German English-medium linguistics journal abstracts over the last 30 years: Quantitative and qualitative structural developments

Marina Ivanova

Chemnitz University of Technology

ABSTRACT

Abstracts are central to academic writing as they summarise and promote publications – this paper shows that the widespread use of abstracts started in the 90s and increased rapidly, becoming a standard. It analyses 593 articles with 555 abstracts from nine linguistics journals, affiliated with the Web of Science to German institutions. The focus is on global rhetorical structures adopted – from introduction/research questions to methods, results, discussions, and conclusions. Additionally, I identify trends in writers' stance expression through selected metadiscourse features as expressed in subjects and verbs. The analyses demonstrate that abstracts from Germany have become more unified towards the scientific IMRAD model. This model, however, has been adapted to the advertising function of the abstract with a stronger emphasis on the authors' contributions and article's importance (especially in introductions, methods, and results) and rare discussions of conclusions and limitations. Thus, general academic writing structures have been adapted to genre-specific functional practices over the last 30 years.

Keywords: journal abstract, metadiscourse, stance, rhetorical structures, IMRAD.

1. Introduction

Abstracts are an indispensable part of journal publications today as they are used to summarise publications and promote them to a specific research community (see Hyland 2000: 64). However, this has not always been the case – abstracts have changed both quantitatively and qualitatively in the

DOI: 10.25951/11260

last 30 years. This article aims to show that the use of abstracts in linguistics journals started in the 90s, becoming a standard today. It also explores the genre developments in terms of rhetorical structure and word length.

Although the journal abstract is attached to a research article, it is traditionally viewed as a genre of its own due to its different structure and purpose (Gillaerts – van de Velde 2010: 135; Hyland 2000: 64). The genre characteristics have been systematized in different typologies (Bhatia 1993; Gillaerts 2013; Hyland 2000; Swales – Feak 2009), mostly through the classification into rhetorical moves. Still, notable disciplinary, cultural and diachronic differences have been recognized (see among others Bondi – Lorés Sanz 2014; Friginal – Mustafa 2017). This paper explores abstracts' developments in terms of global rhetorical structures, in particular in the use of the common IMRAD structure (Introduction, Method, Results and Discussion) together with Research Questions, Conclusions and Limitations. IMRAD is a common structure for research articles, which has also been applied to abstracts (e.g., Lorés Sanz 2004).

In order to avoid confounding variables and potential cultural differences, the focus is on abstracts by scholars affiliated with German institutions. The texts are compiled from the Web of Science (WoS) (2022) database and comprise 593 entries from nine high-impact international English-medium linguistics journals. The academic writing discourse is therefore explored in a context where authors want to reach a relatively broad audience and attract readers from their research communities.

After an overview of the general diachronic development (1969-2021), the paper turns to a quantitative and qualitative analysis of moves in 100 randomly sampled abstracts. The distribution of rhetorical structures is viewed in terms of the overall sample, its development throughout the last 30 years, and its distribution in the selected journals. Then, the qualitative analysis highlights small-grain tendencies in move structure. This prompts a discussion of the extent to which the tendencies are explained by the time articles were published and by the article approach.

The paper also turns to the most frequent subjects and verbs and explores how they are used as metadiscourse markers to express stance and attitude. It tests whether authors prefer personal and positive markers, which would be the logical hypothesis for the advertising function of abstracts. Finally, it combines the insights into the global rhetorical structure and the local metadiscourse expressions to draw conclusions on the genre developments of abstracts throughout the last 30 years.

2. Global rhetorical structure of journal abstracts

The global rhetorical structure of genres has most commonly been analysed through a classification into moves. A move can be seen as a rhetorical stage in a genre. It has a minor communicative purpose to fulfill, which in turn serves the major communicative purpose of the genre (Dos Santos 1996: 485).

Several authors have classified the moves of the journal abstract with a different level of detail. Based on Bhatia's summary, Move 1 'Introducing purpose' is used to express aims and (hypo)theses; Move 2 'Describing methodology' indicates the experimental design and scope; Move 3 'Summarizing results' comprises the findings and problem solutions and Move 4 'Presenting Conclusions' interprets the results and provides applications (Bhatia 1993: 148-149).

Dos Santos provides options for submoves and even substeps to these submoves: Move 1 'Situating the research'; Submove 1A 'Stating current knowledge' and/or Submove 1B 'Citing previous research' and/or Submove 1C 'Extended previous research' and/or Submove 2 'Stating a problem' (Dos Santos 1996: 485).

Hyland (2000) presents 'Purpose' as a separate move and frames the results as a 'Product' in his classification of abstract moves: 'Introduction', 'Purpose', 'Method', 'Product' and 'Conclusion' (Hyland 2000: 67). This terminology suits the advertising function of abstracts stressed in his study, which presents the use of moves as strategies for rhetorical persuasion (2000: 68), claiming significance (2000: 75) and insider credibility (2000: 78).

Gillaerts (2013: 52) summarizes abstract moves and intermediate steps as follows: Move 1 'Situating the research' (1a 'Current knowledge', 1b 'Specific references', 1c 'Indicating the gap'); Move 2 'Presenting the research' (2a 'Purpose', 2b 'Research question or hypothesis'); Move 3 'Describing the methodology', Move 4 'Summarizing the findings' and Move 5 Discussing the research' (5a 'Conclusions', 5b 'Recommendations'). He recognizes several 'methodological issues with the categorization of Moves both in terms of separating overlapping moves and in terms of identifying them in the abstracts.

There are several instructional works on abstract Moves. For instance, Salager-Meyer gives a critical dimension to abstract structure and argues that the move structure is one of the most important features of a well-written abstract and it should include "purpose, methods, results (or data synthesis) and conclusions (optional in case reports)" (Salager-Meyer 1990: 370). These Moves are set in correlation with the overall paper structure. Swales and

Feak's (2009) textbook summarises the abstract moves in a similar manner: Move 1 'Background/ introduction/ situation'; Move 2 'Present research/ purpose'; Move 3 'Methods/ materials/ subjects/ procedures'; Move 4 'Results/ findings'; Move 5 'Discussion/ conclusion/ implications/ recommendations' (Swales – Feak 2009: 5).

Abstracts have also been compared to other genres, for instance by applying the CARS model (Creating A Research Space) designed for introductions (Swales 1990, 2004) to abstracts (Gillaerts 2013; Lorés Sanz 2004). Lorés Sanz (2004) found both examples of CARS and IMRAD structure in linguistics journal abstracts. Comparing abstracts to the research articles, Gillaerts and van de Velde discuss how abstracts have started functioning as "mini articles" (2010: 136).

