freq	G^2	Lemma	Freq	G^2
1	Dr.	99	56.13	Eltern ‘parents’	16	41.78
2	Herr ‘Mr.’	60	38.9	Familie ‘family’	17	37.31
3	Prof.	38	24.5	Freund ‘(boy)friend’	11	28.66
4	Doktorvater ‘advisor’	19	12.19	liebevoll ‘loving’	6	15.6
5	Projekt ‘project’	19	12.19	Mutter ‘mother’	6	15.6
6	Betreuung ‘supervision’	27	11.31	Schwester ‘sister’	6	15.6
7	Möglichkeit ‘possibility’	13	8.33	Liebe ‘love’	5	12.99
8	Durchführung ‘implementation’	10	6.4	Rückhalt ‘support’	5	12.99
9	Klinik ‘clinic’	10	6.4	Studium ‘studies’	8	12.07
10	konstruktiv ‘constructive’	10	6.4	glauben ‘believe’	4	10.39
11	Labor ‘laboratory’	10	6.4	Herz ‘heart’	5	8.22
12	med	10	6.4	Brüdern ‘brothers’	3	7.79
13	Professor	10	6.4	Ehemann ‘husband’	3	7.79
14	Anregung ‘suggestion’	8	5.12	Freundin ‘(girl)friend’	3	7.79
15	wissenschaftlich ‘scientific’	8	5.12	haben ‘have’	3	7.79
16	Arbeitsgruppe ‘working group’	7	4.48	Kraft ‘power’	3	7.79
17	Beginn ‘beginning’	7	4.48	Lebensweg ‘path of life’	3	7.79
18	Bereitstellung ‘provision’	7	4.48	Unterstützung ‘support’	25	6.8
19	Danksagung ‘acknowledgment’	14	4.21	Ausbildung ‘education’	4	6.02
20	Einblick ‘insight’	6	3.84	Ermutigung ‘encouragement’	4	6.02
21	Engagement ‘commitment’	6	3.84	letzter ‘last’	4	6.02
22	Hilfestellung ‘help’	6	3.84	Motivation ‘motivation’	4	6.02
23	Überlassung ‘assignment’	6	3.84	bedingungslos ‘unconditional’	5	5.89
24	durchführen ‘conduct’	5	3.2	danken ‘thank’	29	3.69
25	Frage ‘question’	5	3.2	Geduld ‘patience’	6	3.61

also points to a high degree of conventionality. To a large extent, these results are in line with the findings of previous research, which has shown that acknowledgment sections have become a text type in their own right over time (e.g., Wesian 2015). Also, the high degree of formulaicity is not surprising, given the fact that in everyday language, thanks are usually expressed in what Watts (2003: 186) calls formulaic

and ritualized expressions of procedural meaning. As some of the examples cited in the present paper show, there is still considerable variation in how thanks are expressed. While we could not explore this variation in detail in the present paper, this is clearly an aspect that deserves more attention in future follow-up studies.

Our analysis goes beyond previous research in using a relatively large, fairly homogeneous dataset to provide an in-depth account of contemporary acknowledgments in doctoral theses in one single discipline, medicine. One main result of our study is that the acknowledgments in our dataset typically show a clear division between a professional and a private part that differ significantly in the ways in which the acknowledgments are phrased. Overall, our results indicate that the professional acknowledgments are usually phrased in a relatively distanced tone, with relatively few emotionally loaded terms. This also shows in the way the acknowledged persons are addressed: The authors tend to use titles (*Prof. Dr.*) and full names or last names. When the persons are directly addressed, the polite address pronoun *Sie* is usually used; (6) is a salient exception, in which the author uses first names and the informal pronoun *du*.

(6) Axel, vielen Dank für die Möglichkeit, in deinem Labor zu promovieren, deine stets offene Tür und Unterstützung, auch fernab des Labors. Prof. C.-K. danke ich für die freundliche Übernahme der Co-Betreuung. André, herzlichen Dank für die super nette, zuverlässige und kompetente Betreuung. Auch für den Anstoß, mal den eigenen Kopf zu benutzen, habe ich zu danken. (118)

'Axel, thank you for the possibility to do an MD in your lab, for the fact that your door was always open and for your support even beyond the lab. I thank Prof. C.-K. for kindly agreeing to be my co-supervisor. André, my heartfelt thanks for the super nice, reliable, and competent supervision. I also have to thank you for the suggestion to use my own head sometimes.'

