Evaluative Meaning in Scientific Writing
Macro- and Micro-Analytic Perspectives Using Data Mining

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Philosophischen Fakultäten
der Universität des Saarlandes

vorgelegt von
Stefania Degaetano-Ortlieb
aus Darmstadt
Saarbrücken, 2015

Der Dekan: Univ.-Prof. Dr. Ralf Bogner
Erstberichterstatter: Frau Univ.-Prof. Dr. Elke Teich
Zweitberichterstatter: Herr Univ.-Prof. Dr. Erich Steiner
Drittberichterstatter: Herr Univ.-Prof. Dr. Augustin Speyer
Tag der letzten Prüfungsleistung: 18.12.2014

Abstract

In this thesis, we elaborate characteristics of evaluative meaning of different scientific disciplines and trace their diachronic linguistic evolution. A main focus lies on newly emerged disciplines, such as computational linguistics, which emerged through contact between two other disciplines, such as computer science and linguistics. Here, we consider (1) whether these newly emerged disciplines have created characteristics of their own over time, showing a process of diversification, and (2) whether they have also adopted characteristics from their disciplines of origin, reflected in a linguistic imprint, and if this might have changed over time. The newly emerged disciplines considered are computational linguistics, bioinformatics, digital construction and microelectronics, which have emerged through contact between computer science and a further discipline (linguistics, biology, mechanical engineering, and electrical engineering, respectively).

In terms of theory, this work is grounded in a linguistic theory rooted in sociolinguistics, Systemic Functional Linguistics (SFL; Halliday (2004)), which with its functional perspective on language allowed us to position evaluative meaning within a linguistic theory and to create a model of analysis to trace choices made in the semantic system on the level of lexico-grammar. Moreover, its notion of register, concerned with functional variation, i.e. variation according to language use, combined with the sociolinguistic perspective made it possible to compare the linguistic choices made according to different social contexts, to which the disciplines belong. This allowed us to trace register diversification processes and registerial imprint of evaluative meaning across disciplines.

In terms of methods, we apply classification as a data mining technique, taking a macro- and micro-analytic perspective (cf. Jockers (2013)) on the results. Doing so we gain insights on the degree of diversification and imprint (macro-analysis) and the kind of diversification and imprint (micro-analysis). Studies so far have considered either the macro- or the micro-analytic perspective. By considering both, we are able to investigate generalizable trends as well as detailed linguistic characteristics of evaluative meaning across disciplines and time.

The approach presented in this thesis draws its strength from being grounded in a linguistic theory, which proved to be extremely useful in defining and testing hypotheses and interpreting results. Moreover, an empirical analysis of evaluative meaning across disciplines and time was possible by combining corpus-based methods with data mining techniques.

Kurzzusammenfassung

In der vorliegenden Dissertation werden Bewertungscharakteristiken verschiedener Wissenschaftsdisziplinen erarbeitet und ihre diachrone linguistische Entwicklung untersucht. Ein Hauptfokus liegt auf in neuerer Zeit entstandenen Disziplinen (z. B. Computerlinguistik), die sich durch Kontakt zwischen zwei anderen Disziplinen gebildet haben (z. B. Informatik und Linguistik). In diesem Zusammenhang wird erforscht, (1) ob diese neu entstandenen Disziplinen diachron ihre eigenen Charakteristiken entwickeln und somit einen Diversifikationsprozess aufzeigen und (2) ob sie auch Charakteristiken der Ursprungsdisziplinen übernehmen und somit eine linguistische Prägung aus der Ursprungsdisziplin vorweisen und ob sich diese möglicherweise diachron verändert hat. Die untersuchten relativ neu entstandenen Disziplinen sind die Computerlinguistik, Bioinformatik, Bauinformatik und Mikroelektronik, die durch Kontakt zwischen der Informatik und einer anderen Disziplin entstanden sind, in unserem Fall entsprechend aus der Linguistik, Biologie, dem Maschinenbau und der Elektrotechnik.

Die Arbeit basiert auf der soziolinguistischen Theorie der Systemisch Funktionalen Linguistik (SFL; Halliday (2004)). Aufgrund ihrer funktionalen Perspektive auf die Sprache war es uns möglich, das semantische Konzept der Bewertung in eine linguistische Theorie zu positionieren und ein Analysemodel zu entwickeln, um die Auswahl aus dem semantischen System auf der lexico-grammatischen Ebene nachzuverfolgen. Besonders wichtig ist hierbei auch das Registerkonzept aus der SFL, das sich mit funktionaler Variation befasst, d.h. Variation in Bezug auf den Sprachgebrauch. Die Kombination aus funktionaler Variation und soziolinguistischer Perspektive hat es erlaubt, die linguistischen Entscheidungen in Bezug auf Bewertungen, die in unterschiedlichen sozialen Kontexten (d.h. den verschiedenen Disziplinen) gefällt wurden, zu untersuchen und diese zu vergleichen. Dadurch konnten für die untersuchten Disziplinen registerspezifische Diversifikationsprozesse und Prägungen bezüglich Bewertungen ausgemacht werden.

Methodisch wurde aus dem Bereich des Data Mining die Klassifikation angewandt, die es erlaubt hat, die Ergebnisse aus einer makro- und mikro-analytischen Perspektive (vgl. Jockers (2013)) zu erforschen. Dadurch konnten Erkenntnisse erlangt werden in Bezug auf den Diversifikations- und Prägungsgrad (Makro-Analyse) sowie der Art der Diversifikation und Prägung (Mikro-Analyse). Studien haben bislang entweder die makro- oder die mikro-analytische Perspektive angewandt. Durch den Einbezug beider Ebenen ist es uns gelungen, sowohl generalisierbare Tendenzen festzustellen als auch detaillierte linguistische Charakteristiken und diachrone Veränderungen von Bewertungsausdrücken in verschiedenen Disziplinen zu untersuchen.

Die Stärken des in der vorliegenden Dissertation präsentierten Ansatzes liegen darin, dass er in einer linguistischen Theorie fundiert ist, die sich sehr hilfreich erwiesen hat bei der Hypothesenaufstellung und beim Testen der Hypothesen sowie auch bei der

Interpretation der Ergebnisse. Darüber hinaus hat der Ansatz eine empirische Analyse von Bewertungen in wissenschaftlichen Disziplinen durch das Zusammenspiel von korpus-basierten Methoden und Techniken aus dem Data Mining ermöglicht.

Acknowledgments

At several stages during the last three years or so in which I worked on my thesis, I received enormous support from the following people:

First and foremost, I would like to thank Elke Teich for her great support during the production of this thesis as well as for her valuable advice and continuous encouragement throughout these years, for always taking time to go through things (even though there was no time) and for incredibly constructive comments. She taught me a lot about writing, teaching, and research. Thank you for showing me how to create a "story" out of research findings, how to always keep in mind the bigger picture without forgetting about the details, and how to become a passionate researcher.

Out of many colleagues who have helped me in various ways at different stages of my thesis, I'm really grateful for Hannah Kermes' helpful hints and suggestions, especially concerning technical matters, but also for inspiring chats on our way home, which I will really miss, and to Peter Fankhauser who taught me out of his immense knowledge invaluable things on data mining, the ways to keep things structured when doing research, and to challenge findings, especially if they look too good. I'm also indebted to the GECCo-Team for having given me the opportunity of gaining new inputs and for your patience and understanding during the final phase of this thesis, and to Margaret De Lap for proof-reading. Any remaining misconceptions and errors are, of course, solely mine.

A special thank you goes to my brother Tiziano Degaetano, whose spirit of invention and richness of ideas always inspires and encourages me. Thank you for the numerous opportunities you gave me to discuss my work whenever I needed, giving me invaluable input, I will miss our discussions over lunch and coffee breaks.

Heartfelt thanks go to my parents, "grazie di aver sempre avuto fiducia in me, qualsiasi direzione o decisione io prendessi e per il vostro immenso supporto senza il quale non sarei mai arrivata dove sono ora". A big thank you goes to my siblings Daniela and Domenico Degaetano, who always remind me to look at things from different perspectives, and my best friend Elnaz Allen, who always has a sympathetic ear for my complaints and encourages me to never give up. I would also like to thank my parents-in-law and further family and friends, "danke für Eure Geduld und Euer Verständnis in der Schreibphase, in der wir uns leider wirklich selten gesehen haben".

Last but not least, my deepest thanks go to Alex, for supporting me around the clock, for his willingness to listen to my thoughts about work and my complaints when things don't work the way I want, for reminding me to take time for myself and relax even though I think I can't, for always being there for me no matter what.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Goals

In the present research we aim to provide answers to selected issues regarding the evaluative characteristics of research articles and possible differences and commonalities between disciplines in this respect. The main questions we pursue answers for are:

- Do disciplines differ in the expression of evaluative meaning, i.e. are different lexico-grammatical features involved at the interpersonal level that vary across disciplines?
- Considering recent diachronic change
  - Have possible evaluative characteristics of disciplines changed over time, i.e. are the meanings expressed different and does a discipline use different lexico-grammatical features over time regarding evaluative meaning, or are the same features used?
  - Do newly emerged disciplines, such as computational linguistics, which emerged through contact between two other disciplines (linguistics and computer science), develop their own characteristics, or do they adopt the characteristics of their disciplines of origin, and if so, to what extent?

The leading hypotheses to these questions are a diversification of scientific disciplines over time in the expression or realization of evaluative meaning and a possible imprint of the disciplines of origin on the newly emerged discipline regarding evaluative meaning.

- *Diversification*:
  Scientific disciplines differ from one another in their patterns of language use. Linguistically and with respect to evaluative meaning, this is reflected in a diversified use of specific evaluative lexico-grammatical features. As scientific disciplines evolve, they will become increasingly distinct from one another. Linguistically and with respect to evaluative meaning, the clusters of evaluative lexico-grammatical features used will become more distinct.

- *Imprint*:
  As a discipline develops out of two other disciplines, it will adopt some characteristics of its disciplines of origin. Linguistically, specific lexico-grammatical features are similarly used by the newly emerged discipline and one of the disciplines of origin (e.g., computational linguistics and linguistics). Diachronically, depending on how the newly emerged discipline evolves, it might change the characteristics it adopts. Linguistically, while it uses specific lexico-grammatical features similarly to one discipline of origin in one time period (e.g., computational linguistics using features similarly to linguistics), it uses other features similarly to the other discipline it has originated from in another time period (e.g., computational linguistics using features similarly to computer science).

To approach these hypotheses, we have to consider several theoretical and methodological issues.
Regarding theoretical issues, first of all, we have to define what evaluative meaning is and how it is used in research articles. From a semantic perspective, which components do evaluative acts have? From the perspective of lexico-grammar, how are the components of evaluative acts expressed in terms of lexico-grammar, i.e. which are the lexico-grammatical features that realize evaluative meaning? By establishing what evaluative meaning is and how it is realized lexico-grammatically, i.e. which features are involved, we have a basis on which we can answer our hypotheses, but also more concretely investigate if disciplines differ in a similar way from each other or if different features are involved in the distinction. In terms of recent diachronic change, we will investigate whether there are some diversification trends of disciplines that can be traced over time regarding the use of evaluative meaning. Moreover, as new scientific disciplines develop, do they show more variation of features initially and develop a more distinct evaluative character over time?
Methodologically, we relate our work to approaches used so far to detect and analyze evaluative meaning, considering their strengths and weaknesses in order to develop a methodology suitable to investigate evaluative meaning in a quantitative way, yet accounting also for qualitative data. Here, we will develop an analytical cycle which encompasses a combination of macro- and micro-analytical steps. Moreover, we make use of (1) contrastive corpus-based approaches, in order to account for the hypotheses of *diversification* and *imprint*, i.e. contrasting text productions across disciplines and time, and (2) data mining techniques, as a methodology that allows

us to adopt a macro- as well as a micro-analytical perspective on the data under analysis.
This will enable us to carry out a register analysis of highly specialized scientific disciplines with respect to the use of evaluative meaning.

The main goal is to elaborate characteristics of evaluative meaning of different scientific disciplines and to trace their linguistic evolution. To achieve this, we develop a theoretical and methodological framework for the investigation of evaluative meaning in scientific written discourse and its diachronic development, focusing on recent change, by means of corpus-based and data mining techniques.

1.2 Motivation

The present study is motivated by the overarching question of whether scientific disciplines vary in their use of evaluative meaning, and if so, how they differ and whether they create their own characteristics over time. The focus lies on linguistic variation from a functional perspective, looking at the level of lexico-grammar, i.e. which lexico-grammatical features are used to evaluate within scientific writing and whether there are differences or commonalities between different disciplines in the use of these lexico-grammatical features. This implies considering four fields of investigation:

(1) the interpersonal function of language, focusing on evaluative meaning in scientific discourse,

(2) the notion of register, defined as variety according to language use with respect to 'fine-grained' text classes,

(3) diversification over time, and

(4) the methodological approach for tracing and analyzing evaluative meaning and diversification in highly specialized scientific fields.

While the first three points are related to the subject of investigation, i.e. the theoretical issues, the fourth point is related to the methodological issues, which run parallel to the theoretical ones.

As for the first point, *evaluative meaning* is a non-trivial phenomenon to grasp linguistically, as it can be expressed in a variety of linguistic forms and is highly context-dependent. This makes the phenomenon rather hard to be traced on a large-scale basis. There have been numerous studies on evaluative meaning that range from very detailed single-text-based studies to corpus-based and computationally based ones (e.g., Martin and White (2005); Hunston (2011); Wilson et al. (2005a) just to mention some major works for each approach; see Section 2 for a more detailed account). Nonetheless, we are still looking for a comprehensive picture of how evaluative meaning works in linguistic terms. The question of which methods might be best suited to move toward a comprehensive understanding of evaluative meaning is still a highly debated one. The challenge is to find a balance between (a) high-quality data, to capture the phenomenon of evaluative meaning as comprehensively as possible, and (b) quantitative methods, to allow making generalizations about the phenomenon. In this work, we propose a methodology that considers both aspects in order to move toward a comprehensive understanding of evaluative meaning. This can complement or inform corpus-based studies on evaluation as well as computationally based methods of sentiment analysis and opinion mining, but more generally it can also provide insights on how to find an appropriate balance between qualitative and quantitative perspectives on a linguistic phenomenon.

Regarding the first and second points, we raise the question of how *highly specialized disciplines* might differ in terms of evaluative meaning. This will possibly complement studies in the field of linguistic variation that move toward more fine-grained differences between registers, rather than similarities, as in genre studies that originated in the 1980s. In contrast to the idea of one general 'academic English', i.e. one monolithic entity that is merely differentiated by topics, scientific communities à la Kuhn (1962) differ not only in what they find worthwhile to communicate (topicality), but also how they communicate it, etc. This is related to the fact that disciplines are rooted in specialized social contexts and are thus prone to being influenced by specific contextual factors which can differ between disciplines, topicality being here only one of many features distinguishing between disciplines (Argamon et al., 2007; Teich and Fankhauser, 2010; Degaetano-Ortlieb et al., 2014a). Communication is a leading force in the dissemination of knowledge generated in the scientific communities and the majority of scientific communication is transmitted in the form of scientific research articles, which form the data basis of our study. We are interested in how knowledge is made persuasive, adding insights to cross-disciplinary variation on a more fine-grained scale as we focus on highly specialized fields, rather than on a more general distinction such as 'hard' and 'soft' sciences (see, e.g., Hyland (2000)).

As we put forward the hypotheses of diversification and imprint of scientific disciplines, we also have to consider the *time dimension* (point three above). The time dimension will help to understand whether scientific disciplines are relatively stable over time regarding evaluative meaning or if they change diachronically. Although in this study we focus on recent diachronic change, the insights gained might be of interest for historical sociolinguistics, which aims at reconstructing a language's past in order to account for diachronic linguistic changes and developments (cf. Hernández-Campoy and Conde-Silvestre (2012)).

For (2) and (3), the methodological approach should be a comparative one that allows comparing linguistic realizations of evaluative meaning across disciplines and time. The elaboration of a methodology to approach the hypothesis of diversification of highly specialized scientific disciplines in terms of evaluative meaning is of interest for all of the related fields mentioned above. However, cross-linguistic studies might also profit from the insights gained from this work.

1.3 Approach

To approach the above-raised hypotheses of *diversification* of scientific disciplines and *imprint* left on newly emerged disciplines by their disciplines of origin, we consider previous work on the interpersonal function of language, register studies, and recent diachronic change.

The interpersonal function of language is one of the three functions of language described by Systemic Functional Linguistics (SFL; Halliday (2004)): ideational (topic of a discourse), interpersonal (roles and attitudes of participants in the discourse), and textual (textual presentation of the discourse). Evaluative meaning belongs to the interpersonal function of language, as evaluations presented in a discourse are associated with the participants and their roles and attitudes. In the case of scientific writing, the writer of a research article evaluates abstract and concrete objects which are part of the disciplinary discourse. Thus, evaluative meaning is expressed in evaluative acts and these acts have particular semantic components, which consist of the writer of a research article, the evaluative meaning expressed by the writer toward a target, the target itself, and the reader of the research article. Note that the writer could be explicitly mentioned in the discourse, but does not have to be. Consider Example (1), where the writers evaluate *competitive analysis* as being important. In this example, it is quite clear that the writers present their own belief. However, evaluative expressions such as in Example (2) can also be found, where the writer is not explicitly mentioned in the discourse. The focus of the utterance seems to be more on the evaluation itself than on who expresses it. Besides that, the evaluative expressions shown in (2) are quite down-toning (e.g., evidence suggests rather than proves, and the modal verb *can* is used). While (1), which originates from a computer science text, is almost perceived as a fact put forward by the writers, (2), which is taken from linguistics, is more related to reasoning. These short examples already seem to suggest that disciplines vary in their use of evaluative meaning, and we can also see that different evaluative components are used (e.g., different evaluative word forms, different evaluative meanings, different evaluative lexico-grammatical patterns) which are realized lexico-grammatically.

(1) ***We believe*** *that* ***competitive analysis*** *gives* ***important*** *insight to the performance of these queuing policies.* (Computer science)

(2) ***There is*** *also experimental* ***evidence suggesting*** *that irregular verbs* ***can*** *be semantic attractors, but it is* ***unclear*** *whether this is the result of idiosyncratic analogy or whether it is grounded in the language system itself.*

> *The aim of the present article is to* ***argue*** *that indeed the local attraction effects are grounded in subtle systematic distributional differences in semantic density between regulars and irregulars.* (Linguistics)

In the course of this study, we will try to capture how evaluative meaning is expressed in scientific writing, relying on previous work (e.g., Hunston (1989, 1994, 2011); Hyland (1998, 2000, 2005)), but also on insights gained in our own previous research (Degaetano and Teich, 2011; Degaetano-Ortlieb et al., 2014b). Here we aim to create a model of how evaluative meaning is expressed in scientific writing by looking at lexico-grammatical features that are used to evaluate. What we do not intend is to develop a detailed system as in Martin and White (2005), which was originally meant for the detailed annotation of single texts. Our aim is to be able to make generalizations on a more abstract level for possible differences across disciplines and to determine whether these differences hold over time. Thus, we have to develop a model that allows for these kinds of generalizations.

Besides SFL, our work is also rooted in register theory (Quirk et al., 1985; Biber et al., 1999), which has shown in numerous studies that particular situational contexts have linguistic correlates at the level of lexico-grammar giving rise to registers, i.e. clusters of lexico-grammatical features which occur non-randomly (most prominently the work of Biber and colleagues, e.g., Biber (1988, 1993, 2006, 2012)). We adopt this by considering that particular distributions of lexico-grammatical features associated with evaluative meaning might characterize different scientific registers. As language use continuously adapts to changing social contexts (Ure, 1971, 1982) and considering that social contexts are reflected linguistically in specific distributions of lexico-grammatical features, we should also be able to trace differences by comparing the use of lexico-grammatical features across time periods. Social contexts that are possibly subject to recent change in the scientific domain are newly emerged disciplines, such as bioinformatics and computational linguistics. Diachronic studies on registers are quite sparse (Halliday, 1988; Banks, 2008) and linguistic studies on recent diachronic change have investigated mainly the Brown corpus family (cf. Kučera and Francis (1967); Hundt et al. (1999)), yet the use of linguistic features to investigate recent diachronic change is also gaining interest in social and behavioral science (e.g., Danescu-Niculescu-Mizil et al. (2013)).

In addition to the theoretical considerations, dealing with evaluative meaning (or interpersonal meaning in the general sense), registers and recent diachronic change implies various methodological aspects that have to be considered:

- *Comparative corpus linguistics*
  The investigation of diversification trends in the scientific domain requires a comparative approach which allows the comparison of text productions belonging to different disciplines. Obviously, a corpus is needed that includes text productions from several disciplines. To trace diversification over time, disciplines that might be subject to diachronic change have to be considered

(such as bioinformatics) and the corpus has to be diachronic, i.e. represent at least two time periods.

- *Micro- and Macro-Analyses*
  Detailed analyses are needed to investigate which evaluative lexico-grammatical features are employed in scientific writing. This is accomplished by using sub-corpora and concordances which are analyzed in detail. However, to be able to make generalizations on diversification trends, large-scale results are needed. These are obtained by techniques from information extraction and data mining. Methodologically, we consider micro- and macro-analytical techniques (cf. Jockers (2013)) that will allow us to find a balance between detailed analyses and generalizations on the register level. An analytical cycle is designed with possible recursive steps when switching between micro- and macro-analysis.

1.4 Sections overview

Chapter 2 introduces some major research strands relevant to the present context. We discuss in more detail work on evaluative meaning, register theory and diversification over time moving toward presenting methodological approaches used so far to analyze evaluative meaning. Strengths and weaknesses are discussed that will serve as input on the one hand, and motivate our own research on the other.

The aim of Chapter 3 is to introduce a model of evaluative acts grounded on semantics and lexico-grammar. This will help to position evaluative acts within linguistic theory to allow explanations of the results.

In Chapter 4, two major methodological aspects are presented: (1) the analytical cycle used to approach our hypotheses of diversification and imprint, which will present in more detail the interplay between macro- and micro-analyses, and (2) the corpus-based methods, i.e. identification, annotation and extraction of evaluative lexico-grammatical features, and data mining techniques employed to allow data analysis.

Chapter 5 presents the register analysis, which will start with a macro-analytical perspective and move toward a micro-analytical one for each of the two hypotheses investigated. While macro-analysis encompasses text classification, micro-analysis is concerned with feature analysis. In the first analysis, we macro-analytically investigate whether highly specialized scientific disciplines undergo a diversification process over time. Micro-analytically, we investigate which lexico-grammatical features are involved in diversification, i.e. which features are characteristic of a discipline. In the second analysis, we exploit, on a macro level, the overlap between newly emerged disciplines and their disciplines of origin, according to the usage of evaluative meaning in general. In micro-analytical terms, we look at which features are adopted by the newly emerged disciplines from their disciplines of origin.