Overall, the common IMRAD structure seems to be reflected in all studies listed above. Although it compensates with some detail of the move substeps, this is a good trade-off for a qualitative and quantitative analysis of the structure of 100 abstracts. Therefore, for this study, I will focus on the IMRAD structure and investigate how it is used in linguistics abstracts.

Still, it should be noted that it is difficult to make generalisations on the structure of abstracts. Broadly speaking, the hard and soft field scientists are shown to suggest credibility in their abstracts in different ways. While the soft sciences reference the community and discuss background knowledge from the literature, the hard sciences focus more on the methods and procedures (Hyland 2000: 83). The edited volume by Bondi and Lorés Sanz (2014) has provided evidence for other prominent cultural, disciplinary and diachronic variations of abstracts. For instance, comparing English and Italian abstracts, Diani (2014: 83) discerns a correlation between the size of the community and the use of moves – writers to larger audiences with more competition (English) focus more on the results whereas those are often omitted in the Italian sample. The personalisation differences observed by Diani (2014) are also evident in the study by Friginal and Mustafa (2017), who compare US and Iraqi PhD students' research article abstracts. They suggest that non-native speakers with Iraqi background express more distance through passive and non-personal constructions (Friginal – Mustafa 2017). In terms of discipline, Cavalieri (2014) shows that medicine abstracts focus more on the background (situating research move) than linguistics abstracts. Medicine researchers also use more personal patterns than linguists, putting the writer in the foreground (Cavalieri 2014: 174). In terms of diachronic changes in Economics, Linguistics and History journal article abstracts from 1990 to 2010, Bondi (2014) shows an increase in personal markers putting an accent on the author (*we*). The study also demonstrates the growing need to

prove significance and novelty by using modals and evaluative adjectives. All these "voice markers" are shown to underscore the individuality and authority in writers' contributions (Bondi 2014: 268).

This paper will therefore also limit its focus in terms of culture, discipline and time. It examines the changes in abstract writing by authors with German university affiliations in English-medium linguistic journals during the last 30 years.

3. Corpus and method

3.1 The Web of Science database

The corpus used for this study was compiled from the Web of Science (WoS) online database. The database was filtered for the following journals: *Linguistics*, *Journal of Phonetics*, *Cognitive Linguistics*, *Applied Psycholinguistics*, *Language and Speech*, *Phonetica*, *Linguistic Review*, *Applied Linguistics*, and *Language Learning & Technology* (*LLT*). The 15,130 entries were batch exported in a rich CSV format, which also included the available abstracts. The distribution of paper entries per journal correlated with the age of the journal – the journals with the oldest first entry in the WoS database logically had more entries (see Table 1).

The author information section was filtered to include "Germany" as part of the author affiliations. Although this does not exclude authors from non-German background and international multiple-author papers, it can be expected that national academic standards are reflected in the papers and their abstracts. Since this research focuses on broader rhetorical structure, which is likely to be influenced by national conventions, and does not focus on foreign language-specific features like grammar, this limitation should not largely interfere with the validity of the results. It is unfortunately impossible to isolate a true national German style because most scholars have international experience, so research in academic writing is often influenced by the interference of these confounding variables. Still, the 555 abstracts in this study should provide a relatively comprehensive overview of the rhetorical tendencies of researchers from Linguistics departments in Germany. Future studies can attempt to replicate these results by manually filtering the data to include a sample only with works by scholars affiliated with German institutions.

Table 1 presents the distribution of German papers per journal and the proportion of papers to abstracts. The table also describes the whole corpus

for comparison and shows that the German sample accounts for only 5.66% of all entries in the linguistics WoS database. Thus, the WoS database has great potential for future research on journal abstracts from other countries. The WoS database is a representative example of the new technological affordances and societal demands in academic publishing (Schmied et al., this issue) – the database links a vast number of publications from the numerous individual journal portals and thereby allows researchers to get a quick overview of the rapidly growing field.

Table 1. Journals and German abstracts in the collected WoS corpus

Journal	since	N D	% journals D	N abstracts D	% abstracts D	N all	% all	% D / all
Linguistics	1969	200	33.73	171	85.5	2627	25.08	1.91
Journal of Phonetics	1993	111	18.72	111	100	1576	15.05	1.06
Cognitive Linguistics	1998	61	10.29	61	100	1461	13.95	0.58
Language and Speech	1991	59	9.95	59	100	1283	12.25	0.56
Applied Psycho-linguistics	1992	54	9.11	54	100	1127	10.76	0.52
Phonetica	1972	48	8.09	40	83.33	977	9.33	0.46
Linguistic Review	1995	41	6.91	40	97.56	561	5.36	0.39
Applied Linguistics	1992	10	1.69	10	100	473	4.52	0.10
Language Learning & Technology	2003	9	1.52	9	100	389	3.71	0.09
		593		555		10474		5.66

However, one limitation of the WoS database should be noted – a substantial part of existing abstracts has not been compiled, in particular in the journals *Applied Psycholinguistics*, *Language and Speech*, *LLT*, and *Journal of Phonetics*. Some journals like *Phonetics* have included abstracts from their first issue release, but many of them are not included in the database. A lot of these problems arise when publishers provide only PDFs as this complicates

the automatic creation of paper entries. In order to provide a truthful representation of the abstract distribution throughout the years, I manually checked if the missing abstracts exist online and noted this in the corpus. Future studies can use web scraping and character recognition techniques to complete the WoS database.

In order to increase the representativeness of the abstract distribution overview (Section 4.1), I filtered out genres which typically do not have an abstract: Bibliography, Biographical Item, Book Review, Correction, Item About an Individual, Note, Meeting Abstract, Editorial Material, Letter, News Item, Software and Hardware Review. For the whole corpus, these were 10,474 out of 15,130 papers (31%). For the German section, the 137 filtered papers accounted for 19% of the corpus. They had only 5 abstracts, which proves that these genres typically do not have abstracts. The remaining 81% (583 articles and 10 review articles) constituted the 593 texts in the 100,266-word corpus of journal abstracts written by German authors in English (WoS-D).

3.2 Methodology

The whole WoS-D corpus was tagged with Part of Speech (PoS) tags and parsed in terms of dependencies using spaCy (Honnibal et al. 2020) in Python (Van Rossum – Drake 2009). This was done in order to easily determine the most frequent subjects and verbs in the corpus and analyse the distribution of these metadiscourse markers. The en_core_web_trf language model used has high accuracy in PoS-tagging (98%) and dependency parsing (94%) (spaCy 2022) which is sufficient for an initial overview of the most widespread subjects and verbs. The tagged data was lemmatised with the *textstem* package (Rinkler 2018) in R (R Core Team 2020) for corpus analysis. Again, the scope of this paper does not allow an analysis of many potentially interesting tendencies in the distributions of parts of speech and dependencies. These can be explored in future studies focusing more on metadiscourse features.