A number of potential follow-up questions emerge from this perspective, and from the obvious limitations of the present study. Before we turn to the desiderates for future research, let us first spell out the limitations in more detail: We have zoomed in on a very specific sample of doctoral dissertations, and the sample that we used for qualitative analysis was comparatively small. In addition, our quantitative analysis draws on off-the-shelf automatic taggers with all the potential problems this entails. Also, the annotation for the qualitative analysis was conducted by only one person (A.L.), without checking for inter-coder reliability.

Turning to open questions and potential follow-up studies, it would, firstly, be interesting to compare the acknowledgments conventions across disciplines in

more detail. While a large part of previous studies has in fact focused on this contrastive aspect, it seems promising to use the qualitative and quantitative methods applied in the present study to examine whether there are differences between disciplines regarding the degree of conventionalization of acknowledgments, and regarding the specific conventions. For instance, one may hypothesize that our observation that we usually find a strict division between the professional and the private domain, with a more distanced style in the professional part, may apply to a lesser degree in disciplines that typically have a less hierarchical structure than medical faculties usually have. Secondly, the formulaic patterns that are used in different parts of acknowledgment sections could be identified using a combination of manual annotation and automatic bottom-up pattern-detection methods. This could in turn help to find correlations between specific formulaic patterns and the functions they serve (e.g., expressing gratitude in the professional or in the private domain). Thirdly, the historical point of view taken by Wesian (2015), who took a largely qualitative perspective, could be extended, taking much larger datasets into account. In particular, it would be interesting to trace the diachronic development of the present-day conventions that the current study could only begin to describe. Finally, going beyond the corpus-based perspective taken here, it seems promising to complement this line of research with questionnaire and/or interview studies to get first-hand impressions of the motivations that play a role in the choices that authors make when writing acknowledgment sections.

Acknowledgments

We are grateful to the anonymous reviewers for their thorough and helpful comments that have significantly improved the reader-friendliness of the present paper. Remaining errors are of course our own responsibility.

Data availability

Although all data we used are publicly available, their licenses vary. Hence, for copyright reasons, but also out of privacy considerations, we decided against making the full data publicly available. However, the analysis script and aggregated data (e.g., collexeme lists) are available at https://osf.io/az9j4/. We are of course happy to provide the full data upon request.

References

Aijmer, Karin. 1996. Conversational routines in English: convention and creativity (Studies in Language and Linguistics). London: Longman.

Ben-Ari, Eyal. 1987. On acknowledgements in ethnographies. Journal of Anthropological Research 43(1). 63–84.

Brants, Sabine, Stefanie Dipper, Peter Eisenberg, Silvia Hansen, Esther König, Wolfgang Lezius, Christian Rohrer, George Smith & Hans Uszkoreit. 2004. TIGER: Linguistic interpretation of a German corpus. Journal of Language and Computation 2. 597–620.

Brown, Penelope & Stephen C. Levinson. 1987. Politeness. Some universals in language usage. Cambridge: Cambridge University Press.

Brown, Robert. 2009. How Scholars Credit Editors in Their Acknowledgements. Journal of Scholarly Publishing 40(4). 384–398. https://doi.org/10.3138/jsp.40.4.384.

Bubenhofer, Noah. 2017. Kollokationen, n-Gramme, Mehrworteinheiten. In Kersten Sven Roth, Martin Wengeler & Alexander Ziem (eds.), Handbuch Sprache in Politik und Gesellschaft, 69–93. De Gruyter. https://doi.org/10.1515/9783110296310-004.

Bühler, Karl. 1934. Sprachtheorie. Die Darstellungsfunktion der Sprache. Jena: Fischer.

Caesar, Terry. 1992. On acknowlegdements. The New Orleans Review 19(1). 82–94.

Croft, William. 2009. Toward a Social Cognitive Linguistics. In Vyvyan Evans & Stéphanie Pourcel (eds.), New Directions in Cognitive Linguistics (Human Cognitive Processing), vol. 24, 395–420. Amsterdam and Philadelphia: John Benjamins.

Cronin, Blaise, Gale McKenzie & Lourdes Rubio. 1993. The norms of acknowledgement in four humanities and social science disciplines. Journal of Documentation 49(1). 29–43.

Cronin, Blaise, Gale McKenzie & Michael Stiffler. 1992. Patterns of acknowledgement. Journal of Documentation 48(2). 227–239.

Cronin, Blaise & Kara Overfelt. 1994. The scholar's courtesy: A survey of acknowledgment behaviour. Journal of Documentation 50(2). 165–196.