Finally, Chapter 6 gives a summary of the major aspects addressed in this work and an assessment of the approach applied, and concludes with an envoi and possibilities for future work.

Chapter 2

State of the art

2.1 Introduction

The aim of the present chapter is to give an overview of some major research strands that are relevant to the present context. The theoretical groundwork on evaluative meaning, register theory and diversification over time is laid. This will allow us to develop a model of evaluative meaning that can be applied to the research hypotheses raised.

In very general terms, this study is related to discourse analysis (Stubbs, 1983; Schiffrin et al., 2003), register studies and sociolinguistics (Ure, 1971, 1982; Quirk et al., 1985; Biber et al., 1999; Biber, 1995, 2012; Ferguson, 1994), as well as historical sociolinguistics and studies on recent language change (Hernández-Campoy and Conde-Silvestre, 2012; Mair, 2006, 2009), as they contribute to the knowledge of how language works, how linguistic expressions are associated with meanings, how speakers/writers indicate their semantic intentions and how hearers/readers interpret what they hear/read as well as how changing social contexts influence linguistic choices and vice versa. Moreover, studies related to scientific discourse, especially those on variation in scientific discourse (Bazerman, 1981; Ventola, 1996; Duszak, 1997; Bhatia, 2002; Hyland, 2007, 2009a), are implicated as they consider discourse, social and cultural context as well as variation and change in the academic context.

In specific terms, this study concerns itself, in particular, with research which directly focuses on the questions raised in this thesis or which is particularly relevant to the hypotheses and methodologies addressed. Thus, the following sections focus on selected contributions to register studies concerned with variation and language change related to scientific writing (Ure, 1982; Halliday and Hasan, 1985; Halliday, 1988; Swales, 1990; Teich et al., 2012) as well as descriptive accounts and approaches to evaluative meaning (Hunston, 1989, 2011; Martin and White, 2005; Hyland, 2005).

2.2 The research article and its developing evaluative character

Research articles as we know them today have their roots in experimental reports, most notably in *The Philosophical Transactions of the Royal Society*, the first scientific periodical, established in 1665. Since then, the research article has undergone many changes before arriving at the stage of what we today recognize as a research article. To begin with, according to Swales (1990), the article length has varied from time to time, with 7000 words in 1893 to 5000 words in 1900 and an increase since then to 10,000 words by 1980. References were common but general in the time period of 1893-1900. By 1910 reference use decreased, yet more recent references were used, including reference dates, and they were of direct relevance to the research. From then on an upward trend was observed and, furthermore, references were no longer concentrated only in the introduction section (cf. Swales (1990)). What was also observed was a decline in the citation of books and an increase in the citation of shorter works. Syntactic and lexical features changed as well: relative clauses declined in frequency, descriptions shifted to explanations, subjects of main clauses became more abstract, reporting verbs decreased whereas active verbs increased, with the aim of bringing the finding/theory into a central grammatical position. Authors, instead, have been given a back seat (e.g., *Here, I report the performance of* vs. *The algorithm performs*). Regarding the organization of research articles, before 1950 only 50% were divided into sections. After 1950 section headings became a regular feature. From 1930 an inclusion of discussion and conclusion sections and an increase in length and complexity was observed. According to Swales (1990), the overall trend regarding the development of the research article is a growing abstraction, the deepening integration of present work with the relevant literature, the increasing foregrounding of research as opposed to researcher, the increasingly uphill strive to incorporate more and more information, and a steadily more focused argumentation.

Additionally, trends have been observed in which the introduction and discussion/conclusion sections have undergone similar developments, whereas the methodology and results sections seemed to remain unchanged. Some of the main developments within the introduction and conclusion sections are also attributed to evaluation. As Hill et al. (1982, 335) put it, "research papers make the transition from the general field or context of the experiment to the specific experiment by describing an inadequacy in previous research that motivates the present experiment. The Discussion section mirror-images the Introduction by moving from specific findings to wider implications". Also prominent in these two sections are syntactic devices used to comment on the work of others, such as that-clauses (e.g., *We have shown* [*that SON neurons are...*]), which are used to make claims about other statements (cf. West (1980)). One may hypothesize that self-citations may be positively evaluated, whereas outside reference may be negatively evaluated. Other studies regarding the introduction and conclusion sections have looked at authorial comment, which is introduced by modal auxiliaries (*may* and *should* most frequently), adverbs, adjectives of probability (*possible*, *certainly*), attitudinal markers (adverbs such as *surprisingly*), marked choice of nouns (e.g., *view*, *hypothesis*), switch to first person, or unusual use of metaphor or analogy (cf. Smith (1984)). Functionally, the most common type of authorial comment is epistemic, relating to the probability of a proposition or hypothesis being true (called 'arguing' in Wilson et al. (2005a), Wilson (2008)). The other three most common ones are recommending, emphasizing and evaluating (cf. Swales (1990, 137)).

With these devices at hand, the author of a research article tries to "Create A Research Space (CARS)" to locate the research into. In this regard, Swales (1990) has introduced the CARS model, which relates to moves taken by the author in order to create a research space for the research. Nwogu (1997) has elaborated Swales' model to show how research articles are structured. From his *moves* it seems that evaluative devices are to be found most often within the introduction and conclusion sections. In the Introduction (1) related research is reviewed and references to limitations of previous research are made (negative evaluation, indication of gaps, etc.), and (2) new research is introduced and the importance of the author's own new research is emphasized. In the Conclusion, (1) consistent observations are indicated by hedging (such as *appears to be misleading*), (2) non-consistent observations are indicated by negative verb phrases (such as *did not reveal*) or negative quantifiers, (3) overall research outcomes are highlighted by preparatory statements (e.g., *the results suggest that/offer clear evidence that*) or by explicit lexemes such as *major aim has been attained*, and (4) specific research outcomes are explained by lexical items signaling significance/importance (e.g., *results are important*) and by preparatory statements for limitations of outcomes related to indicate limitations of previous studies (e.g., *error*, *clearly unable*, *did not mention*). From these observations, one can think of (a) possible expressions of evaluation involved in scientific writing (such as *importance/significance*) and (b) possible objects the evaluation is directed toward (such as *previous/own research, observations, outcomes*, etc.). Thus, evaluative meaning is inherent to research articles.

2.3 Registers of specialized discourse in a diachronic perspective

Text productions in the form of research articles represent scientific writing, and different specialized discourses are realized in different registers. The term *register* is rooted in a sociolinguistic tradition, where it is used as a cover term for varieties defined by situational context (cf. Ure (1982), Ferguson (1994)), thus defining functional variation rather than regional or social variation (cf. Quirk et al. (1985)). The core idea associated with register is that language use is systemically influenced by contextual factors (Halliday and Hasan, 1985), i.e. the situational context influences the linguistic choices made by a language user and vice versa. There is a bidirectional relationship where the situational context influences the choice of linguistic features used, while the choice of specific linguistic features creates the situational context. According to Halliday (1988), registers are clusters of associated features that have a greater-than-random tendency to co-occur and which can be summarized in terms of field, tenor and mode of discourse realizing the situational context. In the situational context of scientific writing, *field* is associated with knowledge — or the scientific topic — which is extended, transmitted or explored, *tenor* with the roles and attitudes of participants, i.e. the relationship between writers and readers of research articles and the attitudes of the writer toward the scientific knowledge presented, and *mode* is associated with the textual presentation of this knowledge. Particular linguistic features associated with field, tenor and mode give rise to the situational context of scientific writing, i.e. to the registers of scientific writing.

Registers can be identified at any delicacy of focus (Hasan, 1973). Projected on scientific writing, the focus can be turned from scientific writing in general to the scientific writing of particular disciplines. Halliday (1988) presents an example of this type of research on the physical sciences, also adopting the historical perspective. As language use continuously adapts to changing social contexts (cf. Ure (1971, 1982)), registers will change accordingly. Halliday (1988) shows how over a time span of approx. 600 years or so, the physical sciences have changed in terms of linguistic features used. Moreover, changing social contexts are also led by a process of modernization, especially in the scientific domain. Due to modernization, the need arises for new kinds of activity and new kinds of discourse, i.e. as part of the process of modernization within a society, new registers develop (Ure, 1982). In science, modernization is shown, for example, in the need to adopt specific techniques in one discipline which have been developed in some other discipline, e.g. statistical techniques, algorithms, etc. In this way, disciplines begin to share knowledge, which in return might lead to creating new knowledge, around which new scientific registers may develop. Computer science is one of the disciplines whose knowledge is adopted by other disciplines. In terms of register theory, two registers come into contact with each other as the sharing of knowledge is also reflected in the configuration of the variables of field, tenor and mode. While both registers have their own configuration of linguistic features, when they share knowledge these configurations might merge due to register contact. This register contact may lead to the creation of new registers, such as the language of bioinformatics or computational linguistics, if we think of register contact between the disciplines of computer science and biology or linguistics, respectively. These new registers, termed as *contact registers* (cf. Teich et al. (2012, 2013)), would have their own clusters of linguistic features, even though they will still share some similarities with the registers they have evolved from. Note that in terms of terminology, contact *registers* relate to configurations of linguistic features used; with contact *disciplines* we refer to the discipline which encounters register contact. To study the evolution of the contact registers, the delicacy of focus shifts from particular scientific disciplines to highly specialized disciplines. The farther the focus is shifted toward more delicate planes, the more fine-grained the clusters of linguistic features become. Thus, differences in the use of linguistic features realizing registers can be identified on a very general level, e.g. between general language and scientific writing, but also on a highly specialized level, such as for contact registers.

Since the 1960s, numerous scholars have studied registers from different perspectives, ranging from studies on single registers to register variation studies, which can be synchronic, diachronic or cross-linguistic. Studies on single registers focus on describing characteristics of one particular register. In general, these studies are based on an exploratory analysis of example texts of the register under study, i.e. a text representative of the register is chosen on which a register analysis of lexico-grammatical features is carried out (cf. Biber and Finegan (1994, 353), Teich (2003, 24)). Studies in this area range from studies of registers of the conversational type, such as Ferguson (1964) on baby talk and Leech (1966) on British television advertising, to studies of the more professional type, such as Danet (1980)'s work on written legal language or research on written medical discourse (Van Naerssen, 1985), as well as studies on scientific discourse, such as Swales (1990) on English scientific research articles (cf. Biber and Finegan (1994)). They are mostly of a qualitative nature, giving an indication of what the register features of an assumed register are, and can be used to decide which lexico-grammatical features to choose for a quantitative analysis (cf. Teich (2003, 24)). Besides studies on single registers, there have also been register variation studies, i.e. studies that focus on the delimitation of registers. These studies compare two or more registers and are mainly driven by quantitative methods, where a corpus compiled of registers from various sources is used to detect boundaries between registers (cf. Teich (2003, 24)). Biber's work is one of the most prominent in the field of quantitative register analysis. In terms of methods, he defines:

> A comprehensive linguistic analysis of a register requires consideration of a representative selection of linguistic features. Analyses of these register features are necessarily quantitative, because the associated register distinctions are based on differences in the relative distribution of linguistic features. (Biber, 1995, 31)

Now that we have corpora at hand, the quantitative dimension becomes one that we can account for, and as Halliday puts it in an interview:

> The quantitative basis of language is a fundamental feature of language: I think that a grammatical system is not just a choice between a or b or c but a or b or c with certain probabilities attached — and you get these probabilities out of the corpus. [...] It is essential to be aware of the notion of global probabilities in language. [...] What [registers] tend to do is to shift the probabilities, so it is the same system but with a different set of probabilities,

> not only in the vocabulary but also in the grammar. (Thompson and Collins, 2001)

Taking this into account, besides distinctions between relatively general registers such as speech and writing, distinctions can be drawn between more fine-grained registers as may be the case for highly specialized scientific disciplines (e.g., bioinformatics), something that has been investigated by Teich et al. (see, e.g., Teich et al. (2012, 2013)).

Considering that highly specialized domains develop over time, the diachronic change will also be reflected in the language use, i.e. in terms of linguistic features. However, the linguistic evolution of highly specialized registers has obtained little attention so far (cf. Teich et al. (2013); Degaetano-Ortlieb et al. (2013)).

2.4 Approaches to evaluative meaning

There are three main strands of approaches that deal with the identification of evaluative meaning:

(1) approaches based on smaller text samples such as the appraisal system (Martin, 1995, 2000; White, 2000, 2003; Martin and White, 2005; Hood, 2004) or Hunston's work on Status, Value and Relevance (Hunston, 1989),

(2) corpus-based approaches using corpora as evidence or for contrastive studies (Biber and Finegan, 1989; Hunston, 1994; Conrad and Biber, 2000; Hunston and Thompson, 2000; Hyland, 2000, 2009b; Hunston, 2004; Biber, 2006; Hunston, 2011; Thetela, 1997),

(3) approaches based on computational methods aiming at automatic detection of evaluation (Wilson et al., 2005a; Wilson, 2008; Narayanan et al., 2009; Xiaowen and Liu, 2010; Liu, 2010; Somasundaran, 2010).

2.4.1 Text-based approaches

2.4.1.1 Appraisal

One major approach working on smaller text samples is the appraisal system (Martin and White, 2005) based on Systemic Functional Linguistics (SFL; Halliday (2004)). In SFL language is seen from a functional perspective as performing three metafunctions: construing experience (ideational), construing social relations between people (interpersonal), and organizing the discourse (textual). Appraisal is situated within the interpersonal metafunction and aims at studying the set of meanings expressed in a text having at choice a variety of linguistic resources. To display this wide

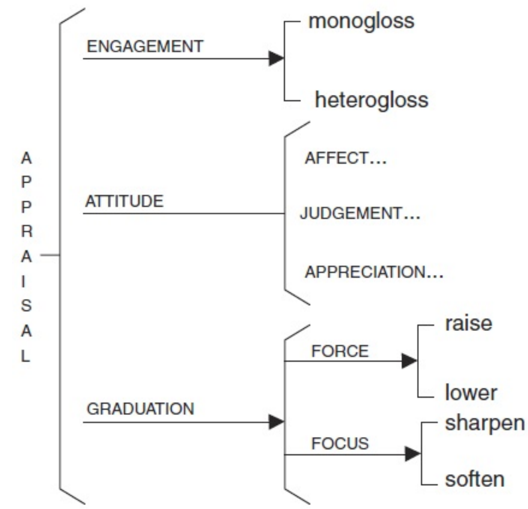


Figure 2.1: System network for the model of Appraisal

variety of choices, Martin and White (2005, 42) have created a system of meanings, adopting the concept of system networks from Halliday. Figure 2.1 shows the model of Appraisal from Martin and White (2005, 38). As can be seen from the categories displayed in the system, appraisal is mainly designed to be a discourse semantic resource for meaning (Martin and White, 2005, 11), rather than being a resource of specific linguistic features (Hunston, 2011, 20).

There are three interacting domains inherent to appraisal: (1) Attitude, concerned with feelings including emotional reactions, judgments of behavior and evaluation of things, (2) Engagement, concerned with sourcing attitudes and the play of voices around opinions in discourse[1], and (3) Graduation, which attends to grading phenomena whereby feelings are amplified and categories blurred (cf. Martin and White (2005, 35)). Attitude is further subdivided into Affect, which deals with positive/negative feelings, Judgement, which deals with attitudes toward behavior, and Appreciation, which involves evaluations of semiotic and natural phenomena (Martin and White, 2005, 42–43). See Examples (1)–(3) taken from Martin and White (2005).

(1) [...] *that was a* **very sad** *day for me. I was* **very unhappy** [...]. (Affect)

(2) [...] *it is often to the lawyer's interest to make* **wrong** *seem* **right** [...]. (Judgement)

(3) [...] *I find nothing* **so pleasant** *as sitting on a* **comfortable** *chair* [...]. (Appreciation)

The strength of the appraisal taxonomy is that it makes it possible to annotate in

[1]Note that this notion of Engagement differs from Hyland's definition (see Section 2.4.2).

detail evaluative meaning within texts, and the relationships established between semantic categories in Appraisal are theoretically motivated. Affect, for example, is subdivided into un/happiness, in/security, dis/satisfaction, each of these further subdivided, e.g., unhappiness into misery and antipathy, happiness into cheer and affection. The more detailed the system, however, the more difficult it becomes to pin down exactly to which category a particular expression belongs. Besides its interpretive nature, the approach is a qualitative one, oriented toward in-depth analyses of a relatively small number of texts rather than to quantitative corpus-based studies (cf. Hood (2005)). Work on appraisal within scientific writing has been done, for example, by Hood (2005). She looks at attitude in undergraduate scientific writing and focuses on the introduction section, this selection motivated by the fact that writers contextualize their own research by positioning it within a topic and in relation to a body of theory and research within the introduction.

2.4.1.2 Status, Value, Relevance

Another approach investigating individual texts is Hunston's early work on evaluation, which focuses on academic prose (Hunston, 1989). Hunston identifies three functions of evaluation: the identification and classification of an object to be evaluated (Status), ascribing a value to that object (Value), and identifying the significance of the information (Relevance/Significance) (cf. Hunston (2011, 21)). Hunston's model predates that of Martin and White (2005) and to some extent has been overtaken by their model (the part on Value). However, while Martin and White (2005) propose a general taxonomy of attitudinal meaning, Hunston proposed a taxonomy of the value meanings construed by research articles (cf. Hunston (2011, 22)), such as her parameters of certainty described in Hunston (1989). These are taken up by Hunston and Thompson (2000) and elaborated into parameters of evaluation: good-bad, certainty, expectedness, and importance. The good-bad and certainty parameters are *real-world-oriented* as they express the writer's view of the status of propositions and entities (in Halliday's terms they are experientially oriented). The expectedness and importance parameters have an additional *text-oriented* function, as they can serve to guide the reader toward the intended coherence of the discourse (cf. Hunston and Thompson (2000, 24)). The good-bad parameter, also known as the positive-negative parameter, is implicitly connected with the other parameters. Evaluations along the certainty, expectedness and importance parameters are related to social/cultural values and can be good or bad, i.e. 'possible' or 'not possible', 'obvious' or 'not obvious', 'important' or 'not important' with possible gradations in between.

2.4.2 Corpus-based approaches

Besides her text-based work, Hunston studies evaluation in a corpus-based fashion as well (see, e.g., Hunston (2004, 2011)). As she points out, there are two traditions in corpus research which are also adopted for the investigation of evaluative meaning: (1) qualitative corpus research, with the aim of identifying and describing the evaluative words and phrases used and their evaluative meanings in context, and (2) quantitative corpus research, which draws on theories of register and variability in language by investigating the distribution of sets of words and phrases across corpora (cf. Hunston (2011, 50)).

2.4.2.1 Evaluation

Qualitative corpus research on evaluation is most prominently contributed by Hunston (see Hunston and Thompson (2000) and Hunston (2011)). Her definition of evaluation is the expression of an attitude toward a person, situation or other entity which is both subjective and located within a societal value-system (cf. Hunston (1994, 210)). In addition to the function of expressing the writer/speaker's opinion reflecting the value system of that person and community, according to Hunston and Thompson (2000), evaluation realizes two other functions: to construct and maintain writer-reader/speaker-hearer relations, and to organize the discourse.

There are many studies that deal with the investigation of these evaluative functions in a qualitative manner. Channell (2000), for example, focuses on individual lexical items or phrases, investigating their collocational information to provide evidence for specific connotations of these expressions, mainly negative or positive connotations. The evaluative function of the selected expression is derived from investigating concordance lines. Thompson and Zhou (2000) investigate evaluative disjuncts, arguing that besides their interpersonal function, they also fulfill a textual function contributing to coherence in text, i.e. evaluative disjuncts organize the discourse. Similar to Channell (2000), they use the corpus to provide evidence for their argument.

There has been also qualitative corpus research of evaluation for scientific writing. Thetela (1997), who based her work on Hunston (1994), has worked on a corpus of human sciences (history, economics, psychology, and applied linguistics) and identified cases of evaluation and categorized them (1) by the evaluated entity such as *studies*, *evidence*, and *results*, and (2) by the ascribed value, i.e. parameters of value which express a particular meaning (parameters of evaluation in terms of Hunston), such as *importance* or *usefulness*. Hyland (1998) uses a similar categorization but also makes a distinction between the sections of a research article (introduction, results, and discussion). He observes that prominent evaluative devices seem to be found in the introduction and discussion section (such as clusters of pronominals, verbs of reasoning, that-nominals, adverbs, adjectives, and modals qualifying assertion) (cf. Hyland (1998)). Charles (2003) contrasts theses from material science and political science, giving quantitative information, for example, about stance-bearing anaphoric nouns, reporting clauses with human subject or with *it*, or on how authors of both disciplines use anaphoric nouns to interpret and express a stance toward information given previously, but varying in the polarization of the evaluation, showing differences of use across the two disciplines. Hunston (2011) takes a closer look at modals (e.g., *you should/must*) and modal-like expressions (e.g., *it is essential to*), status-indicating nouns (*assumption/idea that*) in scientific writing, or patterns of that-clauses and wh-clauses, giving several examples of patterns involved in expressing evaluation.

2.4.2.2 Stance and Engagement

Quantitative corpus research on evaluative meaning is mostly associated with Biber (e.g., Biber and Finegan (1989); Biber et al. (1999); Conrad and Biber (2000)) and Hyland (e.g., Hyland (2000, 2009a); Hyland and Tse (2009)), both having a main interest in register variation for what they term stance and engagement. While Biber conducts contrastive studies on the differences between written and spoken as well as broad register types, such as newspaper language, fiction, academic prose and conversation, Hyland's focus of attention lies on scientific registers. Biber defines stance to encompass Affect, which involves the expression of a wide range of personal attitudes, emotions, feelings, etc. (according to Ochs and Schieffelin (1989)), and Evidentiality (according to Chafe (1986)), which refers to the writer/speaker's attitudes expressed toward knowledge, i.e. toward its reliability, mode of knowing, and adequacy of its linguistic expression (cf. Biber and Finegan (1989)). Later on he divides stance into three major domains: epistemic stance, commenting on the certainty (or doubt), reliability, or limitations of a proposition (e.g., realized by lexical items such as *probably* or *according to*); attitudinal stance, conveying the speaker/writer's attitudes, feelings, or value judgements (realized, e.g., by modal adverbs such as *surprisingly* or *unfortunately*); and style stance, describing how the information is being presented (expressed, e.g., by adverbs such as *briefly* or *honestly*) (cf. Conrad and Biber (2000)).