For the manual analysis of moves, a random sample of abstracts (n = 100, see Appendix) was drawn with R and manually tagged with INCEpTION (Klie et al. 2018) with active learning assistance from the built-in Sentence Classifier (OpenNLP Document Categorizer). The tags used corresponded to the IMRAD categories: Introduction, Method, Results, and Discussion. In addition, the tags Research Questions (RQ), Conclusion, and Limitations were also used in order to test whether they are part of the abstract genre. The output was converted from XML/XMI to TSV in Python

(Van Rossum – Drake 2009) and analysed with R (R Core Team 2020) and the *tidyverse* packages (Wickham et al. 2019). Finally, a qualitative analysis of the abstracts explored the move distribution in a selection of cases in more detail.

4. Results and discussion

4.1 Overview

The raw distribution of German English papers with and without abstracts in Fig. 1 clearly shows the rise of the abstract in journal articles. The abstract gained popularity in the 90s and has become a standard part of publications. This is also confirmed by Fig. 2, which shows the percentage of German English papers with and without abstracts per year. The error bars show the standard error of each bar and depict how uncertain its information is – the smaller the error bar, the smaller the uncertainty and the higher the reliability. Before the 90s, papers without abstracts have smaller error bars than those with abstracts, which used to be the exception. Then, papers from 1990-1996 all have abstracts but also large error bars because the visualisation is made based on little data. Finally, from 2000 to 2021, the few papers without an abstract have large error bars and the papers with an abstract have small error bars. This shows that the papers with abstracts have become a natural part of journals.

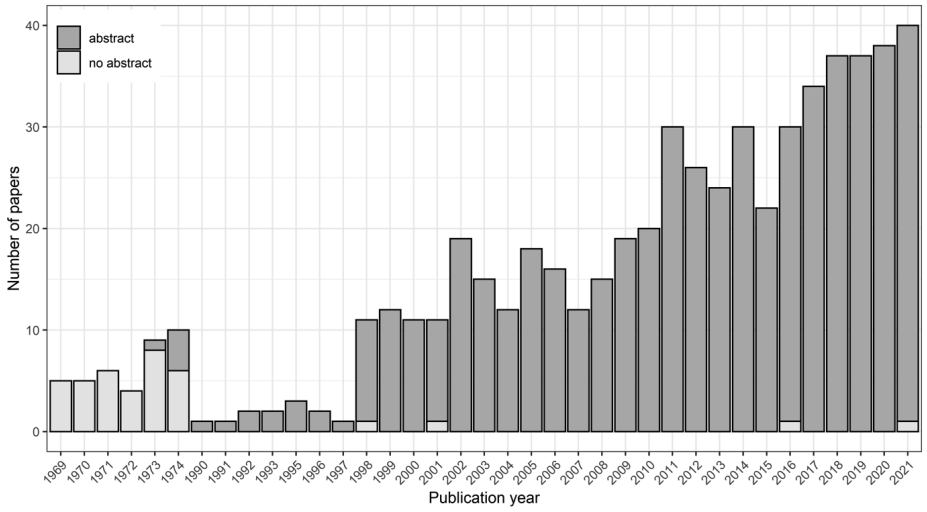

Figure 1. Number of papers with abstracts per year

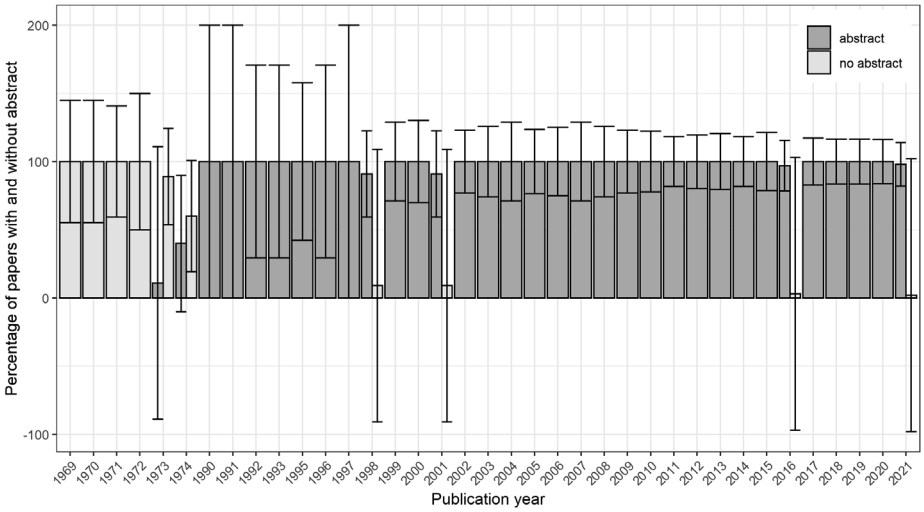

Figure 2. Percentage of papers with abstracts per year

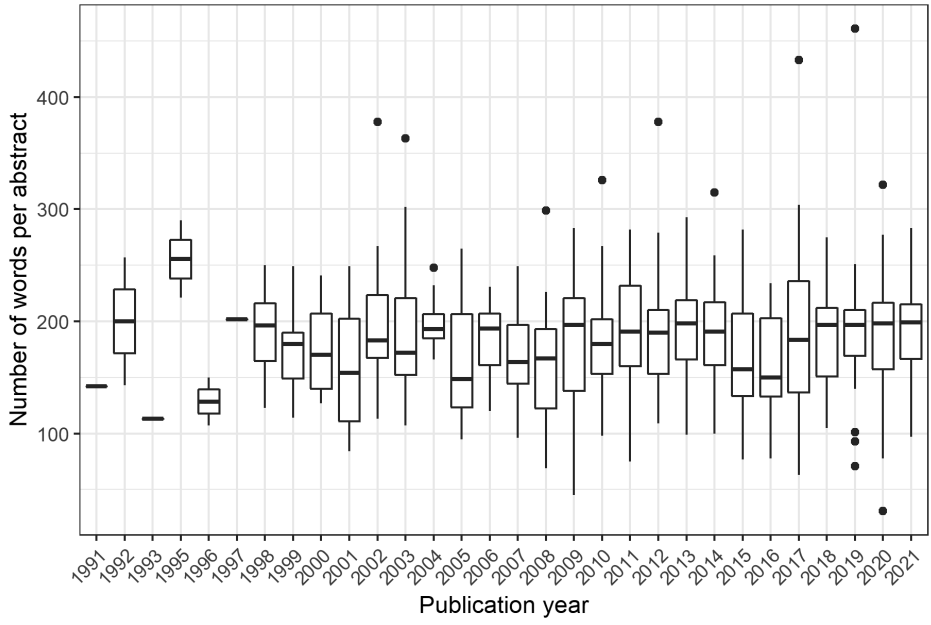

Figure 3. Word count of German English abstracts per year

Fig. 3 visualises the changes in word count of German English abstracts per year. Here it is again obvious that abstracts have become more uniform throughout the years. While in the 90s the median abstract length could

differ between 100 and 200 words, in the last four years, abstracts have had the same median worth length of 200 words. The data also partly reflect the established "increase in the average length of the RA abstracts" (Gillaerts – van de Velde 2010: 136). However, the increasing standardisation in word length is a more prominent observation here.