Culpeper, Jonathan, Samuel J. Oliver & Vittorio Tantucci. 2021. Politeness reciprocity in Shakespeare's dialogue: The case of *thanks*. Journal of Historical Pragmatics 22(2). 202–224. https://doi.org/10.1075/jhp.00053.cul.

Della Constanza, Mario. 2013. Die Performanz: Brücke von der Konstruktionsgrammatik zur Textlinguistik? In Sabine De Knop, Fabio Mollica & Julia Kuhn (eds.), Konstruktionsgrammatik in den romanischen Sprachen, 299–325. Frankfurt: Peter Lang.

Desrochers, Nadine, Adèle Paul-Hus & Jen Pecoskie. 2017. Five decades of gratitude: A meta-synthesis of acknowledgments research. Journal of the Association for Information Science and Technology 68(12). 2821–2833. https://doi.org/10.1002/asi.23903.

Eckert, Penelope. 2012. Three Waves of Variation Study: The Emergence of Meaning in the Study of Sociolinguistic Variation. Annual Review of Anthropology 41(1). 87–100. https://doi.org/10.1146/annurev-anthro-092611-145828.

Evans, Vyvyan. 2019. Cognitive Linguistics. A complete guide. Edinburgh: Edinburgh University Press.

Flach, Susanne. 2021. collostructions: An R Implementation for the Family of Collostructional Methods. www.sfla.ch (last checked 24/03/2024).

Genette, Gérard. 1997. Paratexts: thresholds of interpretation (Literature, Culture, Theory 20). Cambridge ; New York, NY, USA: Cambridge University Press.

Giannoni, D. S. 2002. Worlds of Gratitude: A Contrastive Study of Acknowledgement Texts in English and Italian Research Articles. Applied Linguistics 23(1). 1–31. https://doi.org/10.1093/applin/23.1.1.

Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending Collostructional Analysis: A Corpus-Based Perspective on "Alternations." International Journal of Corpus Linguistics 9(1). 97–129. https://doi.org/10.1075/ijcl.9.1.06gri.

Hansson, N., Moll, F., Halling, T. *et al.* 2019. Scientific language trends among Swedish urologists and surgeons 1900–1955. World Journal of Urology 37. 975–982. https://doi.org/10.1007/s00345-018-2451-z

Halliday, M.A.K. 1978. Language as a social semiotic. The social interpretation of language and meaning. London: Edward Arnold.

Hart, Christopher. 2014. Discourse, grammar and idelogy. Functional and cognitive perspectives. London: Bloomsbury.

Hoffmann, Thomas. 2022. Construction grammar: the structure of English (Cambridge Textbooks in Linguistics). Cambridge: Cambridge University Press.

Honnibal, Matthew, Ines Montani, Sofie Van Landeghem & Adriane Boyd. 2020. spaCy: Industrial-strength natural language processing in python. Zenodo. https://doi.org/10.5281/zenodo.1212303.

Hyland, Kenneth. 2004. Graduates' gratitude: the generic structure of dissertation acknowledgements. English for Specific Purposes 23(3). 303–324. https://doi.org/10.1016/S0889-4906(03)00051-6.

Hyland, Ken. 2003. Dissertation Acknowledgements: The Anatomy of a Cinderella Genre. Written Communication 20(3). 242–268. https://doi.org/10.1177/0741088303257276.

Jakobs, Eva-Maria. 1997. Textproduktion als domänen- und kulturspezifisches Handeln. Diskutiert am Beispiel wissenschaftlichen Schreibens. In Kirsten Adamzik, Gerd Antos & Eva-Maria Jakobs (eds.), Domänen- und kulturspezifisches Schreiben (Textproduktion Und Medium 3), 9–29. Frankfurt am Main ; New York: P. Lang.

Jakobson, Roman. 1960. Linguistics and poetics. In Thomas A. Seboek (ed.), Style in Language, 350–377. Cambridge: MIT Press.

Jautz, Sabine. 2013. Thanking formulae in English: explorations across varieties and genres. Amsterdam: John Benjamins.

Koller, Werner. 2001. DANK und DANKSAGUNG – eine annäherung. In Eva-Maria Jakobs & Annely Rothkegel (eds.), Perspektiven auf stil, 267–304. Berlin, Boston: Max Niemeyer Verlag. https://doi.org/doi:10.1515/9783110941524.267.