Hyland's approach is focused in particular on scientific writing. His work has its origins in Metadiscourse, where the interaction between writer and reader is prioritized. He defines Metadiscourse as "the linguistic resources used to organize a discourse or the writer's stance toward either its content or the reader" (Hyland and Tse, 2004). For academic interaction, in addition to stance, he also takes engagement into consideration (see Figure 2.2). His definition of stance relates to Biber's definition by accounting for the attitudinal dimension of conveying judgments, opinions, feelings, commitments, etc. To this notion he adds what he terms Presence, i.e. the extent to which a writer chooses to project him/herself into the text by the use of first person pronouns and possessive adjectives (such as *I*, *we*, *our*, etc.). Besides self-mention, Hyland relies on three other main elements of writer-oriented

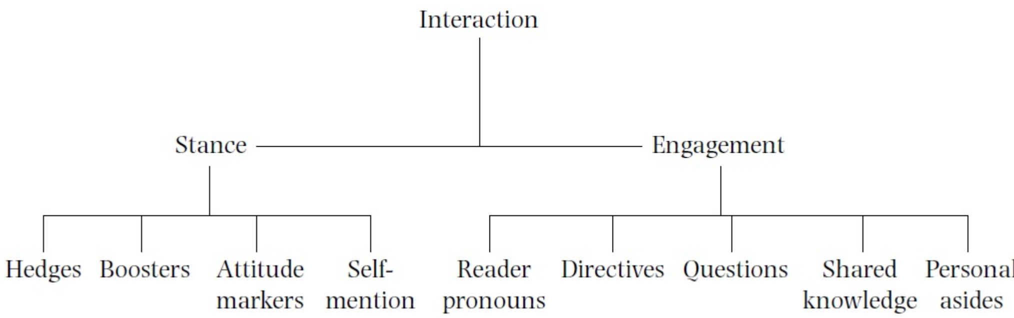


Figure 2.2: Key resources of academic interaction according to Hyland (2005)

features to investigate stance: hedges, boosters and attitude markers (cf. Hyland (2005, 178)). Hedges indicate that the writer presents information as an opinion rather than an accredited fact, withholding complete commitment to a statement. These can be realized lexically, for example, by modal verbs (*may*, *might*, etc.) or verbs (e.g., *suggest*) as shown in Example (4). Boosters, instead, allow writers to express certainty about their statements while effecting interpersonal solidarity (cf. Hyland (2005, 179)) as shown in Example (5). Both hedges and boosters relate to epistemic stance in Biber's terms. Attitude markers relate to the writer's affective attitude to propositions, conveying different attitudinal meanings such as importance, interest, desirability, etc. Lexical markers can be verbs (e.g., *agree*, *prefer*), sentence adverbs (e.g., *unfortunately*, *hopefully*) and adjectives (e.g., *appropriate*, *remarkable*) (cf. Hyland (2005, 180)). Even though Hyland (2005) does not mention nouns explicitly, they can convey this function as well (consider, e.g., *importance*, *interest*). In Biber's terms, attitude markers convey attitudinal stance (see Example (6)).

(4) *Our results suggest that certain microsatellites may act as cis-regulatory elements, controlling gene expression through transcription factor binding and/or secondary DNA structure formation. Due to their high polymorphism and abundance, they might represent an important source of quantitative genetic variation.* (SciTex[2], Biology)

(5) *The gain is obviously due to balancing of work load during run time using the proposed dynamic load-balancing algorithm as is clearly evident from the analysis shown in Figs. 19 and 20.* (SciTex, Digital construction)

(6) *This emphasizes the importance of our attempt to find the simplest sketch construction which has the best guarantees and smallest constants.* (SciTex, Computer science)

[2] A corpus of scientific research articles introduced in Section 4.3 (cf. Teich and Fankhauser (2010); Degaetano-Ortlieb et al. (2013)).

Besides expressing stance, writers "bring readers into the discourse to anticipate their possible objections and engage them in appropriate ways" (Hyland, 2005). Here Hyland refers to what he terms Engagement, which he considers to be part of academic interaction (see Figure 2.2). Two main purposes are related to Engagement: (1) address readers as participants in the discourse, linguistically realized by reader pronouns (e.g., *you*; see Example (7)) or explicit reader mentioning (e.g., *the reader*) and personal asides, i.e. by briefly interrupting the argument to offer a comment on what has been said (see text in parentheses in Example (7)), and (2) rhetorically position the audience by the use of questions, directives and references to shared knowledge (see Examples (9)–(11), respectively). In contrast to Martin and White (2005), who consider Engagement as being concerned with how resources such as projection, modality, etc. position the writer with respect to the value position and potential responses to the value position (cf. Martin and White (2005)), Engagement in Hyland terms deals with the ways writers overtly refer to readers by asking questions, making suggestions and addressing them directly (Hyland, 2001). Thus, according to Hyland, Engagement is more closely related to how the reader is engaged within the discourse rather than how the writer is positioned within the discourse. The linguistic features for Engagement are, therefore, reader-oriented features in Hyland terms.

(7) *This is, in some sense, a geometric data structure (if <u>you</u> think of the pairs as points in N2), so techniques from computational geometry (such as those for orthogonal range searching) may be useful.* (SciTex, Computer science)

(8) *For a lucid discussion of the terms 'language documentation' and 'language description' we refer <u>the reader</u> to Himmelmann 1998.* (SciTex, Linguistics)

(9) *In other words, is there a set of at most k vertices (which may be either black or white) that dominates the set of black vertices?* (SciTex, Computer science)

(10) *<u>Note</u> that GF may consist of several "black components", connected only to white vertices.* (SciTex, Computer science)

(11) *This is <u>commonly known</u> as the OSV structure, 'topicalisation' and 'left-d islocation', being interpreted by some linguists as 'passive'.* (SciTex, Linguistics)

In contrast to Martin and White (2005), both Biber and Hyland are not concerned with providing a comprehensive description of resources available in English to express evaluation. Their interest lies in exploring and explaining the use of evaluative meaning in particular circumstances of use, i.e. how evaluative meaning differs across registers. Thus, while qualitative approaches use the corpus mainly to provide evidence on the usage of expressions, quantitative approaches make use of the corpus to show differences in use across registers. However, one approach does not exclude the other. They can be used complementarily. All qualitative approaches described above can be further investigated across registers, and the insights gained by the quantitative approaches can be explored more closely in a qualitative way.

2.4.3 Computationally based approaches

2.4.3.1 Sentiment Analysis and Opinion Mining

Computational approaches to evaluative meaning are mainly driven by the need for practical applications in the industry. This has led to an increase in research on sentiment analysis, especially since 2001, out of which several research strands have arisen.

2.4.3.1.1 Sentiment classification One main strand is sentiment and subjectivity classification, grounded in natural language processing and treated as a text classification problem, which accounts for two main sub-topics: (1) *document-level classification*, i.e. classifying documents in expressing a positive or negative opinion, and (2) *sentence-level classification*, i.e. classifying sentences in expressing positive, negative or neutral opinion (cf. Liu (2010, 637)).

In document-level classification, the aim is formally the following: given an opinionated document d which comments on an object[3] o, determine the opinion orientation oo of the opinion expressed on o (cf. Liu (2010, 637)). The assumption made here is that the document expresses opinions on a single object and that there is a single opinion holder. This assumption obviously restricts the domain on which document-level classification can be used to customer reviews of products and services as the object (e.g., a camera) and the opinion holder (the review writer) are known. The task can be formulated as a supervised learning problem (as e.g. in Pang et al. (2002)) with two classes (positive and negative). As is common in supervised learning, a training set labeled with the respective classes and an unlabeled test set have to be designed in order to train a model on one dataset and test it on an unseen dataset. Considering that the approaches are meant to be implemented into applications of sentiment detection for the industry, testing the model on an unseen dataset is quite important, as producers want to know how their products are evaluated in new reviews. As reviews are already labeled by their ratings (4–5 stars are considered positive, 1–2 stars negative), they form a readily available dataset. The features used for the classification task are (a) terms (individual words or n-grams) related to the topic and their frequency, (2) part-of-speech tags of adjectives, as important indicators of opinions, (3) opinion words and phrases, and to some extent syntactic dependency and negation. Besides the supervised learning problem,

[3]Note that in this section we adhere to the terminology used in this research field for the thing the evaluation/sentiment/opinion is directed toward, i.e. we use *object*, instead of *target*, which we use in the remainder of this work.

unsupervised learning is also used (Turney, 2002), where several approaches exist. One approach uses an algorithm which extracts phrases containing adjectives and adverbs expressing some form of subjectivity and which have to adhere to specific patterns of part-of-speech (e.g., JJ+NN/NNS+smth as in *beautiful picture*). In a second step, the algorithm estimates the orientation of the extracted instance using pointwise mutual information (PMI) based on the probabilities calculated by issuing queries to a search engine (see Turney (2002)). Finally, the review is classified according to the average opinion orientation of all phrases in the review.

In sentence-level classification, the task involves determining if the sentence is subjective or objective, and if it is subjective, whether it expresses positive or negative opinion. Traditionally, sentence-level classification is performed by supervised learning (see, e.g., Wiebe et al. (1999)). Note that in most applications it is necessary to know beforehand the object (or features of the object) the opinion is directed toward. While in document-level classification labeled datasets are readily available in terms of reviews and their ratings, for sentence-level classification a great manual effort has to be invested in labeling datasets for training the model. To avoid the time-consuming manual labeling, some studies use bootstrapping approaches to label training data automatically (Riloff and Wiebe, 2003; Riloff et al., 2003). Besides subjective words or n-grams, in the iterative process syntactic patterns are learned which are also used to identify more subjective and objective sentences (e.g., subject+passive-verb as in *customer was satisfied*, subject+active-verb as in *customer complained*). However, the assumption is again that sentences express a single opinion from a single opinion holder. This implies that sentence-level classification is not suitable for compound sentences as shown in Example (12) (cf. Liu (2010)). Here, two opinions are expressed: a positive one toward *picture quality* and a negative one toward the *viewfinder*. Nevertheless, besides the mere classification of positive, negative and neutral sentences, some studies work on determining the strength of opinions (neutral, low, medium, high) (Wilson et al., 2004) and contextual polarity (Wilson et al., 2005b) by considering contextual sentiment influencers such as negation (e.g., *not*, *never*) and adversative conjunctions (e.g., *but*, *however*).

(12) *The picture quality of this camera is amazing and so is the battery life, but the viewfinder is too small for such a great camera* (taken from Liu (2010)).

Both document- and level-based classification approaches are based on opinion lexicons, which provide the subjective words used in both approaches. These are built by using bootstrapping techniques to learn subjective words either from a dictionary, such as WordNet (Fellbaum, 1998) from which synonyms and antonyms are essentially used to enlarge the lexicon, or from corpora (Hatzivassiloglou and McKeown, 1997). While the dictionary approach enables the creation of quite large lists, its major drawback is that it cannot account for domain specificity, for which the corpus-based approach is perfectly suited, as corpora can be built to suit specific needs such as domain specificity. Here, seed lists of subjective words are used as well as rules/patterns based on connectives (e.g., *and*, *or*, *but*), which make it possible to find either words with the same orientation (*well-known and important*) or opposite orientation (*counterintuitive but well-known*).

2.4.3.1.2 Feature-based sentiment analysis As seen so far, the above approaches account for a quite general assignment of opinions, i.e. the opinion is attributed to the document or sentence, but not explicitly to the object that is evaluated. This is approached in feature-based sentiment analysis. Here, first of all, the object the sentiment is directed toward is identified (e.g., *picture quality* in Example (13) or *efficiency* in Example (14)). Second, the polarity of the sentiment is determined (i.e. positive, negative or neutral). However, while in Example (13) the polarity of the sentiment toward *picture quality* is clearly positive considering the adjective *amazing*, in Example (14), which is taken from a research article in computational linguistics, the polarity is not as easily assignable to positive or negative. Efficiency is evaluated as being important, which at first glance might indicate a positive polarity. However, as efficiency has to be obtained and might implicate a great effort, it cannot be said to be positive in every sense, while it can be said to be important, whether this is positive or negative is quite hard to determine and is maybe not even intended to be expressed. Due to the specific tasks approached in sentiment analysis, which are mostly concerned with online product reviews and the like, the assignment of more fine-grained classes of sentiment has received less attention so far.

(13) *The **picture quality** of this camera is **amazing**.* (taken from Liu (2010); modified)

(14) *The **efficiency** of the different algorithms is **important** in the parsing problem.* (taken from the SciTex corpus)

What has been accounted for is the extraction of the object evaluated from reviews of different formats (reviews with separate pros and cons or free format reviews) by applying supervised pattern learning approaches (see, e.g., Lafferty et al. (2001); Liu et al. (2005)). In Liu et al. (2005), e.g., rules are generated based on part-of-speech patterns used to extract the object evaluated (the product feature). An example pattern is an adjective followed by a to-particle and a verb (JJ-TO-V), such as *easy to use*, where the verb *use* is determined as the object feature. After identifying the objects, i.e. the things that are evaluated, the opinion orientation is identified, i.e. whether the object is evaluated as positive or negative. The lexicon-based approach, for example, is used to identify positive or negative words from sentences or phrases that contain a previously identified object. Scores of +1 are given to a positive word and -1 to a negative one, while content-dependent words get 0. In addition, negations and but-clauses are accounted for. Finally, the opinion is aggregated, i.e. a score is calculated on the object features in a sentence. If the score is positive, the sentence is labeled as positive; if it is negative or neutral, it is labeled accordingly. Moreover, rules are generated that capture more complex opinions, such as in Examples (15) and (16), in order to classify concepts such as the value range as in (15) or reduction as in (16) into positive or negative.

(15) *This drug causes low blood pressure.* (negative, taken from Liu (2010))

(16) *This drug reduced my pain significantly.* (positive, taken from Liu (2010))

Furthermore, some studies (e.g., Jindal and Liu (2006a,b); Ganapathibhotla and Liu (2008); Narayanan et al. (2009)) also approach sentiment analysis of comparative sentences (see Example (17), where the first algorithm is positively evaluated and the second one negatively, or conditional phrases as in Example (18)). Again, several rules are created to capture the relevant objects and opinions.

(17) *Accordingly, we conclude that the* ***GenSAT algorithm is better than the EA*** *as being an auxiliary of the QCS algorithm.* (taken from the SciTex corpus)

(18) ***If*** *the cell phone was robust, I would consider buying it.* (taken from Narayanan et al. (2009))

As we have seen, the task of sentiment analysis ranges from determining the orientation of documents to sentences and phrases of a particular type. What the approaches to sentiment analysis have in common is the following:

(1) the industrial motivation of finding sentiment expressions within text, mostly reviews[4],

(2) determining the orientation of the opinion, i.e. positive, negative and neutral,

(3) confinement to one particular domain, and

(4) some share the use of dictionaries and rule-based identification of sentiment expressions and objects.

Besides investigating the orientation of specific reviews, i.e. positive, negative or neutral, the application of sentiment analysis is expanding to the analysis of emotion and the like, most prominently within social media, where more fine-grained distinctions are made (emotions/meanings/frames/moods) (see, e.g., Strapparava and Mihalcea (2008); Bollen et al. (2011); Roberts et al. (2012)).

[4]However, some research has extended to other fields (e.g., Somasundaran (2010), who has looked at discourse-level relations within product debates and political debates, or Xiaowen and Liu (2010), who approaches resolving coreference).

However, considering diversification, sentiment analysis aims at finding opinions mostly within one and the same domain (reviews, tweets, etc.), which is clearly related to the purpose of the application. If we think of diversification over time, the time dimension has not been accounted for yet, as has already been pointed out by Devitt and Ahmad (2013, 489), but would be interesting especially for the social and behavioral sciences (cf. Danescu-Niculescu-Mizil et al. (2011)). Moreover, the focus is directed toward classifying the documents, sentences, and phrases correctly, rather than toward the investigation of why they are classified differently, i.e. which features contribute to the distinction. Obviously, if the features are mostly based on lexicons of positive and negative words, the insights gained will not be so revealing. However, with features motivated by deeper linguistic knowledge, such as the part-of-speech rules used for the identification of sentiment expressions and the objects but also stylistic features (as presented in Danescu-Niculescu-Mizil et al. (2011)), insights could be gained on whether text productions differ in their use. However, while the usage of linguistic features in Danescu-Niculescu-Mizil et al. (2011) (e.g., 1st person singular pronoun, certainty meaning, negation) is motivated by psychological studies, features motivated by a linguistic theory such as SFL, which also accounts for the situational context, will lead to interpretations relevant to sociolinguistics studies.

2.5 Summary and conclusions

Research articles have an inherent evaluative character associated with their aim of presenting some kind of new knowledge. In terms of diversification, however, we hypothesize that the evaluative character differs across disciplines, i.e. each discipline has its own evaluative character. To see whether this is the case, we consider the sociolinguistic concept of register, which allows the comparison of different text productions by the investigation of the usage of lexico-grammatical features, focusing on features associated with evaluative meaning. Yet, each existing approach to the study of evaluative meaning presented above has its strengths and weaknesses.

Text-based approaches such as Martin and White (2005) are very detailed in nature and based on a functional language theory, thus allowing very elaborate interpretations on the use of evaluative meaning. However, generalizations on the insights gained are quite hard to make due to a limited amount of evidence (few texts). Its detailedness also bears the problem of being quite time-consuming, which hinders analysis of larger amounts of text. To analyze diversification of different disciplines some kind of generalization has to be made — not possible on this scale of detailedness.

Corpus-based approaches (e.g., Hunston (2011); Hyland (2009b)) move toward more generalizable statements by inspecting a greater amount of text through corpora. However, they lack some grounded theoretical embedding within a theory of language in order to arrive at sound interpretations associated with a sociolinguistic perspective on functional variation. Moreover, the strength of the detailedness of text-based approaches cannot be accounted for in corpus-based approaches, mainly due to the phenomenon of evaluative meaning itself, which is quite hard to pinpoint down to realizations in a text on a large-scale basis. Nevertheless, corpus-based approaches allow the study of diversification in more generalizable ways.

Similarly to corpus-based approaches, computationally based approaches aim to find linguistic realizations of evaluative meaning in large amounts of text. However, they pursue a different aim than text-based and corpus-based approaches, as they are quite application oriented. The focus lies on the correct identification of realizations and their classification into categories such as positive or negative, i.e. the performance of the application is most relevant. There is no motivation toward a sociolinguistic interpretation. While computationally based methods are aware of domain-specific variation, they are not interested in analyzing this kind of variation. This goes hand in hand with the kind of features and approaches used to obtain the best-performing results, which make interpretations in linguistic terms quite difficult.

For the purpose of analyzing diversification and imprint across scientific disciplines, the corpus-based approach is best suited, as it allows (1) generalizable observations and the identification of linguistic realizations of evaluative meaning with some extent of detailedness, and (2) taking a comparative perspective, in our case across disciplines and time.

Still, the theoretical ground has to be established in order to allow interpretations in sociolinguistic terms with an underlying language theory and an appropriate model of evaluative meaning rooted in that theory (see Chapter 3). The model of evaluative meaning has to consider the following:

- the model must be rooted in a theory of language that accounts for the interaction between participants in a discourse and relates the language system to the realizations of the interaction, i.e. how evaluative meaning is construed and realized in language;
- the model must account for resources for the interaction between participants to enact a social and intersubjective relationship, i.e. account for discursive roles and the expressions of evaluations within evaluative acts.

In terms of methods, it is necessary to consider methodologies and techniques which make it possible to account for registerial and diachronic variation (see Chapter 4). Thus, considerations have to be made of the following:

- a corpus is needed that contains different scientific disciplines as well as highly specialized registers and at least two time periods to observe recent diachronic change;

- a methodology is needed to identify evaluative meaning in the corpus;
- an analytical cycle needs to be designed so as to account for macro- and micro-analytical steps to arrive at conclusions on register diversification and registerial imprint;
- in macro-analysis, appropriate techniques have to be considered that make it possible to generalize quantitative observations on diversification and imprint;
- in micro-analysis, it is necessary to consider appropriate techniques that allow a detailed investigation of qualitative observations on diversification and imprint.

Chapter 3

Theoretical approaches and model of analysis

3.1 Introduction

In the present chapter, we present a model for the analysis of evaluative meaning that can be used in research on interpersonal meaning, register variation, cross-linguistic variation, sociolinguistics and the like. In more specific terms, we use the model to approach our hypotheses raised above on the diversification process of scientific disciplines in terms of evaluative meaning over time and a possible imprint left over by the disciplines of origin on the contact disciplines.

In the last few decades, a number of linguistic resources have been shown to express evaluative meaning in scientific writing (e.g., Hyland (2005), Hyland and Tse (2009), Swales (1990), Thompson (2001)). However, we do not yet have a comprehensive model of how evaluative acts are construed. Moreover, only selective aspects are investigated, i.e. the linguistic resources expressing evaluation are still analyzed mostly in isolation rather than being integrated into one model of analysis. Hyland (2005), Hunston and Thompson (2000), and Martin and White (2005) have led the way in this direction, yet the challenge is not to be underestimated and the factors to be accounted for regarding evaluative meaning are manifold (various linguistic forms as well as genre, register and context). Here, we make an attempt to formulate a model for the analysis of evaluation in scientific research articles, adopting some of the ideas and methodologies already established and combining them with new insights, grounding the model within the linguistic theory of Systemic Functional Linguistics (SFL; Halliday (2004)).

As the model is based on a sociosemiotic interpretation of language rooted in SFL, the main underlying concepts of language as a social semiotic are introduced, in particular with respect to the interpersonal component of the semantic system and a particular angle on register. An important assumption in SFL is that (functional) grammar is semantically motivated rather than autonomous (in the sense of formal grammar), i.e. there is a natural and non-arbitrary relation between semantics and lexico-grammar[1], both representing two language strata. It will be seen that semantics works as an interface between the social system and the linguistic system (cf. Halliday (1978)), i.e. social interaction takes a linguistic form through meaning (Section 3.2). Evaluative meaning will be positioned within the linguistic theory of SFL (Section 3.3). In particular, we are interested in the meaning potential of the writer in expressing evaluative meaning. In this sense, the possible selections are displayed that can be made by the writer from the options that constitute the meaning potential (Section 3.4).

3.2 Language as social semiotic

In sociolinguistic terms, social interaction is typically realized in linguistic forms through meanings, i.e. the social context is realized by specific choices of meanings which are further realized in specific lexico-grammatical choices. From a functional perspective, these are considered three strata with three different systems (social, semantic, lexico-grammar) which are interrelated by *stratification* and *instantiation* (see Figure 3.1).