4.2 Rhetorical structures

4.2.1 Quantitative results

The manual annotation of 100 abstracts with one move per sentence gives a comprehensive overview of the tendencies in abstract global rhetorical structure. Fig. 4 presents the raw frequency distribution of IMRAD moves in the annotated sample (full overview of move annotations in the Appendix).

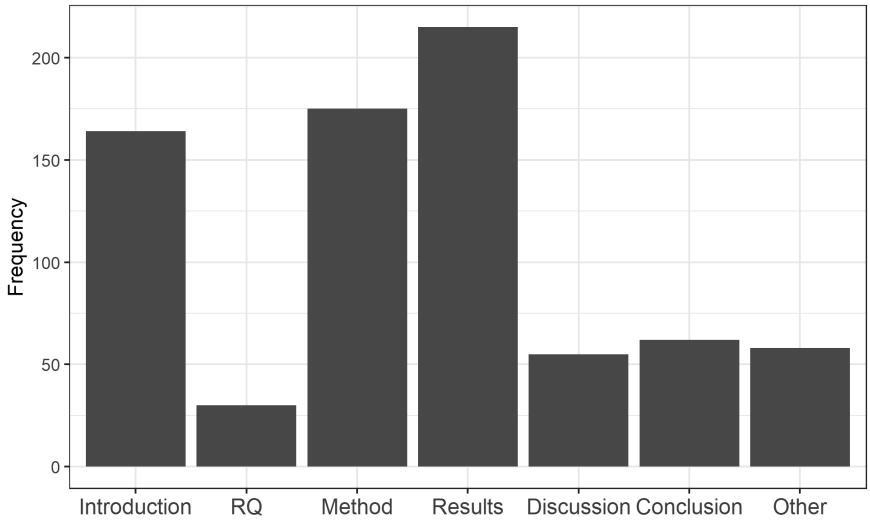

Figure 4. Distribution of Moves in the manually annotated sample (n = 100 abstracts)

The most frequent structure is Results, followed by Methodology and Introduction. This shows that the authors aim to create a research space and address a specific scientific community with the details described in their methodology. Authors summarise the results in the results section but leave limited discussion and conclusion of the findings. There is a nearly equal share of the Discussion and Conclusion moves, which are two of the most infrequent categories. This goes along with the advertising function

of abstracts, i.e. authors use this genre as a teaser to promote their full publications. The category "Other" contains irrelevant text like journal copyright and citations. Research questions or aims have the smallest share on the one hand, because they are mostly implied in the introduction and on the other hand, because of the short abstract length. There were no Limitations sections, which shows that this critical category is reserved for the article. There, authors have an opportunity to clarify the rationale for their decisions, to justify the limitations, and to provide suggestions for future research. The economic size of the abstract does not leave space for these considerations, which would considerably increase the face threat towards the authors and reduce their credibility. Thus, abstracts are left with the summarizing and advertising functions.

Looking at the percentage of IMRAD moves per year in the annotated abstracts (Fig. 5), it becomes evident that the structure of abstracts has become more uniform throughout the years.

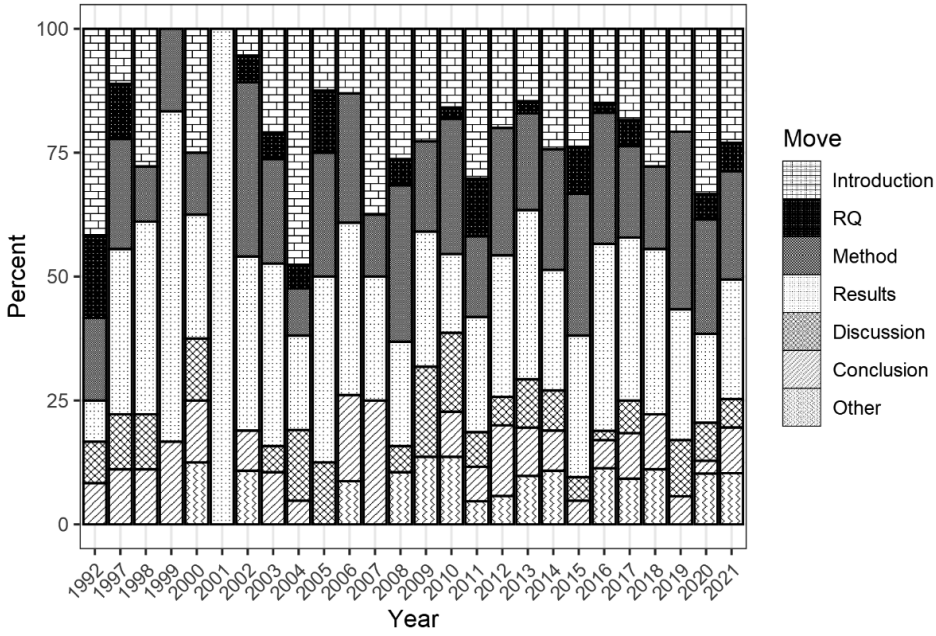

Figure 5. Percentage of Moves per year in the manually annotated sample (n = 100 abstracts)

In the 90s, there are several outliers, e.g. the abstracts from 1999 and 2001 show mostly just Results. However, in the last five years, the moves have had more uniform shares. These results mirror the findings in Hyland

(2000), where moves are shown to become more frequent and more diverse between 1980 and 1997 (Hyland 2000: 82). Similarly, Dontcheva-Navratilova (this issue) also observes a rise of moves in Czech journal articles, showing that academic writing has generally become more complex but also with a standardised, replicable, and thereby easy-to-follow structure. The strong increase in Move 3 "Describing the methodology" in English-medium journal abstracts from *Applied Linguistics* and *English for Specific Purposes* (*ESP*) observed by Diani (this issue) is not that prominent here, but rather a more uniform distribution of moves throughout time. Still, *Applied Linguistics* accounts for only 10 abstracts in my corpus and two abstracts analyzed qualitatively, and *ESP* was not featured, thus a replication study with more samples could provide interesting insights on journal-specific effects.