Langacker, Ronald W. 2008. Cognitive Grammar: A Basic Introduction. Oxford: Oxford University Press.

Mantai, Lilia & Robyn Dowling. 2015. Supporting the PhD journey: insights from acknowledgements. International Journal for Researcher Development 6(2). 106–121. https://doi.org/10.1108/IJRD-03-2015-0007.

Mellado Blanco, Carmen, Fabio Mollica & Elmar Schafroth. 2022. Phrasem-Konstruktionen in der heutigen Forschung: Ein Überblick. In Carmen Mellado Blanco, Fabio Mollica & Elmar Schafroth (eds.), Konstruktionen zwischen Lexikon und Grammatik. Phrasem-Konstruktionen monolingual, bilingual und multilingual, 1–18. Berlin, Boston: De Gruyter.

Merten, Marie-Luis. 2018. Literater Sprachausbau kognitiv-funktional. Funktionswort-Konstruktionen in der historischen Rechtsschriftlichkeit. Berlin, Boston: De Gruyter.

Micciche, Laura R. 2017. Acknowledging writing partners (Perspectives on Writing). Fort Collins, Colorado : Boulder, Colorado: The WAC Clearinghouse ; University Press of Colorado.

Mikołajczyk, Beata. 2013. „Mein Dank gilt in erster Linie meinem Doktorvater" – Danksagungen aus linguistischer Perspektive. Studia Germanica Posnaniensia (33). 77. https://doi.org/10.14746/sgp.2013.33.07.

Morin, Cameron, Guillaume Desagulier & Jack Grieve. 2024. A social turn for Construction Grammar: double modals on British Twitter. English Language and Linguistics. https://doi.org/10.1017/S1360674323000576.

Sanderson, Tamsin. 2005. Kontrastive Fachtextpragmatik deutsch- und englischsprachiger wissenschaftlicher Texte: Danksagungen im interlingualen Vergleich. In Kirsten Adamzik & Wolf-Dieter Krause (eds.), Text-Arbeiten: Textsorten im fremd- und muttersprachlichen Unterricht an Schule und Hochschule (Europäische Studien zur Textlinguistik 1), 61–86. Tübingen: Narr.

Schmid, Hans-Jörg. 2020. The dynamics of the linguistic system: usage, conventionalization, and entrenchment. Oxford: Oxford University Press.

Schmid, Hans-Jörg. 2016. Why Cognitive Linguistics must embrace the social and pragmatic dimensions of language and how it could do so more seriously. Cognitive Linguistics 27(4). 543–557. https://doi.org/10.1515/cog-2016-0048.

Stumpf, Sören. 2015. Formelhafte (Ir-)Regularitäten: korpuslinguistische Befunde und sprachtheoretische Überlegungen (Sprache - System Und Tätigkeit Band 67). Frankfurt am Main: Peter Lang Edition.

Terkourafi, Marina. 2003. Generalised and particularised implicatures of linguistic politeness. In Peter Kühnlein, Hannes Rieser & Henk Zeevat (eds.), Perspectives on dialogue in the new millennium, 149–164. Amsterdam, Philadelphia: John Benjamins.

Terkourafi, Marina. 2015. Conventionalization: A new agenda for im/politeness research. Journal of Pragmatics 86. 11–18. https://doi.org/10.1016/j.pragma.2015.06.004.

Ungerer, Tobias & Stefan Hartmann. 2023. Constructionist approaches: Past, present, future. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781009308717.

Ungerer, Friedrich & Hans-Jörg Schmid. 2006. An Introduction to Cognitive Linguistics. 2nd edn. Harlow: Longman.

VERBI Software. 2024. MAXQDA 2024. Berlin: Verbi Software. https://www.maxqda.com/ (last checked 17/03/2024)

Võ, Melissa L. H., Markus Conrad, Lars Kuchinke, Karolina Urton, Markus J. Hofmann & Arthur M. Jacobs. 2009. The Berlin Affective Word List Reloaded (BAWL-R). Behavior Research Methods 41(2). 534–538. https://doi.org/10.3758/BRM.41.2.534.

Watts, Richard J. 2003. Politeness (Key Topics in Sociolinguistics). Cambridge: Cambridge University ress.

Wesian, Julia. 2015. Danksagungen in Dissertationen: Zur Genese einer Textsorte (Duisburger Arbeiten Zur Sprach- Und Kulturwissenschaft = Duisburg Papers on Research in Language and Culture Band/volume 106). Frankfurt am Main ; New York: Peter Lang Edition.