In terms of instantiation, the social system is defined by *situation types*, out of which different contexts of situation can be chosen as single instances. The available options are structured in terms of *field*, *tenor* and *mode*: the social action, the role relationships involved, and the symbolic or rhetorical channel. These three situational variables, which structure the situation type, form "a conceptual framework for representing the social context as the semiotic environment in which people exchange meanings" (Halliday, 1978, 110). The available options of a semantic system are defined by the *meaning potential*, i.e. the range of choices present in the system for a particular situation type. Constellations of specific meanings chosen from the meaning potential are referred to as *text*, i.e. text being the linguistic form of social interaction (cf. Halliday (1978, 122)), which is embedded in a context of situation. The lexico-grammatical system offers choices that realize the text and in a more general sense the context of situation. However, the relation is not unidirectional but bidirectional, i.e. a particular social context determines the lexico-grammatical choices made, whereas particular lexico-grammatical choices create a social context.

The semantic system acts as an interface between the social system and the lexico-grammatical system. Similarly to the situation type, the semantic system is structured in three metafunctions, *ideational*, *interpersonal* and *textual*: the meaning potential of a writer related to content, the meaning potential of a writer related to participation, and the text-forming potential. These are related respectively to

[1] Note that lexico-grammar in SFL stands for grammar; the lexis being not a separate component but the most 'delicate' end of the lexico-grammar (see, e.g., Halliday (1961); Hasan (1985)).

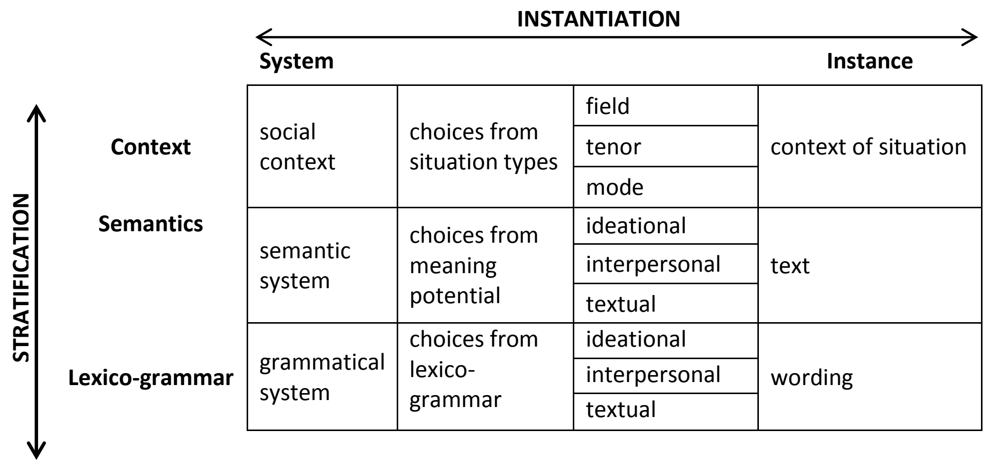


Figure 3.1: Dimensions of stratification and instantiation

the three situational variables of field, tenor, and mode in the sense that the situational variables activate specific options from the corresponding metafunctions which are then realized in lexico-grammar. Thus, particular constellations of lexico-grammatical realizations are reflected in particular constellations in terms of field, tenor and mode. These constellations are referred to as *register* defining functional variation.

The investigation of diversification of scientific disciplines can thus be approached by considering functional variation in terms of register, i.e. by looking at specific constellations of lexico-grammatical realizations in one discipline vs. the other. Therefore, lexico-grammatical differences will point to registerial differences.

As texts constitute the actual linguistic interaction, i.e. the selected choices made from a total set of options, texts are the source for investigating differences in linguistic interaction. In Halliday's terms, text is a semantic concept/unit which is realized in sentences (not composed of sentences), which are characterized by certain lexico-grammatical features. However, text is not a lexico-grammatical unit; it is a semantic unit, i.e. the meanings are encoded in the semantic system and given the form of text (cf. Halliday (1978, 140)). Text is distinguished primarily by organization into functional components, i.e. configurations of meanings of different kinds: ideational, interpersonal, textual. Thus, text is understood as a continuous process of semantic choice and the meanings created by the social system are exchanged by the members of the social context in the form of text (cf. Halliday (1978, 137, 141)). The continuity between text and its sociosemiotic environment is established by the concept of register. Each text belongs to a class of texts defined by the register, which is defined by clusters of field, tenor and mode of the situation and also reflected in the ideational, interpersonal and textual semantic components.

As our focus lies on diversification of scientific disciplines in terms of evaluative meaning, the need arises to determine where evaluative meaning is located within the semantic system to be able to investigate registerial differences in this regard.

3.3 Evaluative acts and the interpersonal metafunction

Obviously, the act of expressing evaluations toward some kind of content is performed by the writer with a persuasive intention toward the reader. This is clearly related to the participatory function of language, i.e. the interpersonal metafunction. However, the interpersonal metafunction includes more than a persuasive intention. It ranges from modality and speech function in the clause to settings affecting the whole of a particular register (like the distance we associate between hard and soft sciences) and it is quite fuzzy to grasp and diffuse as it is realized by different grammatical and lexical features or other devices such as voice quality and intonation contours, etc. (cf. Halliday and Matthiessen (2006, 527)). Nevertheless, evaluative meaning can be positioned within the interpersonal metafunction.

The interpersonal metafunction provides the resources for participants to enact a social and intersubjective relationship, through the assignment of discursive roles and the expression of evaluations and attitudes (cf. Halliday and Matthiessen (2006, 12)). Within the interpersonal metafunction, evaluative meaning is associated with the expression of evaluations and attitudes and the participants involved. One important aspect of interpersonal acts which is also fundamental to evaluative acts is the potential for arguing reflected in the mood system, i.e. an exchange of speech roles among the interactants constructing a range of speech-functional variation. The fundamental types of speech roles are *giving* or *demanding* meaning, i.e. 'inviting to receive' and 'inviting to give'. It is not only the speaker/writer who has to do something, but also the hearer/reader who must do something. It is an interaction between the 'me' and the 'you' which are constructed in language. Considering research articles, a dialog is formed between the writer(s) and the reader(s) and we will stick to these terms further on.

The clause is constructed as a move in an argument either as a *proposition* (question or statement) or as a *proposal* (offer or command), i.e. what is exchanged is either semiotic — *information* which is construed by language itself — or material — *goods-&-services* which exist independently of language (cf. Halliday and Matthiessen (2006, 113–114)). Still, the exchange is concerned with meaning, whether semiotic or material, and is always an act of either giving or demanding meaning. But there are two clear distinctions between the exchange of goods-&-services and information: (1) while for exchanges of goods-&-services the meaning serves to bring the exchange about, for information the meaning itself is exchanged, (2) while the exchange of goods-&-services cannot be argued about, the exchange of information

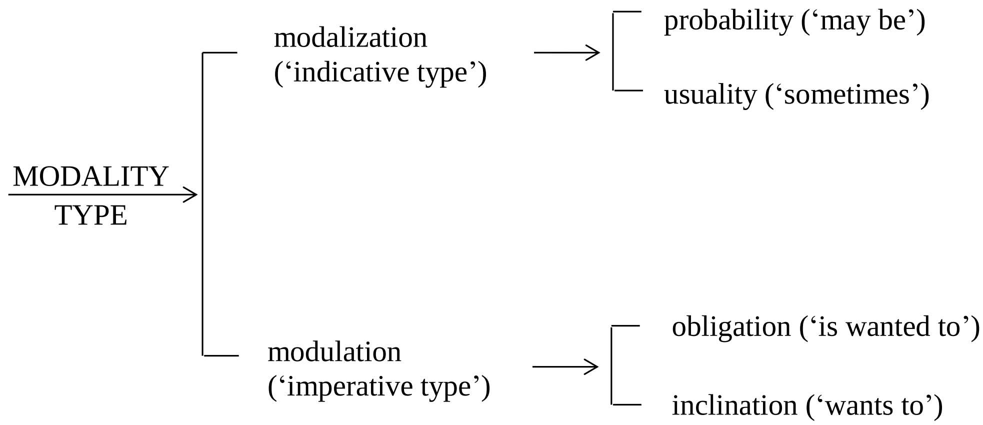


Figure 3.2: System of types of modality (cf. Halliday (2004, 618))

can be argued about, i.e. it can be affirmed, denied, doubted, etc. Thus, evaluative meaning is mostly concerned with the exchange of information rather than with that of goods-&-services, because the latter, being either command or offer, is usually not arguable.

The argumentative scope is provided by the system as an opposition of positive and negative polarity. Interpersonal meaning, however, is not delimited to positive or negative; it offers a whole range of semantic meanings lying in between these two poles. This is referred to as modality, where interactants present their own judgements, opinions, etc. (Halliday and Matthiessen, 2006, 526). The speech function of the clause, i.e. an exchange of information (proposition) or goods-&-services (proposal), determines the range of meanings possibly expressed. Figure 3.2 shows the semantic system of modality with two modality types. While modalization relates to propositions, modulation is related to proposals. Thus, modalization can be considered to realize evaluative meaning, i.e. the writer's stance which can be argued about. If modalization is expressed, i.e. the clause is a proposition realized as indicative in the semantic sense, some degree of either probability or usuality is expressed. There are three values of modality between the positive and negative poles: high, median and low. Zooming into what Halliday (2004) terms probability (see Figure 3.3), toward the positive pole, there are expressions of certainty (e.g., *it certainly is*), while at the median and low levels, there are expressions of likelihood (e.g., *it probably is*, *it may be*). In semantic terms, both certainty and likelihood represent *epistemic meaning* contributing to the stance of the writer and thus realizing evaluative meaning. Besides modality, however, there are numerous other kinds of evaluative meaning, which express desirability, importance, benefit, etc.,

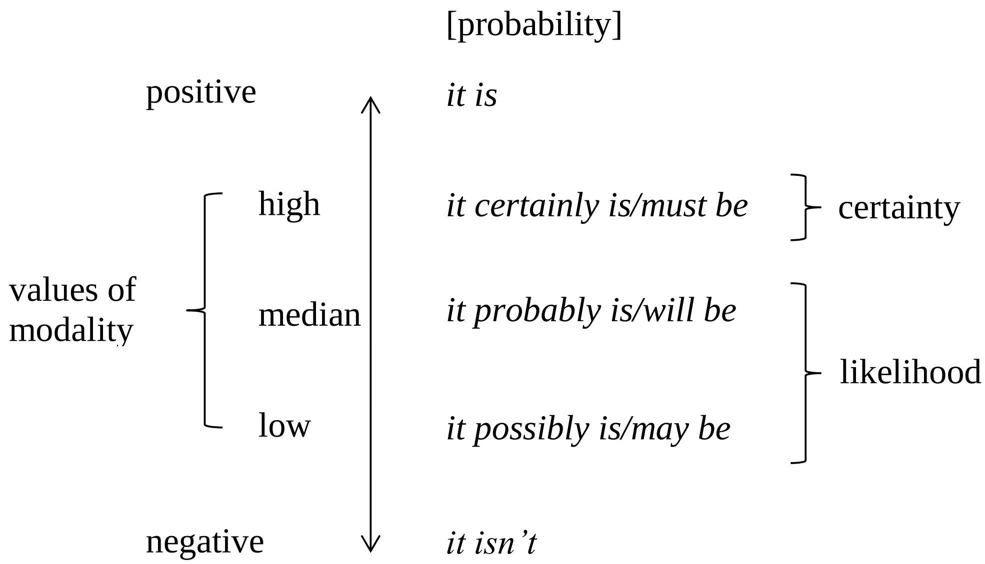


Figure 3.3: Semantic relations of probability to polarity and values of modality

realizing *attitudinal meaning* (rather than epistemic meaning) which is also related to the writer's stance. As we have seen in Section 2.4.1.1, appraisal is one approach that moves toward a comprehensive account of how this faceted picture of different kinds of meanings is constructed.

Considering diversification of scientific disciplines, however, we are not only interested in the different kinds of evaluative meaning, but also in differences related to the entities involved. With entities we relate, first of all, to the participants in the discourse: in scientific writing, the writer(s) of a research article and the reader(s). The role relationship is clearly defined within interpersonal terms. Considering that the writer expresses his own attitudes and judgments and seeks to influence the attitudes and behavior of others (cf. Halliday (1978, 112)), the question arises of what these attitudes and judgments are directed toward, what should the reader perceive as being evaluated? Thus, an additional entity is the 'what the evaluation is attributed toward', i.e. what is the target of some kind of evaluative attribution by the writer and is intended to be perceived by others as receiving the attribution. While appraisal focuses on and gives a very detailed account of the different possible attitudinal meanings (judgement, affect, appreciation) and engagement values (e.g., entertain, attribute, disclaim) possible in an interpersonal act, the entities involved are less considered, especially the thing the evaluative attribution is directed toward. However, the meaning exchanged is constituted not only by the attitudinal meanings and the engagement values, but also by the 'what' these are directed toward in order for the reader to perceive the evaluative act properly and in its full range. Thus, the full meaning potential of evaluative meaning should account for the participants involved, the possible choices in terms of these participants as well as the target of the evaluative act, i.e. what the evaluation is attributed toward.

To investigate registerial differences accounting for the meaning potential in terms of evaluative meaning, considering also the entities involved within evaluative acts, we need a model of analysis that accounts for these considerations.

3.4 Major constituents of evaluative acts in scientific writing and model of analysis

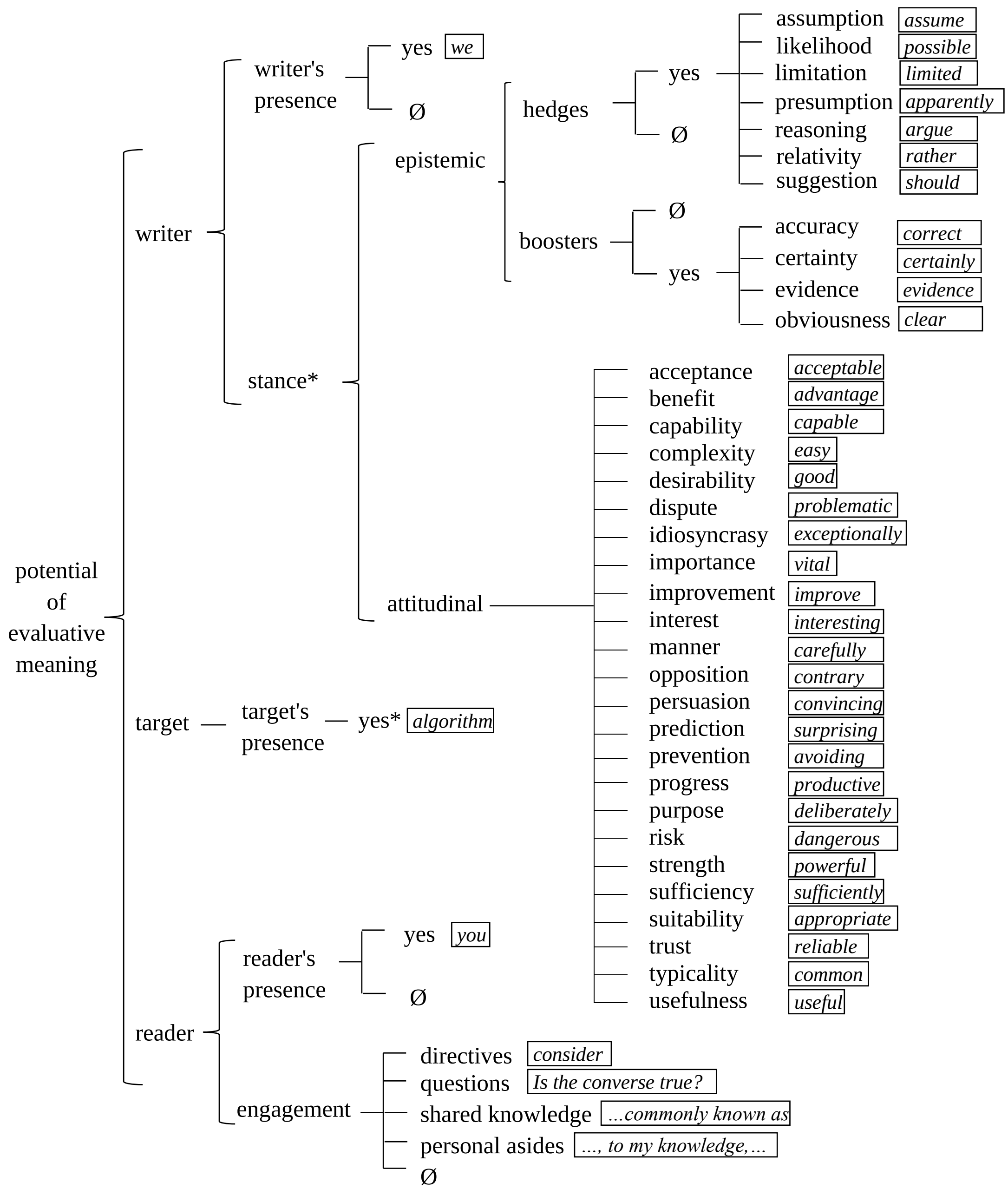
On a semantic level, we have seen in Section 3.3 that an evaluative act is constituted by the participants involved in the discourse, the meaning potential regarding evaluative meaning and the target the evaluative meaning is directed toward. Figure 3.4 shows the potential of semantic choices available in an evaluative act with the main semantic elements being *writer*, *reader* and *target*.[2] When an evaluative act takes place, the writer may be explicitly mentioned in the discourse (see Example (1)), but does not have to be. What is always present in an evaluative act is the stance the writer adopts. The stance expressions can be of the *epistemic* or *attitudinal* type (cf. Conrad and Biber (2000) and Hyland (2005)). As we have seen in the previous section, epistemic expressions are related to the system of modality in terms of SFL, realizing e.g. the meaning of certainty or likelihood. As certainty is positioned toward positive polarity, it functions semantically as a *booster* (see Example (2)). This function can also be realized by the meanings of accuracy evidence and obviousness in scientific writing (see Example (3)). The meaning of likelihood, instead, is located toward the negative pole, functioning semantically as a *hedge* (see Example (4)). As can be seen from Example (5) both hedges and boosters can be combined, positioning the expression toward median polarity rather than high polarity (consider, again, Figure 3.3). Other meanings functioning as hedges are assumption (e.g., *assume that*), limitation (e.g., *is limited to*), presumption (e.g., *apparently*), reasoning (e.g., *argue that*), relativity (e.g., *rather*) or suggestion (e.g., *suggest that*).

(1) [$_{\textbf{writers'-presence}}$ *We*] *prove that our design prevents this from happening* [...].

(2) *This* [$_{\textbf{booster-of-certainty}}$ *certainly*] *constitutes a limiting factor, since the aim of the controller resides precisely in vibration suppression.*

(3) *Each such move is* [$_{\textbf{booster-of-obviousness}}$ *clearly*] [$_{\textbf{benefit}}$ *beneficial*].

[2] For writer and reader, we have adopted some of Hyland's model (cf. Figure 2.2 above).



* If an evaluation takes place, both target and stance are chosen from the system.

Figure 3.4: Potential of semantic choices for evaluative acts with examples

(4) [...] *additional forces arising due to electrostatic charges* [$_{\textbf{hedge-of-likelihood}}$ *might*] *be* [$_{\textbf{importance}}$ *important*].

(5) *This diversity* [$_{\textbf{hedge-of-relativity}}$ *almost*] [$_{\textbf{booster-of-certainty}}$ *certainly*] *neutralizes protein sub-class specific characteristics such as those seen in the enzyme and antibody sets.*

Besides epistemic stance, the writer has the potential to express attitudinal stance. While epistemic stance is grammaticalized as it is related to the system of modality, attitudinal stance is expressed lexically. The meanings expressed with attitudinal stance are far less easily classifiable into more abstract semantic categories than boosters or hedges for epistemic stance. The range of meanings used to express attitudinal stance is quite wide and diversified and is not limited to the meanings shown in Figure 3.4. Therefore, attitudinal stance constitutes in some way an open class.

Both epistemic and attitudinal stance can be realized in a variety of ways in the lexico-grammar, e.g., by different parts-of-speech (adjectives such as *beneficial* or *important* in Examples (3) and (4) or adverbs such as *clearly* in Example (5)). Furthermore, both boosters and hedges can be combined with attitudinal stance. In Example (3), the attitudinal expression (*beneficial*) is enforced by the preceding booster of obviousness (*clearly*). In Example (4), instead, the attitudinal expression of importance is toned down by the preceding hedge of likelihood (*might*). Semantically, the combination of boosters and hedges with attitudinal expressions enables moving the attitudinal expression to either positive and negative polarity.

Essential to an evaluative act is also the target the evaluation is directed toward, be it evaluated with either epistemic, or attitudinal stance or a combination of both. Thus, if an evaluation takes place, the target as well as the stance expression are present in the discourse. Semantically, target and stance expressions are closely related, as the evaluation expressed by epistemic or attitudinal stance is always directed toward a target. In terms of SFL, this semantic relation is also visible in the lexico-grammar, where the realizations of targets and stance expressions appear together in lexico-grammatical patterns. Considering that we are looking at evaluative meaning, we choose to call them *attributive evaluative patterns* which are used to attribute an evaluation by the writer toward the target of the evaluation (see also Hunston and Francis (2000) on evaluative patterns and Hunston and Sinclair (2000) on local grammar of evaluation). In terms of lexico-grammar, our model encompasses five sets of attributive evaluative patterns (see also Figure 3.5):

- the *eval_target* set which comprises patterns where the evaluative expression precedes the target (see Examples (6)–(7)),
- the *eval_rel-v_target* set with patterns that use relational verbs and where the evaluation precedes the target (see Examples (8) and (9)),
- the *target_eval* set comprising patterns where the evaluative expression follows the target (see Examples (10)–(11)),
- the *target_rel-v_eval* set containing patterns with a relational verb where the target precedes the evaluative expression (see Examples (12) and (13)), and

- the *target_v_eval* set which contains patterns that use no relational verb but other types of verbs and where the target again precedes the evaluation (see Example (14)).