Fig. 6 shows a comparison of the moves across journals. Although some of the journals are not strongly represented in the database (e.g. *Applied Linguistics* has only two abstracts in the sample), the charts provide a preliminary overview of some of the rhetorical tendencies in the different journals.

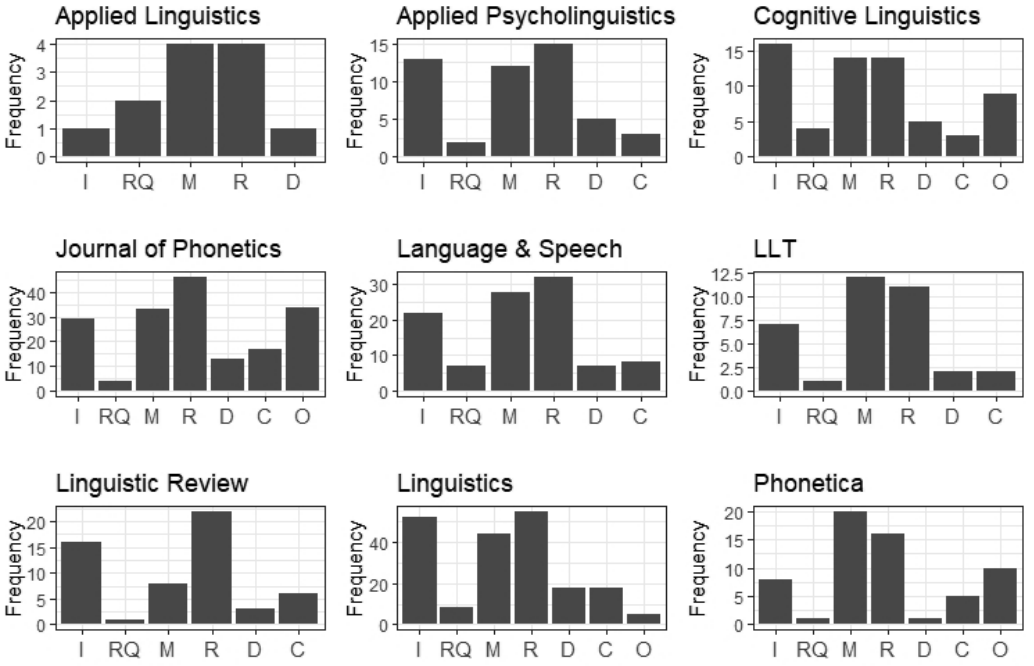

Figure 6. Distribution of moves per journals

For instance, *Linguistic Review* abstracts have the fewest Method moves, perhaps because the journal advertises a focus on theoretical and Generative Grammar articles (*The Linguistic Review* 2022), which already implies some of the methods. Meanwhile, the journals with explicit focus on phonetics (*Journal of Phonetics*, *Language and Speech*, *Phonetica*) share a similar move distribution with few Introduction moves and many Method and Results moves. A similar distribution is evident in *Applied Psycholinguistics* and *LLT* too. *Cognitive Linguistics* and *Linguistics* have the highest shares of Introduction moves, which reveals that their studies need to establish more background knowledge in the field.

4.2.2 Qualitative results

This section looks at qualitative differences between the abstracts throughout the years. As shown in the previous section, Moves have become uniform. For instance, in 2001 there is a *Linguistics* paper abstract which has only Results Moves in its two sentences:

(1)

a) [Results]: This paper argues that the common denominator of topic-comment (TC) constructions in natural languages is not a single functional feature (e.g. aboutness) but rather the fact that they share some salient semantic attributes with prototypical examples of TC.

b) [Results]: The paper tries to disentangle these prototypical semantic attributes of TC and then shows, mainly referring to examples from German, that different TC constructions are characterized by different combinations of these attributes and therefore cannot be analyzed properly by unitary theories of TC function.

In contrast, a 2020 abstract from the same journal incorporates all IMRAD moves in its nine sentences:

(2)

a) [Introduction]: Numerous crosslinguistic studies on motion events have been carried out in investigating the scope of the two-fold typology "path versus manner" (Talmy 1985, 2000) and its possible implications.

b) [Introduction]: This typological contrast is too narrow as it stands, however, to account for the diversity found both within and across types.

c) [Method]: The present study is based on what can be termed a process-oriented perspective.

d) [Method]: It includes the analyses of all relevant conceptual domains notably the domain of temporality, in addition to space, and thus goes beyond previous studies.

e) [Method]: The languages studied differ typologically as follows: path is typically expressed in the verb in French and Tunisian Arabic in contrast to manner of motion in English and German, while in the temporal domain aspect is expressed grammatically in English and Tunisian Arabic but not in German and French.

f) [Method]: The study compares the representations which speakers construct when forming a reportable event as a response to video clips showing a series of naturalistic scenes in which an entity moves through space.

g) [Discussion]: We assume that each of these three cognitive categories is shaped specifically by language structure (both system and repertoire) and language use (frequency of constructions).

h) [Results]: The findings reveal systematic differences both across, as well as within, typologically related languages with respect to (1) the basic event type encoded, (2) the changes in quality expressed, (3) the total number of path segments encoded per situation, and (4) the number of path segments packaged into one utterance.

i) [Results]: The findings reveal what can be termed language-specific default settings along each of the conceptual dimensions and their interrelations which function as language specific attentional templates.

One limitation of this annotation method becomes evident from this comparison – annotating with one move per sentence is difficult because there can be several overlapping moves in a sentence. In example 2, the move is labelled "Results", but it combines information both on the results and the methodology. Move e) clarifies the methodology but also the theoretical background behind it. The methodological issues of move repetition, embedding and identification in genre analysis have also been recognised by Gillaerts (2013: 53).

As can also be seen from the examples above, moves span across several sentences. Many abstracts do not show the linear move structure typical for the sections of papers, but have a shifted order, as in f), g), and h) – Method-

Discussion-Results. There are some examples of abstracts with dynamic moves where the results are interrupted by introduction and method statements for clarification, especially when there are multiple experiments with different methods and results. Dos Santos (1996: 497) similarly identifies differences in the move size (move balance), move blending (move embedding), and move sequence reversal as the main features of abstracts that are used for emphasis of relevant parts and cohesion (Dos Santos 1996: 497). Hyland also observes articles "recycling" moves when several applied methods generate results (Hyland 2000: 69).

As in Gillaerts (2013: 53) considerations, I found it difficult to put some of the moves in definite and unambiguous categories. The Discussion move g) is actually an assumption, but since it is provided with relation to the methods and the results, it was classified as a discussion item.