Ziem, Alexander. 2013. Beyond the sentence: towards a cognitive-linguistic approach to textual reference. Yearbook of the German Cognitive Linguistics Association 1. 39–58.

Appendix

Table 2: Top 100 associated n-grams detected in the data

Rank	n-gram	Freq	G^2	Rank	n-gram	Freq	G^2	Rank	n-gram	Freq	G^2
1	mein arbeit	68	44.49	21	dafür dass	24	18.59	39	mir immer	21	16.26
2	ermöglichen haben	52	40.28	22	für der ermöglichung	24	18.59	41	an mein seite	20	15.49
3	meinen eltern	46	35.63	23	herzlich bedanken	24	18.59	42	bedanken ich sich für der	20	15.49
4	für der bereitstellung	40	30.98	24	insbesondere mein	24	18.59	43	der anfertigung dieser	20	15.49
5	mein schwester	40	30.98	25	konstruktiv kritik	24	18.59	44	der mich während	20	15.49
6	dass ich	33	25.56	26	mein promotion	24	18.59	45	ich sich bei herr prof dr	20	15.49
7	für der bereitstellung der	32	24.78	27	vieler dank für der	24	18.59	46	mein eltern und	20	15.49
8	für der freundlich	32	24.78	28	der gesamt team der	23	17.81	47	sowie mein	20	15.49
9	dieser projekt	40	23.85	29	dieser stelle möchten ich sich bei	23	17.81	48	danken ich für ihr	19	14.72
10	herr univprof	30	23.24	30	für der unterstützung bei	23	17.81	49	dass du	19	14.72
12	der promotion	29	22.46	31	dieser doktorarbeit	22	17.04	50	der cobetreuung	19	14.72
14	bedanken der mich	28	21.69	32	für der übernahme der	22	17.04	51	für ihr unterstützung	19	14.72
15	mein betreuer	28	21.69	33	zu jeder zeit	22	17.04	52	mich während der	19	14.72
16	mein doktorarbeit	28	21.69	34	mit rat und tat zu seite	30	16.67	53	und frau dr	19	14.72
17	mir stets	26	20.14	35	dank gebühren	21	16.26	54	zu gut letzt	19	14.72
18	außerdem danken ich	25	19.36	36	eltern und mein	21	16.26	55	frau pd dr	18	13.94
19	und mich	33	18.81	37	herr pd dr med	21	16.26	56	herr professor dr	18	13.94
20	begleiten haben	39	18.77	38	mein freundin	21	16.26	57	herzlich für der	18	13.94

Table 2 (continued)

Rank	n-gram	Freq	G^2	Rank	n-gram	Freq	G^2	Rank	n-gram	Freq	G^2
58	ich sehr	18	13.94	73	bei der durchführung	16	12.39	88	besonderer dank gelten herr	15	11.62
59	in labor und	18	13.94	74	bei herr prof dr med	16	12.39	89	danken ich frau	15	11.62
60	mein betreuerin	18	13.94	75	dank gelten der	16	12.39	90	danksagung mein besonderer dank gelten	15	11.62
61	möchten ich sich für der	18	13.94	76	danke an	16	12.39	91	dass er	15	11.62
62	danken für der	17	13.17	77	danken der mich	16	12.39	92	der doktorarbeit	15	11.62
63	für der unterstützung bei der	17	13.17	78	danken ich der	16	12.39	93	der experimentell	15	11.62
64	herzlich dank gelten	17	13.17	79	danken ich für sein	16	12.39	94	der mich während der	15	11.62
65	ich dieser arbeit	17	13.17	80	danken ich mein familie	16	12.39	95	dieser arbeit und	15	11.62
66	ich frau dr	17	13.17	81	danksagung an erster stelle	16	12.39	96	dissertation in	15	11.62
67	mich stets	17	13.17	82	der unterstützung in	16	12.39	97	dr thomas	15	11.62
68	möchten ich frau	17	13.17	83	gesamt studium	16	12.39	98	ein ganz besonderer dank	15	11.62
69	möchten ich sich bei mein familie	17	13.17	84	ich mein eltern	16	12.39	99	ein groß dankeschön	15	11.62
70	studium und der	17	13.17	85	mein mutter	16	12.39	100	frau prof dr med	15	11.62
71	von ganz herz	25	13.16	86	unterstützung bedanken	16	12.39				
72	sowie für der	24	12.46	87	weit dank gelten	16	12.39				