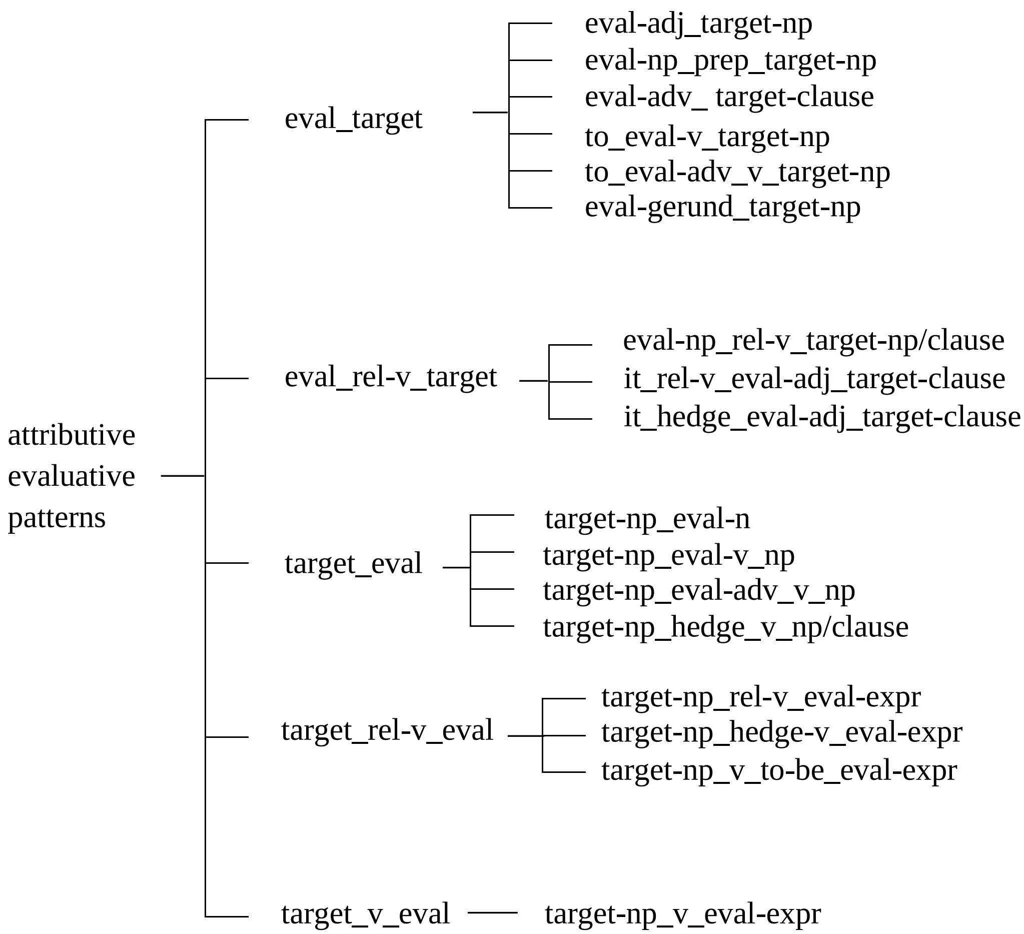


Figure 3.5: Attributive evaluative patterns identified

(6) [...] *three* [$_{\textbf{eval-adj}}$ *important*] [$_{\textbf{target-n}}$ *parameters*] [...].

(7) [$_{\textbf{eval-adv}}$ *Importantly*], [$_{\textbf{target-clause}}$ *it also permits a neat interface*] [...].

(8) [$_{\textbf{eval-np}}$ *One key output variable*] [$_{\textbf{rel-v}}$ *is*] [$_{\textbf{target-np}}$ *area A1 in Fig. 17*].

(9) [...] [$_{\textbf{it}}$ *it*] [$_{\textbf{rel-v}}$ *is*] [$_{\textbf{eval-adj}}$ *essential*] [$_{\textbf{target-clause}}$ *that the train and test set are identical*].

(10) [$_{\textbf{target-n}}$ *A*] [$_{\textbf{eval-v}}$ *fails*] *to be a BPP algorithm.*

(11) [$_{\textbf{target-n}}$ *Word*] [$_{\textbf{eval-n}}$ *importance*] [...].

(12) [...] [$_{\textbf{target-np}}$ *the approach*] [$_{\textbf{rel-v}}$ *is*] [$_{\textbf{eval-adj}}$ *appropriate*].

(13) [...] [$_{\textbf{target-np}}$ *the approach*] [$_{\textbf{hedge}}$ *seems*] [$_{\textbf{rel-v}}$ *to be*] [$_{\textbf{eval-adj}}$ *reliable*] [...].

(14) [$_{\textbf{target-n}}$ *Retrieval*] [$_{\textbf{v}}$ *has played*] [$_{\textbf{eval-np}}$ *a major role*] [...].

Each set can be realized by several lexico-grammatical patterns. The *eval_target* set consists of six patterns, where the evaluation precedes the target evaluated. Table 3.1 shows each pattern as well as examples. The 'eval-adj_target-np' pattern consists of an evaluative adjective (e.g., *important*) within a nominal phrase. In this case, the noun (e.g., *parameters*) is evaluated with the adjective. In the 'eval-np_prep_target-np' pattern, an evaluative noun phrase (e.g., *the importance*) is followed by the prepositions *of*, *in* or *to* and a further noun phrase (e.g., *the algorithm*) on which the evaluation is attributed. The 'eval-adv_target-clause' pattern consists of an evaluative adverb (e.g., *importantly*) which precede a clause that is evaluated by the adverb. The two other patterns 'to_eval-v_target-np' and 'to_eval-adv_v_target-np' have either an evaluative verb in the infinitive (e.g., *to help*) or an evaluative adverb plus an infinitive verb (e.g., *to better understand*) which are followed by a nominal phrase that is evaluated. Finally, the 'eval-gerund_target-np' pattern consists of an evaluative gerund (e.g., *improving*) which evaluates the following nominal phrase.

eval_target patterns	**examples**
eval-adj_target-np	*three important parameters*
eval-np_prep_target-np	*the importance of the algorithm*
eval-adv_target-clause	*Importantly, it also permits a*
to_eval-v_target-np	*to help users identify the best method*
to_eval-adv_v_target-np	*to better understand these complex phenomena*
eval-gerund_target-np	*improving the basic correction algorithm*

Table 3.1: Patterns of the *eval_target* set

The *eval-rel-v-target* set includes three attributive patterns which constitute a relational construction where the evaluation precedes the target evaluated (see Table 3.2). The 'eval-np_rel-v_target-np/-clause' pattern consists of an evaluative noun phrase (e.g., *The importance of this work*) which is followed by the relational verb *be* as well as a noun phrase or a clause which are evaluated. Thus, the evaluation is attributed to what follows the relational verb. The other two patterns consist of an *it* that is followed by a relational verb or a hedge and an evaluative adjective. What follows is a *that* or *to*-infinitive clause which represents the target evaluated.

eval_rel-v_target patterns	**examples**
eval-np_rel-v_target-np/-clause	*The importance of this work is to give* [...].
it_rel-v_eval-adj_target-clause	[...] *it is desirable that the analyses* [...].
it_hedge_eval-adj_target-clause	[...] *it seems difficult to improve them* [...].

Table 3.2: Patterns of the *eval_rel-v_target* set

For the next three sets, the evaluation follows the target. Table 3.3 shows the patterns belonging to the *target_eval* set with examples. The 'target-np_eval-n' pattern consists of a noun phrase (e.g., *memory and attention*) which receives the evaluation from the following evaluative noun (e.g., *limitations*). The 'target-np_eval-v_np' pattern consists of a noun phrase (e.g., *many template-based approaches*) followed by an evaluative verb (e.g.,*lack*) and a further noun phrase (e.g., *a general mechanism for...*). The evaluation is attributed to the first noun phrase. Instead of an evaluative verb, there also might be an evaluative adverb (e.g., *easily*) followed by a verb which is the case for the 'target-np_eval-adv_v_np' pattern. Additionally, the evaluative verb might be a hedge verb (e.g., *suggest*) as for the 'target-np_hedge-v_np/clause' pattern which might then be followed either by a noun phrase or by a clause.

target_eval patterns	**examples**
target-np_eval-n	[...] *memory and attention limitations* [...].
target-np_eval-v_np	*Many template-based approaches lack a general mechanism for* [...].
target-np_eval-adv_v_np	[...] *the designer can easily build* [...].
target-np_hedge-v_np/clause	*The presented experimental evidence suggests that* [...].

Table 3.3: Patterns of the *target_eval* set

The patterns of the *target_rel-v_eval* set include patterns with a relational verb which is followed by an evaluative expression. The target of the evaluation precedes the relational verb. Table 3.4 shows the patterns of this set with examples. The 'target-np_rel-v_eval-adj/np' pattern consists of a noun phrase followed by a relational verb and an evaluative expression (e.g., *relevant*). The evaluation is attributed to the noun phrase preceding the relational verb. Instead of an evaluative verb, there might be a hedge verb (e.g., *seems*) which precedes the evaluative expression or a verb followed by a *to be* (e.g., *proved to be*), as in the 'target-np_hedge-v_eval-expr' and 'target-np_v_to-be_eval-expr' patterns, respectively.

Finally, the *target_v_eval* set consists of one pattern, the 'target-np_v_eval-expr' pattern. An example is shown in (15). Here, a target noun phrase is followed by a verb which is followed by an evaluative expression.

target_rel-v_eval patterns	**examples**
target-np_rel-v_eval-expr	*Ontologies are relevant* [...].
target-np_hedge-v_eval-expr	[...] *this model seems to be appropriate* [...].
target-np_v_to-be_eval-expr	*SVM proved to be the best classifier* [...].

Table 3.4: Patterns of the *target_rel-v_eval* set

(15) *In both situations,* [$_{\textbf{target-np}}$ *multivariate statistics*] [$_{\textbf{verb}}$ *play*] [$_{\textbf{eval-np}}$ *a major role*].

The targets evaluated can be (1) single targets (e.g., *parameters* in Example (6)), (2) double targets obtaining the same evaluation polarity (both positive or both negative), (3) double targets receiving an opposed polarity (one is positive and the other negative), and (4) targets used in a relational construction.

For double targets evaluated with the same polarity, consider Examples (16) and (17). In Example (16) the evaluative attitudinal expression is *improve.* This type of expression evaluates not only the target *accuracy*, which is improved (the Goal), but also the target *background correction*, which performs the improvement (the Actor). The evaluation polarity is clearly positive and is directed toward both participants, in our case targets. The polarity in Example (17) is obviously negative, with the evaluative expression *worsen* relating again to both targets *ABB* and *the SEU rate.*

(16) *In particular,* [$_{\textbf{target-1}}$ *background correction*] [$_{\textbf{hedge}}$ *appears*] *to* [$_{\textbf{eval-pos}}$ *improve*] [$_{\textbf{target-2}}$ *accuracy*] [...].

(17) [...] [$_{\textbf{target-1}}$ ABB] [$_{\textbf{hedge}}$ *can*] [$_{\textbf{eval-neg}}$ *worsen*] [$_{\textbf{target-2}}$ *the SEU rate*] [...].

For double targets evaluated with opposite polarities, consider Example (18), in which the evaluative expression is *outperforms.* In this case, *algorithm* is positively evaluated to outperform *a strict block model*, which is negatively evaluated.

(18) [...] [$_{\textbf{target-pos}}$ *our algorithm*] [$_{\textbf{booster}}$ *consistently*] [$_{\textbf{eval}}$ *outperforms*] [$_{\textbf{target-neg}}$ *a strict block model*] [...].

Additionally, targets can be evaluated with an epistemic or an attitudinal value. Examples (16)–(18) show targets evaluated with attitudinal values. Example (19) shows a target evaluated with epistemic value (evidence). Here, the authors of the research article referred to with *we* state an epistemic fact with *show that* evaluating what follows the *that.* In this case, *we* is the Actor and the *that-clause* represents the Goal.

(19) *In all,* [$_{\textbf{target-1}}$ *we*] [$_{\textbf{eval-epist}}$ *show that*] [$_{\textbf{target-2}}$ *the PQ trees help reduce the number of clusters to be analyzed*] [...].

Engagement features are clearly involved in the interaction of evaluative acts; however, in the lexico-grammar they are not as closely connected to the other two constituents, writer and target. They may occur, for example, in attributive evaluative patterns together with stance features (mostly the case for the reader's presence, see Example (20)), but do not have to, as in the question in Example (21).

(20) [$_{\textbf{writer's-presence}}$ *We*] [$_{\textbf{hedge}}$ *assume*] [$_{\textbf{target-clause}}$ *that* [$_{\textbf{reader's-presence}}$ *the reader*] *is familiar with the basic notions of formal language*].

(21) [$_{\textbf{question}}$ *If it does, how long does it take?*]

Thus, we can say that the target and its evaluation by stance expressions are essential features in an evaluative act, which are always involved in some kind of attribution structure. Engagement features, instead, run parallel to those and are not necessarily involved in an evaluative act. Thus, we will focus on stance and target features for our investigation.

3.5 Summary

In this chapter, we have set the theoretical base to analyze evaluative meaning and created a model of analysis grounded on a sociosemiotic interpretation of language. The semantic choices available for evaluative meaning are realized in the lexico-grammar and can be traced by inspecting lexico-grammatical choices.

Considering that we are interested in the diversification of and imprint left on scientific disciplines, the semantic system of choices presented in Figure 3.4 will allow us to trace the choices made in different disciplines by looking at the lexico-grammatical expressions and evaluative attributive patterns. We might, for example, encounter the case that a register has a preference for intruding the writer into the discourse (e.g., by a frequent use of the writer's presence), for using a particular meaning (e.g., a preference for importance) or attributive evaluative pattern (e.g., relational verb patterns) or for evaluating particular targets (e.g., more nominal vs. clausal targets). Besides looking at the distinct choices of evaluative meaning made across disciplines, we will trace the diachronic development of the contact disciplines to inspect diversification trends and linguistic imprint. The methodology we adopt to pursue these kinds of investigations is presented in the next chapter.

Chapter 4

Methodology

4.1 Introduction

The main goal of this chapter is to design a methodology for the investigation of evaluative meaning across registers and time. As we are interested in diversification trends and linguistic imprint, we focus on contact registers, as introduced in Section 2.3, which have emerged out of contact between two other registers. Additionally, contact registers are prone to show recent diachronic changes in terms of language use, as they have newly emerged.

The starting point is that scientific registers have specific linguistic characteristics that distinguish them from one another. Thus, in our case scientific registers will vary in their use of evaluative meaning, reflected in the distinct use of specific lexico-grammatical features. The main hypothesis is that these characteristics may change over time, showing diversification trends, i.e. registers will become increasingly distinct from one another. A second main hypothesis is that as a contact register has emerged out of two other registers, it will still carry linguistic reflections of evaluative meaning from its seed registers.

In the first section of this chapter, we will make these hypotheses more concrete considering register contact. To test the generated hypotheses, we adopt comparative corpus-linguistic methods. For this purpose, a corpus is needed that covers scientific registers emerged out of register contact (e.g., computational linguistics) as well as the corresponding scientific registers which have come into contact (e.g., computer science and linguistics). In addition, the corpus has to be diachronic for the investigation of recent diachronic change. Here, we consider two time periods, first, the 1970s/80s, approximately the time period where contact registers such as computational linguistics emerged, and second, the early 2000s, as a more recent time period. Furthermore, a methodology is needed to test the hypotheses based on qualitative as well as quantitative analyses of lexico-grammatical features associated with evaluative meaning. Thus, we need detailed qualitative analyses located on a micro-analytical level to investigate which lexico-grammatical features are employed to express evaluative meaning, but also quantitative analyses located on a macro-analytical level to account for more generalizable trends (cf. Jockers (2013)). The methodology has to account for recursive steps between these two analytical levels.

Section 4.2 presents the hypotheses formulated to investigate register diversification of scientific disciplines over time. Section 4.3 presents the corpus used for the analyses and Section 4.4 illustrates the analytical cycle with the micro- and macro-analytical steps and the techniques employed.

4.2 Hypotheses

In general terms, we are interested in how highly specialized scientific registers vary in their use of evaluative meaning and whether these registers become more distinct from relatively closely related registers over time, or if they show some kind of overlap. As explained in Section 2.3, the focus of identifying registers is shifted toward more delicate planes. Contact registers (e.g., computational linguistics) form good candidates for this kind of investigation when compared to their seed registers (e.g., computer science and linguistics). Considering this delicacy of focus, we make our leading hypotheses more concrete as follows:

- H1: *Register diversification.* Scientific registers vary in their use of evaluative features, i.e. they form their own clusters of lexico-grammatical features of evaluative meaning. However, we hypothesize that contact registers are prone to be less well distinguished than their seed registers, especially in the time period they have approximately emerged. Even though both contact as well as seed registers will become increasingly distinct over time, the diversification will be more pronounced for the contact registers, as they have a stronger incentive to become registers of their own than the seed registers, which are already relatively well established as registers.

With respect to the contact registers, however, there is also a secondary hypothesis which has to be considered:

- H2: *Registerial imprint.* As contact registers have emerged out of two other registers, besides their growing distinctness, they will still reflect some linguistic characteristics of their seed registers, i.e. they will show a possible imprint left by the seed registers. Thus, we hypothesize that contact registers will show some overlaps with the seed registers in terms of evaluative meaning. Diachronically, the contact registers might (i) exhibit a shift over time from being more similar to one seed register to being more similar to the other, i.e. by sharing features with one or the other, (ii) show clear tendencies over time to one seed register, or (iii) lie in between the two seed registers.

At the core of H2 is the assumption that a contact register lies somewhere in between its two seed registers, while H1 assumes that the contact registers develop their own clusters of features. The distinction between these two hypotheses is obviously not clear cut, but they have to be tested separately. Additionally, these hypotheses bear some implications for how they can be tested.

For H1, we have to consider how well seed and contact registers are distinguished by evaluative features and whether they are better distinguished over time. For this, several seed and contact registers are considered. It is assumed that all contact registers will be less well distinguished than the seed registers in the 70s/80s (t1), i.e. in the time period the contact registers have approximately emerged. This will change over time when considering the early 2000s (t2), as the contact registers will create their own clusters of features which will contribute to a better distinction.

For H2, a single contact register has to be compared to its two seed registers, i.e. we exhibit a triple comparison of two seed registers and one contact register. While the contact register will have evaluative features of its own, it will also share some features with its seed registers (registerial imprint). The triple comparison will show with which of the two seed registers the contact register shares the most features. The diachronic development of a contact register in terms of registerial imprint is tested by comparing results from observations on t1 and t2. This will indicate possible shifts or consolidation trends.

4.3 Corpus

To investigate the above hypotheses, we use SciTex, the English Scientific Text Corpus (Teich and Fankhauser, 2010; Degaetano-Ortlieb et al., 2013), which was specifically built to investigate register formation in scientific writing focusing on interdisciplinary contact between computer science and other selected scientific disciplines (see Figure 4.1)[1]. The corpus covers nine scientific disciplines: five *seed disciplines* (A subcorpus: computer science; C subcorpus: linguistics, biology, mechanical engineering and electrical engineering) and four *contact disciplines* (B subcorpus: computational linguistics, bioinformatics, digital construction and microelectronics). The corpus contains approx. 34 million words and comprises two time slices, the 70s/80s (SaSciTex) and the early 2000s (DaSciTex), covering a thirty year time span similarly to the Brown corpus family (Kučera and Francis, 1967; Hundt et al., 1999). For the two time slices, SciTex also has two one-million-word subcorpora which are cleaned of erroneous data produced by the OCR conversion from PDF files to text files. In these two subcorpora formulas (mostly from computer science) and examples (as in linguistics) were tagged to exclude them from linguistic

[1]The corpus was built within the project *Registers in Contact* (*Regico*). Work in this project was funded by the *Deutsche Forschungsgemeinschaft* (DFG) under grants TE-198/2 and EXC 284: *Multimodal Computing and Interaction.*

searches. SciTex encompasses full English journal articles from at least two different journals per discipline and wherever applicable, the same journals were used for the two time periods. The corpus has been annotated on the level of tokens, lemmas and parts-of-speech (PoS) using TreeTagger (Schmid, 1994). Additionally, each document has been enriched with meta-information (such as author(s), title, scientific journal, academic discipline, and year of publication) as well as document structure (e.g., abstract, introduction, section titles, paragraphs and sentence boundaries). SciTex is encoded in the Corpus Query Processor (CQP) format (Evert, 2005) and can be queried with CQP by using regular expressions in combination with positional (e.g., PoS) and structural attributes (e.g., sentence, sections).

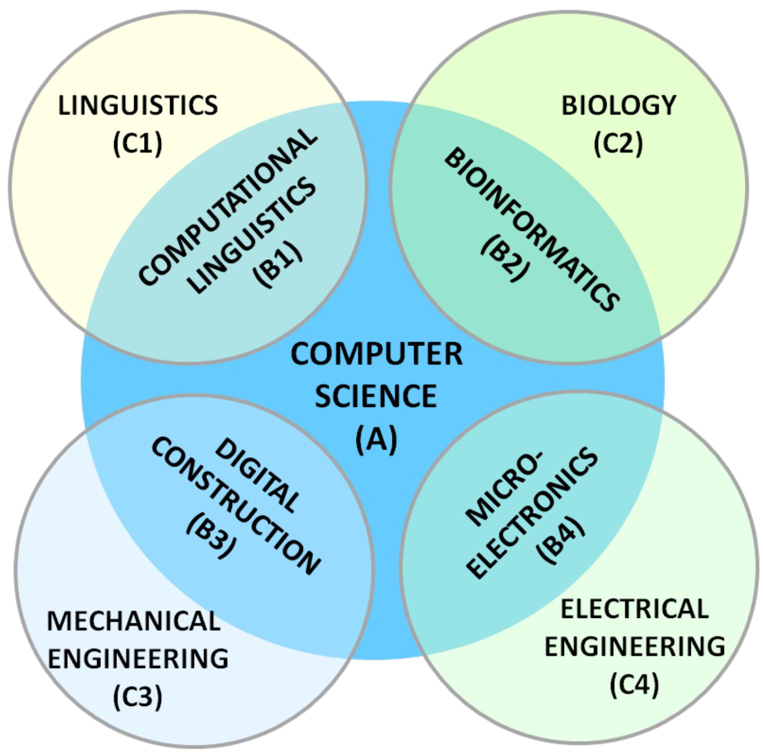


Figure 4.1: Scientific disciplines in the SciTex corpus

In addition, we created a random sample of SciTex, which amounts to approx. 52,000 words, and which was manually inspected and annotated for evaluative lexico-grammatical features to get a first impression of which features might be involved in expressing evaluative meaning. This subcorpus is built out of the abstract, introduction and conclusion sections only. That selection was motivated by Nwogu (1997)'s observations that these sections are apt to include a large amount of evaluative meaning in comparison to the main part of research articles, which was supported during our own corpus inspection: we had started by inspecting whole articles (including the main part), but noted that the main part contained few expressions of evaluative meaning in comparison to the other sections. As manual annotation is a very time-consuming task, this choice of sections allowed us to more efficiently inspect how evaluative meaning is possibly expressed. The annotation was performed by one annotator and helped to formalize the lexico-grammatical features involved. It was performed on texts of 43 different authors, and the subcorpus contains 1212 evaluative and 869 non-evaluative sentences.