So far, I have focused on the diachronic changes in structure. However, it should be noted that analyses of global rhetorical structure should also consider the article's approach. Some recent theoretical, methodological or typological studies like the following 2021 *Linguistics* abstract continue using limited moves:

(3)

a) [Introduction]: Many researchers seem to think that Construction Grammar posits the existence of only wholly idiosyncratic constructions.

b) [Introduction]: However, this misconception betrays a deep misunderstanding of the approach because it glosses over the fact that constructions rarely if ever emerge sui generis.

c) [Introduction]: Rather, Construction Grammar aims to balance the fact that some linguistic uses cannot be fully predicted from other well-established uses with the fact that extensions of a construction, while not predictable, are motivated by other senses in the constructional network.

d) [Method]: This paper illustrates this idea by providing an analysis of the Spanish completive reflexive marker se.

The abstract builds up the background and concludes with one move on the case study. Again, d) in example 3 contains overlapping moves and presents a dilemma to the annotator. It combines method and results with a reference to the background (*this idea*), so both labels could fit the sentence. However, it has been labelled "Method" due to the emphasis on the approach. There

are several other recent abstracts with homogeneous moves in the database, which shows that the approach of the article should also be considered. This idea was also discussed with relation to the hard and soft sciences in Hyland (2000: 70). It can be generalized that empirical articles tend to have a clear IMRAD structure while theoretical, methodological and typological articles often omit moves and focus on those rhetorical structures which suit their purpose.

4.3 Selected metadiscourse features

The choice of subjects and verbs can reveal a lot about the communicative goals of the abstracts. Table 2 shows the top 10 subjects and verbs (roots) in the corpus. No diachronic change in the use of these features was observed, which indicates they are core academic discourse features of linguistics abstracts regardless of the publication period. The most frequent subject is the personal pronoun *we*, which expresses authorial agency. The language model has also coded relative clauses (*that*, *which*) as subjects and they are the second most frequent item, showing that abstract sentences can be relatively complex. The expletive construction or personal pronoun *it* is also popular as it allows researchers to use passive voice or to refer to non-human subjects like the frequent nouns *results*, *paper* and *study*. These nouns put agency on an inanimate object and distance the author from the contribution. The first-person pronoun *I* is relatively low on the list with 97 occurrences, showing 1) the predominantly collaborative nature of journal article writing and 2) single authors' preference to deemphasize their individual contribution. These results go along with Bondi's (2014) findings of the growing personalization through the use of self-referential *we* in Linguistics and Economics articles (Bondi 2014: 254, 257). The locational self-referential nouns like *study* and *paper* were also shown to be prominent (Bondi 2014: 254, 257f).

In terms of the verbs, in addition to the verb *be*, the most frequent verbs are *show*, *argue,* and *suggest*. These verbs are typically carriers of *that*-complement clauses, which are a widespread structure in academic writing (Hyland – Tse 2005; Pho 2013). They incorporate extraposition (Hyland – Tse 2005: 42), which allows the author to provide evaluation of the statement in the main clause. For instance, *suggest that* is weaker than *show that* and *claim that* is more negative than *argue that* (Pho 2013). The most frequent verbs here are all positive or strong: *show* (positive strong), *argue* (ambiguous strong), *suggest* (positive weak), and *find* (positive strong). This reveals that authors focus on the advantages of their work and do not address their or other

studies' limitations in the abstract. This goes along the "linguistic positivity" trend, which suggests that academic writing uses more positively connoted words (Wen – Lei 2021: 20). The top verbs also confirm Bondi's observation of the widespread use of verbs as acts of topic-setting (*investigate*, here on place 5) and claim-making with growing claim-making verbs (*show*, *argue*, *suggest*, here on place 2-4) (Bondi 2014: 262). Similarly, Schmied (this issue) observed a rise in positive strong verbs like *show* in the recent Master's theses, confirming the linguistic positivity trend.

Table 2. Top 10 subjects and root in all abstracts of the WoS corpus (100,266 words)

	Subject	n	Verb	n
1	*we*	539	*be*	477
2	*that*	379	*show*	264
3	*it*	346	*argue*	125
4	*which*	254	*suggest*	107
5	*results*	197	*investigate*	102
6	*they*	194	*find*	96
7	*paper*	159	*reserve*	92
8	*study*	155	*discuss*	90
9	*I*	97	*present*	75
10	*speakers*	93	*examine*	73

All in all, the subjects and verbs show that authors resort to metadiscourse markers to put or take away focus from their individuality and to express their stance and attitude. Together with the rhetorical moves, authors combine global and local structures to promote their work.

5. Conclusion

Abstracts use their global and local structure to condense and advertise the article they introduce. This paper took entries from linguistics journals in the Web of Science database authored by scholars with affiliations to German research institutions and examined 593 papers with 555 abstracts as well as a sample of 100 abstracts quantitatively and qualitatively. It showed that abstracts have become a standard part of articles after the 90s. Abstracts have also become more uniform both in terms of word length and of IMRAD

structure. This goes along with the findings of Busch-Lauer, who noted that German and English native speakers show more awareness of genre norms and have "adapted to international standards" (Busch-Lauer 2014: 60).

In the field of linguistics, abstracts focus on background knowledge (introduction) but also specify their sample and analysis procedures in the methodology and highlight their results. The move distribution usually follows the IMRAD article structure, but authors often switch, combine, and omit moves in order to suit the purposes of their article or journal approach.

All these strategies aim to promote the study, following the "linguistic positivity" tendency identified in abstracts by previous corpus studies (Cao et al. 2021; Vinkers et al. 2015; Wen – Lei 2021). In this paper, linguistic positivity is demonstrated on two levels: the structural level, where there were no Limitations moves, and the metadiscourse level, where there are no negatively connoted verbs. Limitations are almost obligatory in journal articles but absent in abstracts, which also shows that the IMRAD structure of articles is not completely compatible with the structure of abstracts.

The observed developments in abstract length and structure raise the question of the extent to which they are a result of authorial choices and editorial prescriptions. This is difficult to address due to the occluded nature of peer review. Hyland judges from the variety of patterns in his sample that "how writers use such practices is not *determined* by editorial prescription or genre constraints" (Hyland 2000: 75, emphasis in original). Here it can also be claimed that since abstracts have become uniformly structured across nine journals, each with their own requirements, their structure is more indirectly influenced. Still, future studies can compare the abstracts with the official journal requirements to determine how thoroughly authors comply to them. The study can also be extended to abstracts written by scholars affiliated with research institutions in other countries. Such analyses can gain insights in the impact of editorial and cultural conventions on one of the most popular genres in academic writing.