4.4 Analytical cycle: Macro- and micro-analysis

Testing the hypotheses given in Section 4.2 requires considering quantitative as well as qualitative data. While a quantitative approach requires a macro-analytical view, qualitative findings need to be inspected on the level of micro-analysis.

On the macro level of analysis, comparisons can be carried out by looking at trends and patterns of features aggregated over the entire corpus (cf. Jockers (2013, 24)). Note that, here, *pattern* is not meant as, for example, part-of-speech patterns (as our attributive evaluative patterns above), but relates to a paradigmatic or typical usage of specific features across data, in our case a typical/characteristic usage of features realizing evaluative meaning. Thus, we look at patterns of features which reflect the usage of characteristic dispositions of expressing evaluative meaning. By the macro-analytic approach, we consider quantitative data to better understand the context in which individual text, features, occurrences, etc. exist.

The micro level of analysis is associated with what Jockers (2013) terms ‘close reading’ related to qualitative data analyses, i.e. looking closely at individual occurrences of some features or words in a context, be it in a key word in context (KWIC) view or in a word cloud or larger strands of texts (e.g., sentences, paragraphs), but it is also associated with closely inspecting aggregated features in terms of a feature analysis, in which features clustered together or top-ranking features from statistical techniques would be inspected.

Both levels of analysis (macro and micro) complement each other. The macro-scale perspective informs the micro-scale one by providing a fuller sense of the whole context in which specific texts exist (cf. Jockers (2013)). The micro-scale level helps to inspect general trends more closely and enforce the evidence provided by the macro-analysis. By considering each level in separation, one would miss the bigger and more comprehensive picture. Only by combining both levels of analysis can one reach a new and better understanding of the data (cf. Jockers (2013)). Nevertheless, these approaches to analysis differ in their method of accessing texts and gathering facts, as in macro-analysis data is inspected quantitatively, while in micro-analysis it is inspected in a qualitative way. Qualitative analysis allows fine-grained distinctions to be made. The aim is to generate complete detailed descriptions rather than quantification (cf. McEnery and Wilson (2001, 76)). For example, ambiguity is one case that can be fully recognized in a qualitative analysis, as for example the disambiguation of the word *bank*. However, by ‘close reading’, i.e. by the qualitative approach, it is hard to make generalizations over a larger population, such as the distribution of the use of modal verbs across texts. Generalizations can only be obtained by a quantitative analysis, revealing details about the texts that are unavailable to close readers of the texts (cf. Jockers (2013)). Again, this difference does not imply that the two approaches are divergent from each other; quite the contrary, because of these differences they complement each other. “Qualitative analysis can provide greater richness and precision, whereas quantitative analysis can provide statistically reliable and generalisable results” (McEnery and Wilson, 2001). Considering macro- and micro-analysis as being two extremes of a scale, depending on the findings and their force of evidence, one has to move toward one or the other side of the scale to find a balance for the application of the appropriate degree of macro- and micro-analysis.

In addition to finding the right degree of complementation between macro- and micro-analysis, different dimensions of comparison have to be considered that make it possible to properly investigate the phenomenon under study. As we aim to investigate whether there are differences in the use of features of evaluative meaning across scientific registers, the first dimension to consider is the *registerial* one. The registerial dimension can be further subdivided into different levels of comparison for the corpus at hand: (1) all registers against each other (for H1), and (2) groups of registers (for H2). The first level of comparison is relatively straightforward, as one register is compared to the others. Comparisons on this level will possibly lead us to observe tendencies inherent to individual registers that stand out from the others, but might also indicate general tendencies across registers. On the second level, groups of registers might show particular tendencies. Considering the registers covered by SciTex, these groups can be formed by different constellations. What we are particularly interested in is how contact registers relate to their seed registers. Thus, comparisons are carried out for triples of registers as pointed out above (e.g., computational linguistics vs. linguistics vs. computer science). To capture possible diachronic changes related to the use of evaluative features across registers, we also consider the *time* dimension. This enables us to detect whether possible preferences of registers — differences as well as commonalities across registers — have changed over time. The time dimension runs parallel to the registerial dimension. Thus, comparisons made in the registerial dimension can be pursued in a diachronic perspective.

For our purposes, we have designed an analytical cycle which accounts for different recursive macro- and micro-analytical steps (see Figure 4.2). We start with the macro-level, where we perform a classification of documents into the nine registers of SciTex according to feature vectors on both time periods to test H1 (register diversification) and on triples of subcorpora (A-Bs-Cs) to test H2 (registerial imprint). In a second step, we perform a micro-analysis by inspecting the feature weights of the classification results to determine characteristic features of evaluative meaning for each register for H1 and to determine features adopted from the seed registers by the contact registers for H2. A further micro-analytical step is then performed to check for data quality by inspecting the results of the feature analysis in more detail on concordance lines, i.e. by looking at the extraction results of the most distinctive features. This will lead, if necessary, to a revision of the extractions, which will allow us to inform the macro-analysis and the micro-analytical feature analysis with data of higher quality.

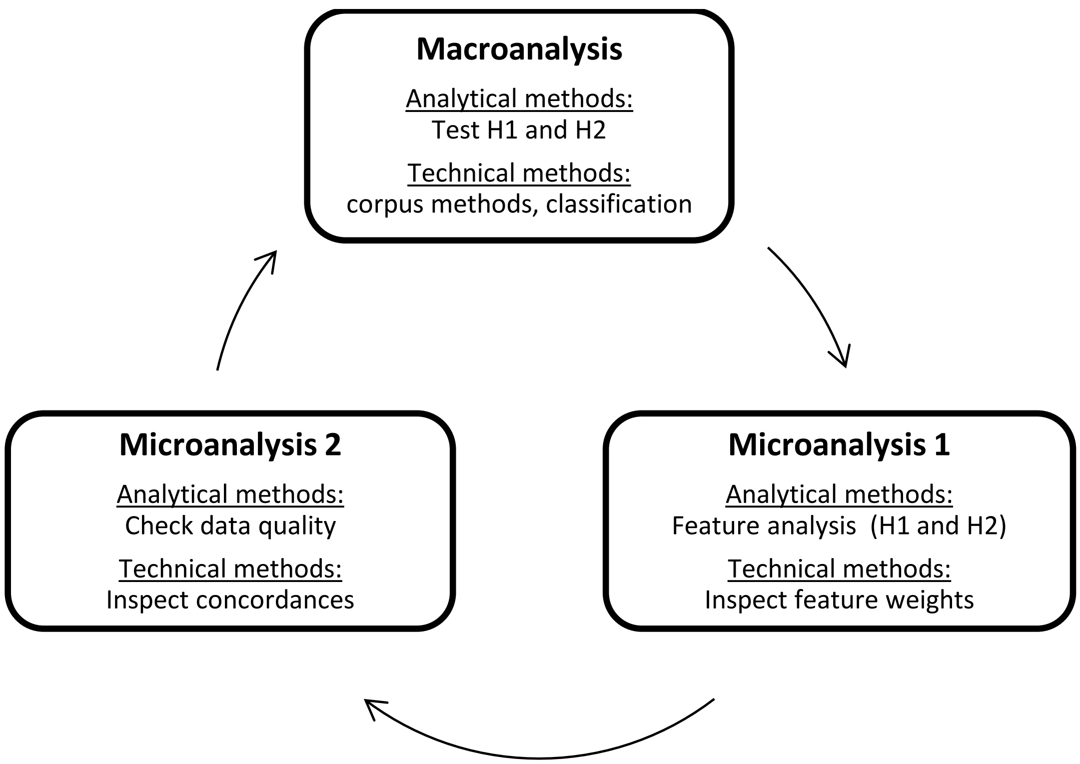


Figure 4.2: Analytical cycle

In the following sections, we describe in more detail the individual analytical and technical steps.

4.4.1 Macro-analysis: Analytical methods

4.4.1.1 Hypotheses testing

In our study, the macro-analysis serves to test our hypotheses of registerial diversification (H1) and registerial imprint (H2) to obtain generalizable observations.

For *register diversification*, we are interested in the distance D between one register (e.g., A: Computer science) and the sum of the other registers in SciTex. We hypothesize that the distance D_{t1} in early text productions (70s/80s) is smaller than the distance D_{t2} in later productions (2000s). For computer science (A), for example, this would be:

$D_{t1}(A, X) < D_{t2}(A, X); X = \{B_1, B_2, B_3, B_4, C_1, C_2, C_3, C_4\}$

Furthermore, we also hypothesize that the difference between D_{t1} and D_{t2} of contact registers will be greater than the difference in distance of the seed registers:

$D_{t2}(B_i, X)$ - $D_{t1}(B_i, X) > D_{t2}(A, X)$ - $D_{t1}(A, X)$
$D_{t2}(B_i, X)$ - $D_{t1}(B_i, X) > D_{t2}(C_i, X)$ - $D_{t1}(C_i, X)$

For *registerial imprint*, we are interested in the distance D between a contact register B_i and its seed registers A and C_i. We hypothesize that B_i will either show a greater distance from A than from C_i or a greater distance from C_i than from A:

$D_{t1}(A, B_i) > D_{t1}(C_i, B_i)$ or
$D_{t1}(A, B_i) < D_{t1}(C_i, B_i)$

This distance might then either be preserved over time ($\approx$), or change, i.e. the distance between the contact register B_i and A might be smaller than the distance between B_i and C_i in t1 but greater in t2, or the distance between B_i and A might be greater than the distance between B_i and C_i in t1 but smaller in t2:

$D_{t1}(A, B_i) \approx D_{t2}(A, B_i)$ or $D_{t1}(C_i, B_i) \approx D_{t2}(C_i, B_i)$

$D_{t1}(A, B_i) < D_{t1}(C_i, B_i)$ while $D_{t2}(A, B_i) > D_{t2}(C_i, B_i)$ or
$D_{t1}(A, B_i) > D_{t1}(C_i, B_i)$ while $D_{t2}(A, B_i) < D_{t2}(C_i, B_i)$

Thus, B_i might be more similar to A in t1, while it is more similar to C_i in t2, or B_i might be more similar to C_i in t1 and more similar to A in t2.

4.4.1.2 Features

As we know from register theory, registers are formed of clusters of lexico-grammatical features which have a greater than random tendency to co-occur. Thus, to measure the distance between registers and test the above formalized hypotheses, we need to consider features which are associated with evaluative meaning and which might be differently used across registers. In our case, we use linguistic features (features of evaluative meaning) generated out of a detailed linguistic model (as described in Section 3.4). The features investigated are shown in Table 4.1. Three different sets of features are used: (1) *stance* features associated with the writer's expression of stance as well as the writer's presence (termed self-mention in the following), (2) *evaluative pattern* features associated with the expression of stance within patterns, and (3) *target* features associated with what is evaluated.

Stance features are indicated by an $S\_$ and have five types: (1) document structure, i.e. the amount of stance expressions across document sections (Abstract, Introduction, Main, Conclusion), (2) stance meaning, i.e. the meanings listed in Section 3.4

feature set	**feature type**	**feature examples**
stance	document structure (S_doc)	*S_Abstract*, *S_Conclusion*
	meaning (S_m)	*S_importance*, *S_complexity*
	pos (S_pos)	*S_NN*, *S_JJ*
	sem (S_sem)	*S_epistemic*, *S_attitudinal*
	type (S_type)	*S_self-mention*, *S_booster*
evaluative pattern	document structure (P_doc)	*P_Main*, *P_Introduction*
	meaning (P_m)	*P_desirability*, *P_reasoning*
	pattern type (P_type)	*P_eval-adj_target-np*
target	token of single target (T1_lemma)	*T1_algorithm*, *T1_model*
	second token of bi-gram target (T2_lemma)	*T2_system*, *T2_approach*
	length (T_length)	*T4_length*, *T7_length*

Table 4.1: Features investigated

in Figure 3.4, (3) the part-of-speech of the stance expression, (4) the semantic type, which can be either epistemic or attitudinal, and (5) the type of the stance expression, which can be self-mention (referring to the writer's presence), booster, hedge or attitude marker.

Evaluative pattern features are indicated by a *P_* and have three types: (1) document structure, i.e. the amount of patterns across document sections, (2) the evaluative meaning expressed within the pattern, and (3) the pattern type (e.g., *eval-adj_target-np*; see Figure 3.5 in Section 3.4 for the other types).

Target features are indicated by a *T_* and have three types: (1) single token targets (e.g., *efficient* [$_{\text{target}}$ *algorithm*]), (2) targets occurring in a bi-gram on the second position, i.e. syntactic heads in a bi-gram, and (3) the token length of the target (uni-gram to 7-gram, e.g. *the lack of* [$_{\text{target}}$ *a fully articulated theory of proper names*]).

In addition, stance and evaluative pattern features are also considered at sentence beginning, which approximates Theme position, i.e. cases in which stance or pattern features appear at the beginning of a sentence. These are indicated by *Sb_* and *Pb_*, respectively.

In total, we consider a list of 271 features. The different steps in our analytical cycle will then help to cope with the number of features in a manner relevant for testing our hypotheses. These steps are described in the following macro- and micro-analytical steps.

4.4.2 Macro-analysis: Technical methods

Inherent to macro-analysis are techniques that make it possible to obtain empirical evidence which leads to generalizable findings. For this purpose, appropriate techniques have to be adopted. In our case, first, corpus-based techniques are employed, which include an annotation procedure to consistently annotate realizations of evaluative meaning in large corpora to be able to extract features from the corpus in a way relevant for testing our hypotheses. Second, statistical techniques are needed that match the requirements to test our hypotheses. Here, we use a classification technique.

4.4.2.1 Corpus-based techniques

In the following sections, we will describe the annotation procedure to consistently annotate realizations of evaluative meaning in large corpora to be able to extract the above-described feature sets from the corpus.

4.4.2.1.1 Technical framework: CWB, CQP and annotation procedures To annotate the full version of SciTex, we use annotation procedures derived from the YAC recursive chunker (Kermes, 2003). For this, the Corpus Workbench (CWB) is used.

In a preliminary step, we have to consider (1) the features to annotate (stance features, evaluative attributive patterns and target features), (2) the attributes we want to annotate for each feature (meaning, type, etc.), and (3) which lexicons have to be built. The lexicons have to capture realizations of different meanings by different parts-of-speech, i.e. the epistemic and attitudinal meanings and their subcategories such as likelihood and obviousness for epistemic expressions and importance and desirability for attitudinal expressions. The procedure is three-fold and encompasses (1) the use of lexical items listed in the Frame Index in FrameNet (Ruppenhofer et al., 2010) for the specific meanings related to the category *quality*, (2) the extraction of lexical items from our manually annotated corpus, and (3) the use of WordNet to find synonyms for the lexical items taken from FrameNet and our own corpus. As the annotation in our procedure is based on the use of part-of-speech tags, we built lexicons for each part-of-speech, e.g., for the importance meaning four lists have been generated, one for each part-of-speech expressed by this meaning (adjective, adverb, noun, and verb). The lists have to be defined in advance as variables in a define_list macro (see Figure 4.3, e.g., `$att-adj-importance`). The variables can then be used in the annotation macros (see Figure 4.4) to look for lemmas of the respective meaning and part-of-speech.

Having the lexicons and the categories to be annotated, the annotation procedure designed for the Corpus Workbench involves the use of (1) queries as rules based

```
MACRO define_list_att-importance(0)

    define $att-adj-importance < 'adj-importance.txt';
    define $att-adv-importance < 'adv-importance.txt';
    define $att-n-importance   < 'n-importance.txt';
    define $att-v-importance   < 'v-importance.txt';
;
```

Figure 4.3: Example macro for the attitudinal expressions of importance

```
MACRO attitude-marker-importance(0)
(
    ([pos="J.*" & lemma=$att-adj-importance])
    |
    ([pos="RB.*" & lemma=$att-adv-importance])
    |
    ([pos="N.*" & lemma=$att-n-importance])
    |
    ([pos="V.*" & lemma=$att-v-importance])
)
;
```

Figure 4.4: Example of a query macro for the attitudinal expressions of importance

on PoS tags and structural attributes that search for a defined feature in the corpus, and (2) Perl scripts that permit delimiting the range of the features found, if necessary, and define the attributes to be annotated. For the first annotation step of defining rules for each feature, consider the query in Figure 4.5 (a query macro) which is used to annotate one prepositional attributive pattern (*eval-np_prep_target-np*). Here an evaluative nominal phrase containing an evaluative noun is followed by the preposition *of* and a further noun phrase, which can be followed by a prepositional phrase, a conjunction or a dash (trying to cover the most common noun phrase dispositions). These rules were defined manually and results were evaluated for precision in the small version of SciTex 2000s (one million words). In the case of low precision, the rules were refined to obtain the best possible results. Especially for the multiple-word features (such as the attributive patterns), this procedure was very important for obtaining good results.

For the second step, a Perl script is used to annotate features into the corpus. After CQP is started by the Perl script, the annotation macro (shown in Figure 4.6) uses the query macro of the prepositional pattern (shown in Figure 4.5) for annotation (see `/eval-of-target[]` in line 1 in Figure 4.6) and starts a subroutine. Here, first, the range of the pattern to be annotated is defined (see line 2 in Figure 4.6). In this case, the last token searched for in the query macro is omitted in the annotation. To make this more comprehensible, consider line 4 in Figure 4.5, where the last token to be matched should not be a preposition (`IN`), a noun (`N.*`) or a conjunction (`C.*`). This token is used in the query macro to allow searching for both single nouns and

```
MACRO eval-of-target(0)
(
    #eval-np followed by "of" and NP
    (/np-eval-n[] "of" /np[] [pos!="IN|N.*|C.*"])
    |
    #eval-np followed by "of" and NP with PPs
    (/np-eval-n[] "of" /np[] /pp[]+)
    |
    #eval-np followed by "of" and NP with conjunction
    (/np-eval-n[] "of" /np[] ([word=","] /np[]){0,2}[pos="C.*"] /np[]
    [pos!="N.*"])
    |
    #eval-np followed by "of" and "-" with NP
    (/np-eval-n[] "of" /np[] "-" /np[] ([pos!="N.*"]|/pp[]))
)
;
```

Figure 4.5: Example of an annotation macro for attributive features

```
1  Macro("/eval-of-target[]", sub {
2      $end = $end-1;
3      Struc(
4      "evaluation" => undef,
5      "evaluation_set" => "eval_target",
6      "evaluation_pattern" => "eval-np_prep_target-np",
7      "evaluation_meaning" => "importance",
8      "evaluation_sem" => "attitudinal",
9      "evaluation_precision" => "98.09%");
10 });
```

Figure 4.6: Example of an annotation rule for attributive features

compounds. If this token were left out of the query, only the first noun of a compound would be matched. In the annotation, however, this token should be excluded, as it is not part of the attributive pattern. This is determined by delimiting the length of the matchend with `$end-1` in line 2 of Figure 4.6. Second, additional structural information in the form of attributes can be added to the annotated feature. As an example, for the prepositional pattern five attributes are annotated: the *evaluation set* (`eval_target`), the *evaluation pattern* (`eval-np_prep_target-np`), the *evaluation meaning* (`importance` in this case, but it could be filled with other meanings), the semantic type of the expressions (abbreviated with *sem* which can be either *attitudinal* or *epistemic*), and the *precision* of the query rule (`98.09%`).

The annotation is performed for each feature (stance, attributive evaluative patterns and target features) on different annotation layers, each with attributes of their own.[2] This allows one to perform queries on multiple annotation layers, so that one can search, for example, for attributive patterns with specific meanings (e.g., only the importance meaning) or used in specific document sections (Abstract, Introduction,

[2]Note that we have not annotated engagement features, as we concentrate on the main devices of evaluative acts in this study. As we have seen in Section 3.4, while stance, targets and the attributive patterns are primarily involved in an evaluative act, engagement is secondary.

etc.).

The annotation is stored as XML tags with attributes and respective values as shown in the example in Figure 4.7. In this case, the importance expression is embedded in the start and end xml-tags for the evaluation pattern (`<evaluation>` and `</evaluation>`) with respective attributes and values. In CQP the annotation is encoded as shown in Table 4.2. Thus, each attribute is encoded in a separate tag.

```
<s> This underlies
    <evaluation set="eval_target" pattern="eval-np_prep_target-np"
    meaning="importance" sem="attitudinal" precision="98.09%">
    the importance of sparse solutions in overcomplete problems
    </evaluation>
    , as is the case for the RTI data.
<s>
```

Figure 4.7: Example of the xml annotation in CQP

4.4.2.1.2 Enhancing data quality In the annotation procedure we also account for possible noise on the extraction results. Two examples are introduced to exemplify errors due to incorrect tagging and the exclusion of features due to very low precision.

Example of incorrect tagging: Self-mention The stance feature *self-mention*, which relates to the writer's presence in the research articles, is annotated by looking at the lemmas of personal pronouns (part-of-speech tag `PP` or `PP$`) as shown in Figure 4.8. Here the singular (*me*, *myself*, *my*, *mine*) and the plural forms (*we*, *us*, *ourselves*, *our*, *ours*) of the first person personal pronouns are used. For the singular form *I*, a separate macro (see Figure 4.9) had to be created to disambiguate the personal pronoun from other usages of the word *I* (e.g., when used as a mathematical variable but wrongly tagged as a personal pronoun; see Figure 4.10 for examples). Here *I* tagged as a personal pronoun (`PP`) should not be preceded by a verb in base form (see `pos!="VV"` in Figure 4.9) and the lemma *table*, but could be preceded by the base form of *do* and *say* and should be followed by either *am*/*was*, *do*/*did*, *have*/*had* and any verb in base form in present and past excluding *are* and *were*.

In this case, two query macros (`/self-mention[]` and `/self-mention-I[]`) had to be created, as the annotation range differs for each case, which has to be considered in the annotation. For the query macro in Figure 4.8, the whole expression is annotated. For the query macro in Figure 4.9, instead, the range of the annotation has to be delimited. The start position is set to +1 to omit the first token of the query (see `[pos!="VV" & lemma!="table"] | pos="VV" & lemma="do|say"]`). The end position is set to -1 to exclude the possible verb forms following the personal pronoun (see `[pos="VBD|VBP|VD|VDD|VH|VHD|VV|VVD" & word!="are|were"]`).

feature	**encoding example**
stance	<stance> <stance_type att-marker> <stance_meaning importance> <stance_sem attitudinal> *important* </stance_sem> </stance_meaning> </stance_type> </stance>
attributive patterns	<evaluation> <evaluation_set eval_target> <evaluation_pattern eval-adj_target> <evaluation_meaning importance> <evaluation_precision 98.61%> *important question* </evaluation_precision> </evaluation_meaning> </evaluation_pattern> </evaluation_set> </evaluation>
targets	<etarget> <etarget_etype np> <etarget_epattern eval-adj_target> <etarget_emeaning importance> <etarget_eprecision 98.61%> *question* </etarget_eprecision> </etarget_emeaning> </etarget_epattern> </etarget_etype> </etarget>

Table 4.2: Encoding examples of attributes for each feature

Example of exclusion: target_eval-n pattern The *target_eval-n* pattern, consists of a noun phrase in which the noun at the end of the expression is evaluative (see *limitations* in Example (1)). For this pattern, disambiguation is very important to obtain relevant results, as many nouns that might be evaluative in one context are not evaluative in another. Consider Examples (2) and (3) taken from computer science in SciTex 2000s. In both cases, *problem* is used as a term or a task in computer science, with no evaluative function. This is relatively explicit here as in Example (2) it is used in a headline and in Example (3) an abbreviation for the term is mentioned (see *CSP*); both variants indicate a non-evaluative usage of *problem*. However, in other disciplines, such as computational linguistics and linguistics, *prob-*

```
MACRO self-mention(0)
(
    ([lemma="me|myself|my|mine|we|us|ourselves|our|ours" & pos="PP.*"])

)
;
```

Figure 4.8: Query macro for self-mention with singular and plural personal pronoun forms

```
MACRO self-mention-I(0)
(
  (
  ([pos!="VV" & lemma!="table"] | [pos="VV" & lemma="do|say"])
  [lemma="I" & pos="PP"]
  [pos="VBD|VBP|VD|VDD|VH|VHD|VV|VVD" & word!="are|were"]
  )

)
;
```

Figure 4.9: Query macro for self-mention for the first person singular pronoun *I*

lem can be evaluative (consider Example (4)). What enforces the fact that *problem* in Example (4) is evaluative is that it is preceded by an additional evaluative adjective (*aggravating*), as well as the verb *leading to*, which in many cases introduces a negative evaluation.

(1) *It could be that unbounded Merge, and whatever else is involved in UG, is present at once, but only manifested in limited ways for extraneous reasons (*<[$_{\textbf{target-np}}$ *memory and attention*] [$_{\textbf{eval-n}}$ *limitations*]> *and the like)* [...].

(2) *Constraint satisfaction* ***problems***

(3) [...] *and graph coloring as a constraint satisfaction* ***problem*** *(CSP)* [...].

(4) *An extracted sentence may include not only common information, but additional information specific to the article from which it came, leading to source bias and* <*aggravating* [$_{\textbf{target}}$ *fluency*] [$_{\textbf{eval-n}}$ *problems*]> *in the extracted summary.*

As we can see, even though there seem to be some rules that might help to disambiguate *problem*, the evaluative or non-evaluative usage of nouns in this pattern can be very discipline-specific. Thus, the disambiguation procedure would have to be carried out for each discipline, separately. For this, the possibly evaluative nouns in this pattern located at the end of the annotated noun phrase could be extracted from the large version of SciTex and analyzed in context to determine which nouns

```
133985: ents vertices of a bridge <I> are the vertices in I th
133990: dge I are the vertices in <I> that are also in C. The
134254: ridge I. The two edges in <I> incident to these vertic
134288:  the vertices in I. Since <I> contains at most R + 1 v
165947: d in each component . Let <I> be the set of indices i
```

Figure 4.10: Extraction results from CQP of *I* not used as a personal pronoun

have an evaluative meaning. For illustration purposes, we have done this for computer science going through a list of approx. 3,000 nouns and checking whether they are used in an evaluative way. As a result, we came up with 10 nouns which we considered to be evaluative in computer science (*improvement*, *benefit*, *advantage*, *interest*, *challenge*, *importance*, *assistance*, *suitability*, *preference*, *simplicity*). If we consider that this procedure led to just 31 occurrences of evaluative expressions (i.e. for the 10 nouns altogether) extracted for computer science for the *target_eval-n* pattern, the amount of work invested is out of proportion to the results obtained. Thus, we decided to exclude this pattern from the annotation procedure.

4.4.2.1.3 Data extraction Having the features and their attributes annotated, the corpus can be queried for stance features, evaluative attributive patterns and targets evaluated, as well as their attributes (e.g., evaluative meanings). The xml-tags encoded in CQP are used for this purpose (see Table 4.2 for a full list). To obtain, e.g., distributional information on stance expressions, we use the <stance> xml-tags. The CQP-query to extract all instances of stance is:

```
<stance>[]+</stance>;
```

The square brackets stand for one token and the plus sign relates to possible additional tokens. Thus, the stance expression can have one, two or more tokens. The query looks for instances, as shown in Example (5), where *crucial* is the stance expression annotated.

(5) *A **crucial** question is the question* [...].

Similarly, we can query for evaluative attributive patterns with the <evaluation> tags and for the targets with the <etarget> tags as shown in the following queries:

```
<evaluation>[]+</evaluation>;
<etarget>[]+</etarget>;
```

Thus, the annotation allows us to extract the features more efficiently from the corpus. Instead of using a whole long query or multiple queries for one feature (as for the *eval-np_prep_target* pattern; see Figures 4.5), a simple query can be used:

```
<evaluation>[_.evaluation_pattern="eval-np_prep_target"]+</evaluation>;
```

After having performed the query, we can obtain distributional information on structural attributes with the CQP functions `group` or `tabulate`. The `tabulate` command is used to extract distributional information on different levels of annota-

```
SASCITEX2> tabulate Last match match text_id, match text_ad, match stance_type,
match stance_sem, match stance_meaning, match div_type, match pos;
Aho1977 A       hedge   epistemic       relativity      Abstract        JJ
Aho1977 A       hedge   epistemic       assumption      Abstract        JJ
Aho1977 A       hedge   epistemic       suggestion      Abstract        PP$
Aho1977 A       hedge   epistemic       relativity      Abstract        NNS
```

Figure 4.11: Extract of the output of the tabulate command for stance

text_id	text_ad	Abstract	Conclusion	Introduction	JJ	JJR	JJS	MD
Aho1977	A	10	46	67	119	5	6	81
Milner1977	A	4	0	65	69	3	0	62
Kwong1977	A	22	0	52	81	2	0	60
Dillon1974	C1	52	30	141	123	1	0	88
Rutenbar1984	B3	0	9	44	139	1	2	55

Figure 4.12: Extract of the matrix produced by the tabulate command in combination with the tbl2matrix Perl script

tion separated by tabulates. The command shown in Figure 4.11, e.g., tabulates information on stance by text (*text_id*), academic discipline (*text_ad*), stance type (*stance_type*), semantics of the stance expression *(stance_sem*), meaning of the stance expression (*stance_meaning*), document section (*div_type*), and the part-of-speech of the stance expression (*pos*) (see, again, Figure 4.11 for the output). The `tabulate` command is very efficient when used in combination with Perl scripts or sorting options offered by command-line systems. In our case, we use a Perl script[3], which summarizes the tabulate output into a matrix as shown in Figure 4.12. The data extraction process is automated in an extraction pipeline (see Kermes and Teich (2012) for a detailed description). The obtained matrix with all relevant features shown in Section 4.4.1.2 can then be used for data analysis.

4.4.2.2 Classification

To test our hypotheses of register diversification (H1) and registerial imprint (H2), we make use of a classification technique in order to measure the distance D for both H1 and H2 as illustrated in Section 4.4.1.1.

Text classification has been widely used in numerous studies, mostly based on bag-of-words representations developed in information retrieval and used, for example, in text categorization and stylometric studies (see, e.g., Joachims (1998); Koppel et al. (2002); Rybicki (2006); Argamon et al. (2007, 2008); Fox et al. (2012)). Basically, documents are represented by the words occurring in them, regardless of their ordering (cf. Joachims (1998)). Besides using simple bag-of-words approaches, other kinds of features can be used, such as features generated out of knowledge-based rules (e.g., Hayes et al. (1988)) or linguistic features based on linguistic theories (e.g., Argamon et al. (2007), Degaetano-Ortlieb et al. (2014a)). In our case, we use

[3]The Perl script tbl2matrix.perl, which was written by Hannah Kermes.

the linguistic features associated with stance, evaluative attributive patterns and targets obtained from our model of analysis and described in Section 4.4.1.2.

For the classification, we use support vector machines (SVM) (Vapnik and Chervonenkis, 1974), as they are known to obtain very good results on many relevant features (Joachims, 1998; Manning et al., 2008). In principle, what SVM does is to perform a binary classification trying to find a *hyperplane* in a multidimensional space that divides two classes from each other, where ideally instances of one class are on one side of the hyperplane and instances of the other class on the other side (cf. Baayen (2008); see Figure 4.13). The hyperplane is maximally far away from any data point and the distance from the hyperplane (decision surface) to the closest data point determines the *margin* of the classifier (see, again, Figure 4.13; cf. Manning et al. (2008)). The data points closest to the hyperplane are referred to as *support vectors* (shown in Figure 4.13). This method of distinction also implies that the decision function for an SVM is specified by a small subset of the data which defines the position of the hyperplane to separate the data (cf. Manning et al. (2008)).

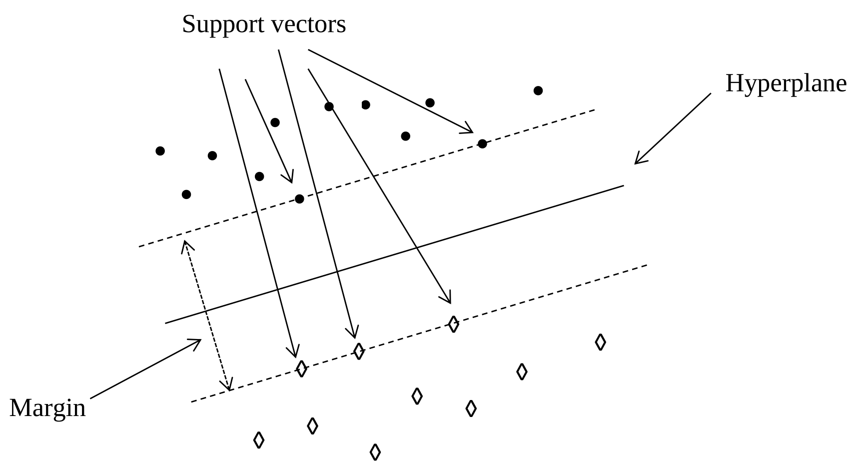


Figure 4.13: Hyperplane, instances and support vectors in a multidimensional space for SVM

For register diversification (H1), the classes are represented by the registers of the SciTex corpus, i.e. we perform 18 classifications (9 for each time period), each one with a single register vs. the other registers taken together (e.g., A vs. REST, B1 vs. REST, etc.). This will show how well one register can be classified from the other registers. Performing the classification on both time periods will then show whether in t2 a register classifies better than in t1, which would mean that the distance between one register and the others would become greater diachronically and the register would have encountered a process of diversification.

For registerial imprint (H2), however, we need to solve a multi-class problem and not a binary one, as we classify triples of A-Bs-Cs subcorpora (e.g., A-B1-C1, A-B2-C2). In this case, we use a pairwise classification, where a set of one-versus-one classifiers are built, i.e. for each triple the classification is performed on each register pair (e.g., A vs. B1, A vs. C1 and B1 vs. C1). This will show how well the contact register can be classified from its seed registers and whether there are diachronic changes by performing the triple classification on both time periods.

The SVM classification is performed with the data mining platform Weka (Witten et al., 2011). As input we use a matrix with normalized values of features for each text per register. The dataset can be manipulated according to specific needs by filtering options (such as smoothing) or for feature engineering by removing features (specified as attributes in Weka). As a filtering option we use smoothing to account for possible outliers in the data set, i.e. instances which vary greatly from the other instances in terms of attribute values. For this the MathExpression filtering is used and set to log(A+1). SVM classification is performed in Weka with the SMO classifier. We use a linear kernel and as test options we choose 10-fold cross-validation, i.e. the full set of available data is partitioned into 10 folds, which ensures that no instance is used simultaneously for training and testing preventing the evaluation results from reflecting only a particularly good/bad choice of the test set (cf. Resnik and Lin (2013, 279)). From the output of the SVM classifier, we consider the overall classification accuracy, i.e. the percentage of correctly classified instances. Additionally, several statistical evaluation measures are given for each class (e.g., precision, recall and F-measure). For our analyses, besides the overall accuracy, we will relate to the accuracy for each single class as well as the F-measure, which is the harmonic mean or weighted average of precision and recall (Van Rijsbergen, 1979; Powers, 2011). Furthermore, we also consider the confusion matrix (included in the output) which illustrates the number of texts classified in each class. The rows show the texts belonging to one class (e.g., A) and the columns represent the class into which the texts have been classified. Table 4.3 shows an example of a confusion matrix obtained by the classification of computer science texts (A) vs. the other registers (REST) as well as the respective accuracies and F-measures. For A, for example, 124 texts are classified correctly, while 78 texts are misclassified into the REST class, which leads to an accuracy of 61.4% for A and an F-measure of 0.703.

	A	**REST**	**total texts**	**accuracy**	**F-measure**
A	124	78	202	61.4	0.703
REST	27	2868	2895	99.1	0.982
overall				96.6	0.964

Table 4.3: Example of a confusion matrix for computer science vs. all other registers (70s/80s)

In our analysis, to test register diversification (H1), we closely inspect the confusion matrices to see not only which registers are relatively distinct (having fewer misclassifications) and which ones show less distinctness (i.e. more misclassifications and thus overlap with other registers), but also how this might have changed diachronically by comparing the classification results of t1 (70s/80s) and t2 (2000s). For registerial imprint (H2), we also closely inspect the confusion matrices which show three classes (e.g., A-B1-C1). Here, we gain insights into how well the contact registers classify when compared to their seed registers, i.e. with which seed registers they have the most overlaps (misclassifications), and whether this changes diachronically.

4.4.3 Micro-analysis: Feature analysis

While macro-analysis gives insights into general trends of register diversification and registerial imprint, we need a micro-analytical step to investigate which features either contribute to a distinction or are shared among registers.

For register diversification (H1), we are interested in whether there are features characteristic of a specific register. The better a given register is distinguished from the others, the more characteristic are the features contributing to the distinction for that particular register. For the micro-analysis of H1, we again need to consider both time periods (t1 and t2). As we hypothesize that the registers will be better distinguished in t2 than in t1, the features contributing to the distinction in t2 will be more characteristic for that register than the ones that contribute to the distinction in t1. Possible outcomes might be related to differences in, e.g., specific realizations of evaluative meaning, i.e. registers might be distinguished by one particular feature set (stance, pattern, target features) in t1, while they are distinguished by a different set in t2. For example, they might be distinguished by differences in evaluating different targets in t1, while they might be distinguished by the stance meanings (e.g., importance, desirability) used to evaluate targets in t2. They might also be distinguished by where the evaluative meaning is expressed, i.e. whether evaluative meaning is expressed in specific sections of research articles or prominently at sentence beginning, etc. To inspect the possible outcomes, we consider the SVM weights of the features for each classification. The classification outputs feature weights for each feature which are either positive or negative. The sign is an indication of the class the feature is associated with, i.e. as we test H1 by a binary classification of a specific register vs. all other registers (e.g., A vs. REST, B1 vs. REST, B2 vs. REST), we have either features associated with that particular register or features associated with the rest. Here, we are particularly interested in the features characteristic of the single register (i.e. A, B1, B2 etc.). The SVM weights represent a feature ranking, i.e. the greater the weight the more distinctive a particular feature is for a register. To determine characteristic features and possible diachronic tendencies, we inspect the top 20 features of each classification and compare them for t1 and t2 for the single registers, i.e. we consider nine diachronic comparisons, one for each SciTex register.

For registerial imprint (H2), we are interested in the features adopted by the contact register from the seed registers. We investigate again the SVM weights for each classification to determine which features are associated with which register and which contribute to a possible overlap between the contact and seed registers. If we take the A-B1-C1 triple, for example, the classification outputs SVM weights for each classification pair, i.e. A vs. B1, A vs. C1 and B1 vs. C1. The negative feature weights are associated with A in both pairs and with B1 in the B1-C1 pair. The positive feature weights are associated with C1 and with B1 in the A-B1 pair. This is the case for all triples. Again, the greater the SVM weight of a feature, the stronger its contribution to the classification of the register the feature is associated with. When looking at the features, we are particularly interested in features characteristic of the Bs and which are also characteristic of either A or the respective C, i.e. which are shared among the contact register and its seed registers. Consider again the A-B1-C1 triple. If some of the top 20 features are the same for B1 (vs. A) and C1 (vs. A), there is an overlap of characteristic features for both registers when compared to A. Looking at the number of features n shared among the top 20 for B1 and C1 and for B1 and A indicates with which seed register the contact register B1 shares more characteristic features. Thus, we say that B_i is more "similar to C_i" if the number of features shared among B_i and C_i is greater than for B_i and A. We also say that B_i is more "similar to A" if the number of features shared among B_i and C_i is smaller than for B_i and A. For this, we can formulate the following:

if $n(B_i, C_i) > n(B_i, A)$ then $B_i \approx C_i$
if $n(B_i, C_i) < n(B_i, A)$ then $B_i \approx A$

In addition, we consider the diachronic perspective, i.e. we inspect the number of features shared by the contact register in t1 (n_{t1}) and t2 (n_{t1}) to observe possible diachronic changes or consolidation trends:

if $n_{t1}(B_i, C_i) > n_{t1}(B_i, A)$ and $n_{t2}(B_i, C_i) > n_{t2}(B_i, A)$ then $B_i \approx C_i$
if $n_{t1}(B_i, C_i) > n_{t1}(B_i, A)$ and $n_{t2}(B_i, C_i) < n_{t2}(B_i, A)$ then B_i shifts toward A
if $n_{t1}(B_i, C_i) < n_{t1}(B_i, A)$ and $n_{t2}(B_i, C_i) < n_{t2}(B_i, A)$ then $B_i \approx A$
if $n_{t1}(B_i, C_i) < n_{t1}(B_i, A)$ and $n_{t2}(B_i, C_i) > n_{t2}(B_i, A)$ then B_i shifts toward C_i

Thus, if the number of features characteristic of B_i and C_i (vs. A) is greater than the number of features characteristic of B_i and A (vs. C_i) in both time periods (t1 and t2), then the contact register uses evaluative meaning similarly to its C seed register. If, instead, the number of features characteristic of B_i and C_i (vs. A) is greater than the number of features characteristic of B_i and A (vs. C_i) in t1 but smaller in t2, then the contact register exhibits a diachronic shift from expressing evaluative meaning more similarly to its C seed register to expressing it more similarly to computer science (A). This can clearly also be the other way around, i.e. the contact register can be more similar to computer science (A) in both time periods or shift to its C seed register in t2, while it shared more features with A in t1.

Furthermore, we investigate which features are "adopted" by the contact register from the seed registers. For this, we consider the 10 most characteristic features of each B register vs. its C register and vs. A and reclassify the triple based on these features.[4] Specifically, we say that a feature f is "adopted from A" if in the comparison of A vs. B_i the feature is typical of A AND if the same feature is typical for B_i in the comparison with C_i. We also say that a feature f is "adopted from C_i" if in the comparison with C_i vs. B_i the feature is typical of C_i AND if the same feature is typical for B_i in the comparison with A. For this, we formalize the following:

if $f(A, B_i) = f(B_i, C_i)$ then f is adopted from A by B_i
if $f(C_i, B_i) = f(B_i, A)$ then f is adopted from C_i by B_i

Adopting this micro-analytical step for both hypotheses will help us to gain a better understanding of how and which kinds of linguistic diversification trends occur for highly specialized registers, and which linguistic differences or commonalities exist between contact and seed registers in terms of registerial imprint.

4.4.4 Micro-analysis: Ensure data quality and gain deeper insights

While the feature analysis of the first micro-analysis shows which features of evaluative meaning are characteristic of a register for H1 and which features are adopted from the contact registers from their seed registers for H2, as well as possible diachronic shifts, a further micro-analytical step positioned closer toward "close reading" serves to confirm the features obtained from the feature analysis. This step not only ensures data quality, but also leads to a deeper linguistic understanding of how evaluative meaning is expressed across registers.

For this purpose, we inspect each of the top 20 features of the classification in detail. This encompasses two steps: (1) inspect features within concordances, (2) look at the distributional information of each feature considering further attributes, i.e. looking at distributional information not only of the features themselves but also, e.g., of the lexical items they are realized by in the case of specific meanings (such as importance), of the evaluative patterns they are realized in, etc.

4.4.4.1 Data quality checks

During this micro-analytical inspection we still encountered some noise problems within our data. With an additional revision step on the features most relevant to each register, data quality was enhanced further, while trying to avoid as much noise

[4] Doing so, we can inspect which features typical for B_i are even more typical for A or C_i.

```
rithm is given whose time <complexity> is linear in the size of
ow that the computational <complexity> of generating efficient
timal code and whose time <complexity> is linear in the size of
is optimal and is of time <complexity> Oip 2s ) , where p is th
 affect the computational <complexity> of code generation ? Is
```

Figure 4.14: Concordances of complexity in computer science (70s/80s)

as possible. What we mostly encountered in terms of noise were problems of lexical items not being used in an evaluative way and data sparsity which would lead to non-representative results. We will illustrate the procedure we used to avoid these problems through examples.

4.4.4.1.1 Non-evaluative lexical items While we more closely investigated top 20 features that are characteristic of each register, some features turned out to still present problems in terms of not being used in an evaluative way. This was particularly the case for the stance features, which were annotated in terms of lists of lexical items, rather than the attributive patterns or targets, which had undergone a more detailed procedure to ensure data quality during annotation. Additionally, the data quality procedure during annotation was focused on the whole corpus rather than on single registers, as it would not have been possible to go through every lexical item (approx. 540 items) annotated as a stance expression in register-specific terms. Instead, we decided to use a more detailed data quality procedure on features that turned out to be characteristic for a register from the classification in the first round of the analytical cycle to enhance data quality on relevant features for a second round, and so forth.

While adjectives and adverbs have an inherent modifying function, nouns and verbs are often ambiguous in terms of evaluative meaning. If we consider the noun *claim* from the assumption meaning in computer science, most of the times it will not be evaluative, as it is related to a formalized statement (see Example (6) in bold). In linguistics, instead, it mostly expresses assumption (see Example (7)). Thus, we would have to disambiguate the evaluative from the non-evaluative function.

(6) ***Claim*** *6.4. At any time the size of the occupied space in the buffer of GREEDY(B , 1) is at most the size of the occupied space* [...].