Overall, German abstracts have become more unified towards the scientific Anglo-American IMRAD model. This general model, however, has some major differences from that of the articles. It has been adapted to the advertising function of the abstract where stronger emphasis is placed on the authors' contributions and article's importance (through the introduction, method and results) and limitations are rarely discussed. Thus, over the last 30 years, common academic writing structures like IMRAD have been adapted to fulfil genre-specific functions.

REFERENCES

Primary sources

Web of Science
2022 https://www.webofscience.com/wos/woscc/basic-search, accessed June 2022.

Secondary sources

Bhatia, V.
1993 *Analysing Genre. Language Use in Professional Settings.* London: Longman. https://doi.org/10.4324/9781315844992.

Bondi, M. – R. Lorés Sanz
2014 "Changing voices: Authorial voice in abstracts". In: M. Bondi – R. Lorés Sanz (eds.) *Abstracts in Academic Discourse: Variation and Change*. Bern: Peter Lang, 243-269.

Busch-Lauer, I.
2014 "Abstracts: Cross-linguistic, disciplinary and intercultural perspectives". In: M. Bondi– R. Lorés Sanz (eds.) *Abstracts in Academic Discourse: Variation and Change*. Bern: Peter Lang, 43-63.

Cao, X. et al.
2021 "Promoting science with linguistic devices: A large-scale study of positive and negative words in academic writing", *Learned Publishing* 34 (2), 82-88. https://doi.org/10.1002/leap.1322.

Cavalieri, S.
2014 "Variation across disciplines. The case of Applied Linguistics and Medicine". In: M. Bondi– R. Lorés Sanz (eds.) *Abstracts in Academic Discourse: Variation and Change*. Bern: Peter Lang, 161-174.

Diani, G.
2014 "On English and Italian research article abstracts: Genre variation across cultures". In: M. Bondi– R. Lorés Sanz (eds.) *Abstracts in Academic Discourse:Variation and Change*. Bern: Peter Lang, 65-83.

Dos Santos, M. B.
1996 "The textual organization of research paper abstracts in applied linguistics", *Text & Talk* 16 (4), 481-499.

Friginal, E. – S. S. Mustafa
2017 "A comparison of U.S.-based and Iraqi English research article abstracts using corpora", *Journal of English for Academic Purposes* 25, 45-57. https://doi.org/10.1016/j.jeap.2016.11.004.

Gillaerts, P.
2013 "Move analysis of abstracts from a diachronic perspective: A case study". In: N. – L. Johannesson – G. Melchers – B. Björkman (eds.)

Of Butterflies and Birds, of Dialects and Genres: Essays in Honour of Philip Shaw. Stockholm: Acta Universitatis Stockholmiensis, 49-60.

Gillaerts, P. – F. van de Velde
2010 "Interactional metadiscourse in research article abstracts," *Journal of English for Academic Purposes* 9 (2), 128-139. https://doi.org/10.1016/j.jeap.2010.02.004.

Honnibal, M. – I. Montani
2017 *spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing*. https://spacy.io/, accessed June 2022.

Hyland, K.
2000 *Disciplinary Discourses: Social interactions in Academic Writing*. Harlow, UK: Longman.

Hyland, K. – P. Tse
2005 "Evaluative *that* constructions: Signalling stance in research abstracts", *Functions of Language* 12 (1), 39-63.

Klie, J.-C. et al.
2018 *The INCEpTION Platform: Machine-Assisted and Knowledge-Oriented Interactive Annotation*. In: *Proceedings of System Demonstrations of the 27th International Conference on Computational Linguistics (COLING 2018)*, Santa Fe, New Mexico, USA. https://inception-project.github.io/, accessed June 2022.

Lorés Sanz, R.
2004 "On RA abstracts: From rhetorical structure to thematic organisation", *English for Specific Purposes* 23 (3), 280-302. https://doi.org/10.1016/j.esp.2003.06.001.

Pho, P. D.
2013 *Authorial Stance in Research Articles: Examples from Applied Linguistics and Educational Technology*. Basingstoke: Palgrave Macmillan.

R Core Team
2020 *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/, accessed June 2022.

Rinkler, T.
2018 textstem: Tools for Stemming and Lemmatizing Text. [Version 0.1.4]. https://cran.r-project.org/web/packages/textstem/index.html, accessed June 2022.

Salager-Meyer, F.
1990 "Discoursal flaws in medical English abstracts: A genre analysis per research- and text-type", *Text* 10 (4), 365-384. https://doi.org/10.1515/text.1.1990.10.4.365.

spaCy
2022 *Available Trained Pipelines for English*. https://spacy.io/models/en, accessed June 2022.

Swales, J. M.
1990 *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
2004 *Research Genres. Exploration and Applications*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139524827.
Swales, J. M. – C. B. Feak
2009 *Abstracts and the Writing of Abstracts*. Michigan: University of Michigan Press.
The Linguistic Review
2022 https://www.degruyter.com/journal/key/tlir/html?lang = en, accessed June 2022.
Van Rossum, G. – F. L. Drake
2009 *Python 3 Reference Manual*. CreateSpace. https://dl.acm.org/doi/book/10.5555/1593511, accessed June 2022.
Vinkers, C. H., et al.
2015 "Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: Retrospective analysis", *BMJ (Clinical Research Ed.)* 351, h6467. https://doi.org/10.1136/bmj.h6467.
Wen, J. – L. Lei
2021 "Linguistic positivity bias in academic writing: A large-scale diachronic study in life sciences across 50 years", *Applied Linguistics* 43 (2), 340-364. https://doi.org/10.1093/applin/amab037.
Wickham, H. et al.
2019 "Welcome to the Tidyverse", *Journal of Open Source Software* 4 (43), 1686. https://doi.org/10.21105/joss.01686.

APPENDIX

Overview of the random sample for manual analysis (n = 100)

Year	Journal	N	all D	%
1	2	3	4	5
1992	*Language and Speech*	1	2	50
1997	*Language and Speech*	1	1	100
1998	*Linguistics*	1	22	5
1999	*Linguistics*	1	20	5
2000	*Linguistics*	1	22	5
2001	*Linguistics*	1	18	6
2002	*Journal of Phonetics, Language and Speech, Phonetica*	3	26	12

1	2	3	4	5
2003	*Language Learning & Technology, Linguistics*	2	22	9
2004	*Language Learning & Technology, Linguistics*	2	18	11
2005	*Cognitive Linguistics, Linguistics*	2	22	9
2006	*Journal of Phonetics, Language and Speech, Linguistics*	3	20	15
2007	*Cognitive Linguistics, Linguistic Review*	2	13	15
2008	*Cognitive Linguistics, Linguistics, Phonetica*	3	19	16
2009	*Cognitive Linguistics, Journal of Phonetics, Linguistic Review*	3	23	13
2010	*Journal of Phonetics, Language and Speech, Linguistics, Phonetica*	4	22	18
2011	*Cognitive Linguistics, Journal of Phonetics, Language Learning & Technology, Linguistic Review, Linguistics, Phonetica*	6	30	20
2012	*Cognitive Linguistics, Language Learning & Technology, Linguistic Review, Linguistics, Phonetica*	5	28	18
2013	*Journal of Phonetics, Linguistic Review, Linguistics*	3	24	12
2014	*Applied Psycholinguistics, Cognitive Linguistics, Journal of Phonetics, Linguistic Review*	4	33	12
2015	*Language and Speech, Linguistic Review, Linguistics*	3	29	10
2016	*Applied Linguistics, Journal of Phonetics, Language and Speech, Linguistics, Phonetica*	5	34	15
2017	*Applied Psycholinguistics, Journal of Phonetics, Language and Speech, Linguistic Review, Linguistics*	5	36	14
2018	*Journal of Phonetics, Linguistics*	2	37	5
2019	*Applied Psycholinguistics, Language and Speech, Language Learning & Technology, Linguistics*	4	37	11
2020	*Applied Linguistics, Cognitive Linguistics, Journal of Phonetics, Linguistics*	4	40	10
2021	*Applied Psycholinguistics, Cognitive Linguistics, Journal of Phonetics, Language & Speech, Linguistic Review, Linguistics, Phonetica*	7	42	17

Overview of the manually annotated moves per abstract (n = 100)

Abs ID	I	M	RQ	R	D	C	Other
1	2	3	4	5	6	7	8
1	1	1	NA	2	NA	NA	NA
10	2	1	NA	3	NA	1	NA

1	2	3	4	5	6	7	8
100	2	4	NA	3	NA	NA	2
11	1	5	NA	2	1	NA	NA
12	2	NA	NA	2	NA	1	NA
13	NA	3	1	3	NA	NA	NA
14	1	2	NA	1	1	NA	7
15	1	2	NA	3	NA	1	NA
16	5	NA	1	1	NA	NA	NA
17	2	NA	NA	1	NA	1	NA
18	2	2	1	1	2	NA	NA
19	1	1	2	1	2	3	2
2	2	1	NA	3	2	NA	2
20	2	1	1	NA	1	1	NA
21	1	2	1	3	1	1	NA
22	2	2	NA	4	NA	1	2
23	1	NA	3	3	NA	1	3
24	1	3	NA	2	NA	1	2
25	1	NA	NA	3	NA	1	NA
26	2	2	1	1	NA	NA	NA
27	3	NA	2	1	NA	NA	NA
28	2	1	1	1	NA	1	NA
29	2	3	NA	2	NA	1	2
3	2	NA	NA	1	1	NA	NA
30	2	1	NA	3	NA	1	2
31	1	NA	NA	1	1	NA	NA
32	2	2	NA	3	1	1	NA
33	2	4	NA	3	NA	2	2
34	5	2	2	1	1	1	NA
35	4	1	NA	4	NA	1	NA
36	1	NA	1	4	NA	NA	NA
37	2	3	NA	1	NA	NA	NA
38	3	1	NA	NA	NA	NA	NA
39	4	1	NA	NA	NA	1	NA
4	2	2	1	1	NA	NA	NA
40	2	2	1	2	1	NA	NA
41	1	1	NA	1	NA	NA	NA

1	2	3	4	5	6	7	8
42	NA	1	NA	4	NA	1	NA
43	1	3	NA	1	NA	1	NA
44	NA	1	NA	4	1	2	NA
45	3	1	NA	2	NA	1	NA
46	NA	2	NA	3	NA	NA	2
47	3	4	1	1	NA	NA	NA
48	NA	2	NA	2	NA	NA	2
49	2	4	NA	2	1	NA	NA
5	1	3	NA	1	1	NA	NA
50	2	2	NA	1	1	NA	NA
51	NA	1	NA	1	NA	NA	2
52	NA	4	NA	1	2	NA	NA
53	1	3	1	4	NA	1	2
54	2	1	NA	1	1	1	2
55	3	2	NA	7	NA	NA	NA
56	2	2	NA	3	NA	NA	NA
57	1	2	NA	3	2	NA	2
58	3	1	NA	2	1	NA	1
59	1	2	NA	2	1	NA	NA
6	NA	NA	NA	2	NA	NA	NA
60	2	1	NA	1	1	1	NA
61	1	1	NA	1	1	NA	NA
62	1	1	1	4	NA	1	NA
63	2	4	NA	2	NA	2	NA
64	3	2	NA	1	1	1	NA
65	4	1	NA	1	NA	1	1
66	NA	2	NA	4	NA	1	NA
67	2	NA	NA	6	1	1	NA
68	2	1	NA	3	NA	2	2
69	NA	2	NA	2	1	NA	NA
7	2	2	NA	1	NA	2	NA
70	1	1	NA	2	NA	NA	NA
71	NA	1	NA	4	NA	1	NA
72	NA	4	NA	2	3	1	NA
73	2	2	NA	3	2	NA	NA

1	2	3	4	5	6	7	8
74	4	1	NA	3	NA	1	NA
75	4	2	NA	NA	3	NA	NA
76	1	3	NA	1	1	1	NA
77	NA	2	NA	2	1	1	NA
78	2	2	NA	1	1	NA	NA
79	NA	4	NA	6	NA	2	2
8	NA	4	1	4	NA	NA	2
80	1	6	NA	5	NA	2	NA
81	NA	3	1	3	1	1	NA
82	2	1	NA	2	1	1	1
83	3	2	NA	2	1	1	2
84	NA	1	NA	3	1	1	NA
85	1	1	NA	1	NA	1	NA
86	2	2	NA	5	NA	2	2
87	NA	1	1	1	1	NA	NA
88	3	1	1	2	1	NA	2
89	2	1	NA	1	NA	NA	NA
9	3	3	NA	3	1	NA	NA
90	1	3	NA	1	NA	NA	NA
91	4	1	NA	5	NA	NA	NA
92	1	1	1	1	1	NA	NA
93	NA	NA	NA	NA	NA	NA	1
94	1	2	NA	1	NA	1	NA
95	4	1	NA	1	NA	NA	2
96	2	1	NA	1	NA	1	NA
97	1	2	NA	1	4	2	2
98	1	1	NA	1	1	NA	NA
99	2	1	3	3	NA	1	NA

Address: Marina Ivanova, English and American Studies, Chemnitz University of Technology, 09107 Germany. marina.ivanova@phil.tu-chemnitz.de.
ORCID code: https://orcid.org/0000-0001-7414-6668.