(7) [...] *sign-language agreement is not linguistic at all, a **claim** with which we firmly disagree.*

To exemplify the methodology, we have looked more closely into the usage of *claim* across registers. We have evaluated all occurrences of *claim* in computer science (598) and linguistics (398) in SciTex 2000s. While for computer science all occurrences are non-evaluative, for linguistics most occurrences are evaluative. Non-evaluative cases in linguistics are shown in Example (8) and Example (9). In Example (8) *claim*, more explicitly *universal claim*, is not used as a claim of scientific discourse by the author of the article. It is rather a term which does not necessarily encompass an evaluative meaning. In Example (9) the meaning of *claim* is equal to *have the right to* and not to *assertion*. Nevertheless, besides these examples (which are not related to the assumption meaning), and wrongly tagged occurrences (in which the verb was mistaken for the noun), 390 out of 398 cases are actually evaluative in linguistics.

(8) [...] *in both cases Maya is seen as offering a counterproposal, replacing Iddo's* ***universal claim*** *with a partial generalization* [...].

(9) [...] *the fact remains that there are other expressions that have an equal* ***claim*** *to be called proper names* [...].

What we have also noted by this closer inspection are some differences in use that distinguish both semantic variants (evaluative from non-evaluative). For the variant used in computer science, the following observations can be made: *claim* is

- very specific, always preceded by the definite article *the*,
- possibly preceded by a non-evaluative adjective (e.g., *inductive*, *second*, *following*),
- used with verbs such as *hold*, *prove* or *follow*,
- used as a headline (e.g., *Claim 6.9*),
- most often used in the singular form (approx. 92%).

In linguistics, on the other hand, *claim* is

- further specified by the complementizer *that* (approx. 40%) or by prepositional phrases (approx. 20%),
- more often used in the plural form than in computer science (compare 27% with 8% in computer science),
- followed by the verb be and evaluative adjectives (e.g., *true*, *wrong*, *right*) or verbs (e.g., *falsified*, *supported*).

Some of the observations can be operationalized in terms of macros for the annotation, such as the use of *claim* in a headline in computer science. Others cannot be easily operationalized. The use of the complementizer or the plural vs. the singular form, for example, is not really indicative of the evaluative or non-evaluative use. The distinction is more a semantic one, where in computer science *claim* is used as a term, while in linguistics it is used to indicate an assumption. From the observations made by this investigation, we can decide to annotate claim as an assumption in linguistics but not in computer science. However, for the other registers, especially the contact registers, the distinction is not as clear, although they show a tendency. In computational linguistics, for example, most of the occurrences are evaluative (84%), while the others are either wrongly tagged (*claim* tagged as a noun even though it is a verb) or non-evaluative. In this case, we decided to annotate it as evaluative.

Even though this kind of procedure enables us to enhance data quality, it is too time-consuming to be applied on all lexical items for each feature investigated. Thus, we decided to deal with this on the most distinctive features for each register to ensure high-quality data for the features that predominantly characterize a register. This micro-analytical step will lead to new data which is then used to inform the macro-analysis in a new round of the analytical cycle.

4.4.4.1.2 Data sparsity As we are interested in features characteristic of a given register, it is essential that the features are somehow representative for that register. However, one additional issue we have to deal with is data sparsity, i.e. even though features are characteristic by the SVM weights for a particular register, they still cannot be representative of a register, if they occur, e.g., just once in the subcorpus.

Consider, e.g., the stance meaning of purpose (*S_purpose*) which was among the top 20 features for computer science, but occurred just once in the dataset for A (see Example (10)). Even though the feature is among the top 20, we have to omit it from the computer science dataset, as it cannot be said to be characteristic of that register in terms of our study.

(10) [$_{\textbf{eval-adv-of-purpose}}$ *Deliberately*], *no attempt was made to achieve simplifying reductions.* (SciTex 70s/80s, A: Computer science)

By this closer inspection of features, we have seen that the dataset has to be revised to obtain a more reliable analysis. To obtain this, we revise the extraction of the features for each register by excluding lexical items that have shown to be used mostly in a non-evaluative way in a particular register, as well as features which occur less than five times. Note that in these cases, we also consider that the features should not be present in only a single text. This is to avoid presenting characteristic features of a register that in reality would be characteristic of just one single text.

This procedure is adopted for both hypotheses (H1 and H2). As our analytical cycle is based on recursive steps, the top 20 features are checked until our criteria of data quality and number of occurrences are met.

4.4.4.2 Concrete linguistic insights

Besides testing for data quality, the inspection of concordance lines and distributional information also leads to deeper linguistic insights into how evaluative meaning is expressed across registers. This will be illustrated for each register within the analyses. Considering our hypotheses, for register diversification (H1), we expect to gain a better understanding of the diversification process of the registers, i.e. how they differ in linguistic terms from the other registers and how they have evolved over time. For example, the feature of *S_self-mention*, which relates to the writer's presence in the discourse, might be distinctive for more than one register, however, there might be differences in how self-mention is realized across registers and time periods. Just to give a very simple example, we might encounter a preference for the singular form in one register (see Example (11)), while the plural form (see Example (12)) is preferred in another register, or we might encounter some kind of preference in one time period as opposed to the other.

(11) *In the earlier article* [$_{\textbf{self-mention-sg}}$ *I*] *noted that, for some speakers, basic forms such as* [...]. (SciTex 70s/80s, C1: Linguistics)

(12) [$_{\textbf{self-mention-pl}}$ *We*] *ignore constraints that may be introduced into the dag because of side effects.* (SciTex 70s/80s, A: Computer science)

For our hypothesis of registerial imprint (H2), we expect to gain insight into specific linguistic developments of the contact registers with respect to their mixed nature. By inspecting the features quite closely, we might find that some contact registers adopt, e.g., a specific feature from one of their seed registers, but the usage might be different. Consider, e.g., the target *algorithm* which might be evaluated both in computer science and the contact registers, due to the fact the contact registers adopt it from computer science. Even though the target itself is the same, the evaluation of the target might differ. Consider, here, Examples (13) and (14), where the target *algorithm* is said to perform best in computer science, while it is said to be appropriate for use in computational linguistics.

(13) *If $m = n$, the new algorithms given in Examples V and VI* [...] *are the* [$_{\textbf{eval-adj-of-desirability}}$ *best*] [$_{\textbf{target}}$ *algorithms*] *known.*
(SciTex 70s/80s, A: Computer science)

(14) *Moreover, our following observations and suggestions are meant only to serve as a point of departure for the computational linguist and the computer technician concerned with the practical application of linguistic analysis to linguistic computation (i.e., the devising and programing of the* [$_{\textbf{eval-adj-of-suitability}}$ *appropriate*] [$_{\textbf{target}}$ *algorithms*]*)* [...].
(SciTex 70s/80s, B1: Computational linguistics)

4.5 Summary

To obtain a more comprehensive picture of how evaluative meaning may differ across registers, i.e. by looking at possible diversification trends, and how contact registers are affected by possible diachronic trends with respect to their seed registers, we designed an analytical cycle which encompasses different levels of macro- and micro-analysis.

The different macro- and micro-analytical steps are not considered separately but are interrelated and inform each other. In this sense, the designed analytical cycle allows us to adopt recursive quantitative as well as qualitative analytical steps and methods that make it possible to create a balance between generalizable trends and fine-grained insights gained by the analyses. The adopted methodology also allows us to enhance data quality in a way that is manageable considering the number of features investigated (approx. 270) and the dimensions of comparison adopted (single registers, register triples, and time). This will allow an empirical study on evaluative meaning which also accounts for fine-grained linguistic insights on how evaluative meaning is expressed across registers.

In terms of macro-analysis, we adopt text classification as a method, which enables us to consider possible differences and commonalities between registers. Here, we rely on empirical observations of how well text productions of one register classify into the right class. Performed on both time periods to test H1 (register diversification), this reveals possible diversification trends, i.e. whether registers classify better over time, which would make them more distinct linguistically. Additionally, by testing H2 (registerial imprint), classification is performed on triples of subcorpora (A-B-C) to inspect overlaps between contact registers and their seed registers.

In terms of micro-analysis, we first employ feature analysis, which allows us to observe characteristic features that lead to register diversification as well as features that reflect the contact registers' mixed nature, i.e. that point to a possible registerial imprint passed on by the seed registers. Second, the more detailed analysis of features within concordances and the account of distributional information on different levels not only leads to enhanced data quality, but also allows a fuller understanding of the expression of evaluative meaning across scientific registers, as well as its linguistic development.

Chapter 5

Register analysis

5.1 Introduction

In this chapter, we empirically analyze how scientific registers differ in terms of evaluative meaning and whether this has changed diachronically. The aim is to investigate our two hypotheses of register diversification (H1) and registerial imprint (H2). By our model of analysis, we have specified how evaluative meaning is possibly expressed in scientific writing and which lexico-grammatical features are involved (see Section 3.4). Moreover, based on our analytical cycle of macro- and micro-analyses, we are able to empirically investigate these two hypotheses (see Section 4.4). We pursue answers for H1 and H2 in two separate analyses.

For the investigation of register diversification (H1), we pose the following questions:

- Are highly specialized registers distinct from one another in terms of evaluative features, i.e. do evaluative features contribute to a distinction between registers?
- Does their distinctness increase over time, i.e. can we observe register diversification trends?

Answers to these questions are pursued in terms of macro- and micro-analysis. In macro-analysis, we consider how well the registers are classified diachronically considering classification accuracy, precision, recall and F-measure as well as confusion matrices. On this level, we describe possible diversification trends on a relatively general base, i.e. we consider the *degree of diversification*. In micro-analysis, the SVM weights of the top 20 features are considered and analyzed in detail. The top 10 features are then considered to be characteristic, giving insights into concrete tendencies of diversification related to specific lexico-grammatical features of evaluative meaning characteristic of a register. This will point to the *kind of register diversification*.

For the investigation of registerial imprint (H2), the following questions are investigated:

- Do contact registers share commonalities with their seed registers, i.e. by which seed register is the contact register influenced most?
- Are these commonalities preserved over time or do they change, i.e. does the influence of one seed register diminish, while it increases for the other seed register?

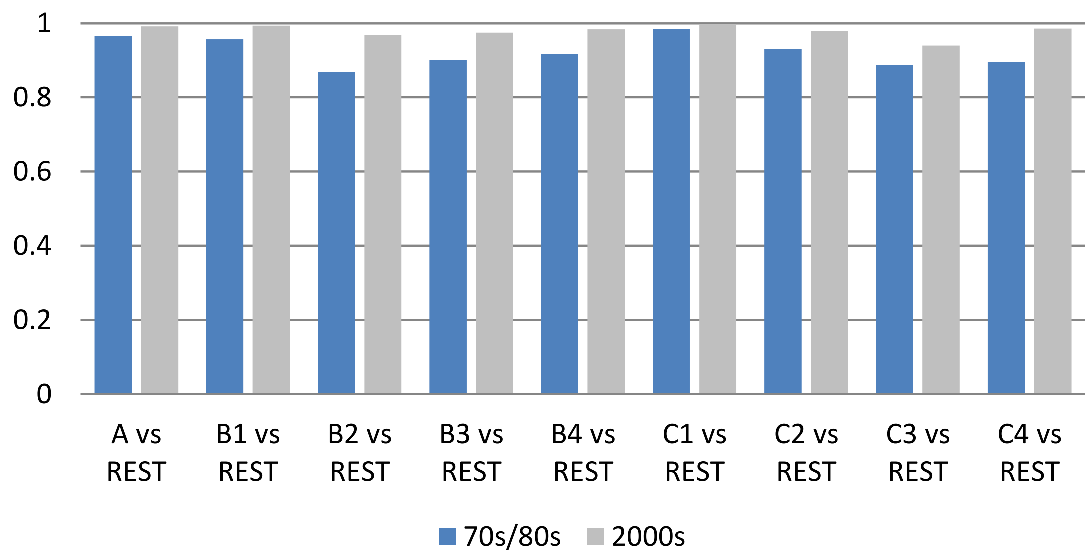
Again, we adopt a macro- and micro-analytical perspective to answer these questions. On the macro level, we consider how well contact registers are classified when compared to their seed registers. Here, we look at whether they show more overlaps with one seed register in comparison to the other and if this changes diachronically. Therefore, we consider the *degree of registerial imprint*. On the micro level, we investigate further which commonalties are shared between a contact register and its seed registers, i.e. which lexico-grammatical features of evaluative meaning are used similarly. Here, we consider the *kind of registerial imprint*, which is also investigated diachronically. In addition, besides similarities, i.e. features that are shared between contact and seed registers, we also consider whether features are adopted from one or the other seed register by the contact register. Note that if we would just consider the features that are shared, we still would not know if they are adopted by the contact register from the seed register.

5.2 Register diversification

5.2.1 Degree of register diversification

The study of the degree of register diversification is at the macro level of analysis. We empirically analyze register diversification by classification of texts with SVM as described in Section 4.4.2.2. For both time periods, we perform nine classifications each (18 classifications in sum), classifying a single SciTex register (e.g., A, B1, C1) vs. all other registers taken together (REST). This approach will indicate the distance of one register vs. the others and whether this distance increases over time, pointing toward register diversification.

We first consider the overall classification accuracies shown in Figure 5.1. We can see that all pairs have a quite high overall accuracy of 87–97% for the 70s/80s and 94–99% for the 2000s. Diachronically, the classification accuracies increase over time for each register. However, in order to consider how well the single registers classify in comparison to the rest, we inspect precision (misclassified texts from the REST class placed into a given register class), recall (correctly classified texts for



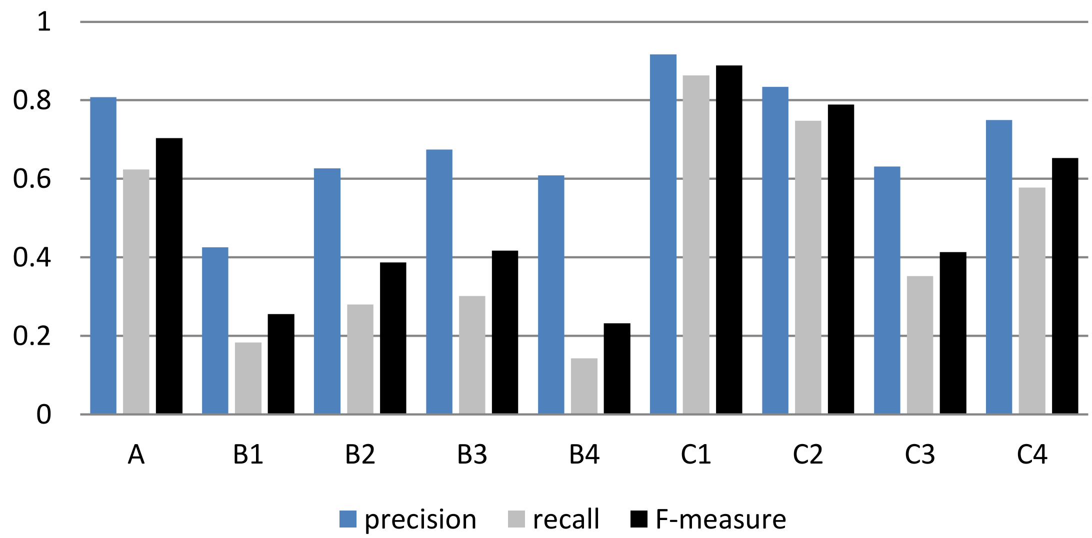
A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

Figure 5.1: Overall classification accuracies across registers for both time periods

a given register class) as well as the F-measure (which accounts for both recall and precision). In the 70s/80s, we see from Figure 5.2 that especially for the contact registers, recall is quite low, which is also reflected in a relatively low F-measure. In the 2000s, on the other hand, Figure 5.3 shows how precision, recall and the F-measure are relatively high for all registers.

Considering the F-measure for the single registers diachronically (see Figure 5.4), we clearly see how all contact registers have a very low F-measure in the 70s/80s (B1: 0.24; B2: 0.39; B3: 0.42; B4: 0.23), being less well distinguished from the other registers in terms of evaluative features. The seed registers, instead, show higher F-measures, being better distinguished. Thus, the contact registers do not present a distinct use of evaluative features in the 70s/80s, while their seed registers do. Note that the best results are achieved by the seed register C1 (linguistics). In the 2000s, however, the F-measure is quite high for all registers. Thus, there is clearly some kind of diversification process that shows up in all registers. Still, the contact registers (Bs) in particular show a high increase in F-measure, by approx. 50–70%, while the seed registers (A and Cs) have increased by only approx. 8–40% (see Table 5.1).

As the registers are better classified in the 2000s in comparison to the 70s/80s, they have moved toward a more distinct use of lexico-grammatical features of evaluative



A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

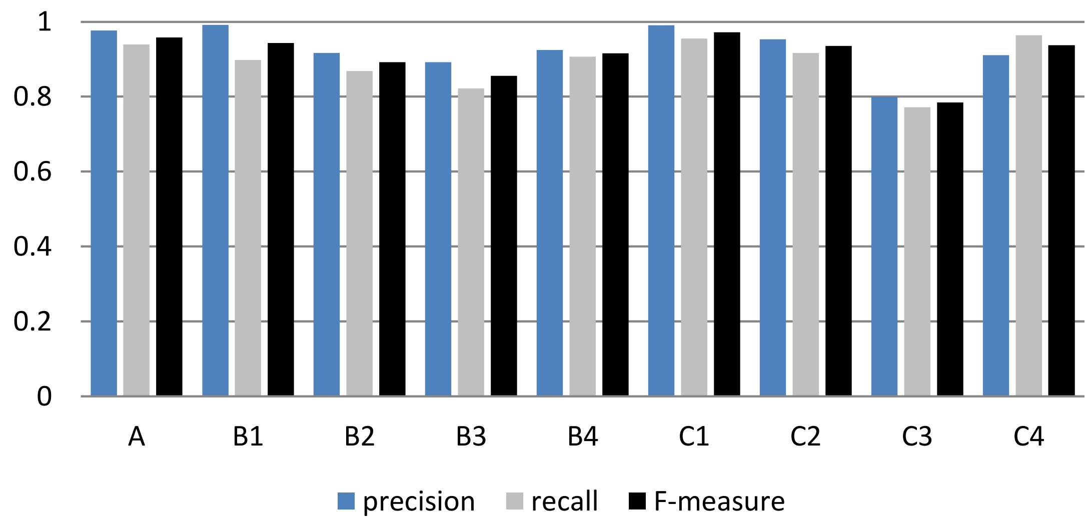
Figure 5.2: Precision, recall and F-measure across registers for the 70s/80s

	70s/80s	**2000s**	**increase in %**
A	0.704	0.962	25.80
B1	0.256	0.943	68.70
B2	0.387	0.906	50.05
B3	0.417	0.888	47.10
B4	0.232	0.903	67.10
C1	0.889	0.972	8.30
C2	0.789	0.935	14.60
C3	0.413	0.785	37.20
C4	0.653	0.937	28.40

A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

Table 5.1: Increase of F-measure from 70s/80s to 2000s

meaning. Thus, they have undergone a process of register diversification over time and their distance from the other registers has increased. This is strongly reflected by the contact registers computational linguistics (B1), bioinformatics (B2), and mi-



A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

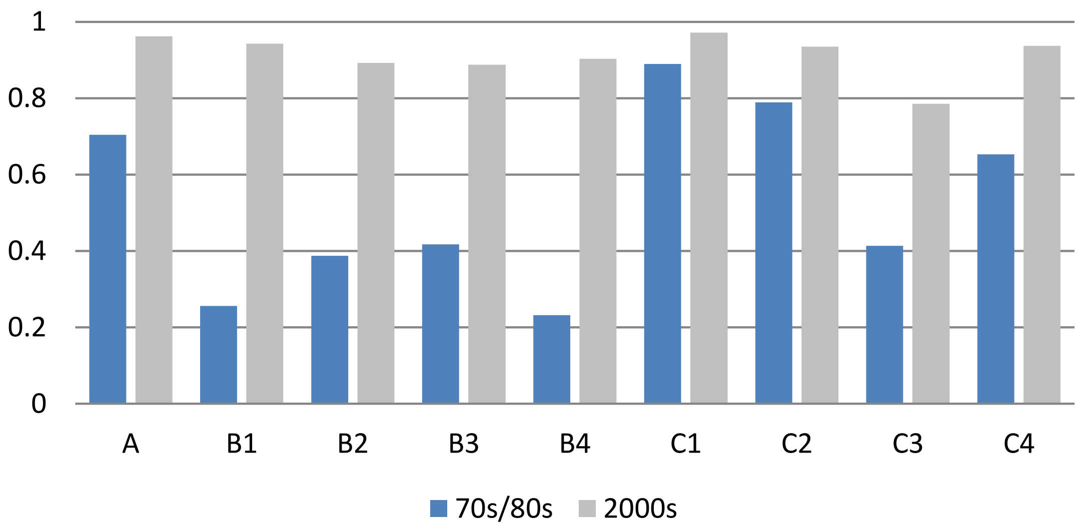
Figure 5.3: Precision, recall and F-measure across registers for the 2000s

	REST	**B1**	**misclassifications in %**
REST	2940	31	1.04
B1	103	23	81.74

Table 5.2: Confusion matrix for computational linguistics (B1) in the 70s/80s

croelectronics (B4) as they achieve an F-measure above 0.90 (see, again, Table 5.1).

This tendency is also reflected in the confusion matrices for each classification pair, again, especially for the contact registers. Consider, e.g., the confusion matrices for B1 (see Tables 5.2 and 5.3). In the 70s/80s, most texts were misclassified from B1 into the REST class (81.74%), i.e. 103 texts could not be identified as text productions of computational linguistics in terms of features of evaluative meaning. Additionally, 31 texts (1.04%) were wrongly classified as text productions of computational linguistics from the REST class. This changes in the 2000s, as only 10.22% of the texts are misclassified from B1 into the REST class and 0.05% of texts are wrongly misclassified into B1 (see, again, Table 5.3). Therefore, computational linguistics seems to develop a distinct use of evaluative features over the 30-year time span investigated. The same holds for B2, which in the 70s/80s has 71.99% misclassified text and only 13.17% in the 2000s, B3 with 69.78% misclassifications in the 70s/80s and 14.46% in the 2000s, and B4 with 85.66% misclassifications in the



A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

Figure 5.4: F-measure across registers for both time periods

	REST	**B1**	**misclassifications in %**
REST	1983	1	0.05
B1	14	123	10.22

Table 5.3: Confusion matrix for computational linguistics (B1) in the 2000s

70s/80s and only 11.22% in the 2000s (see Table 5.4). Thus, bioinformatics (B2), digital construction (B3) and microelectronics (B4) also develop a distinct usage of evaluative features, presenting a quite strong diversification process over time.

Comparing the percentages of misclassifications of contact and seed registers (refer again to Table 5.4), we can see that all contact registers have almost 10–15% of misclassifications, while the seed registers have around 8–3%, except for C3 which has the highest amount of misclassification in the 2000s with approx. 23%. Thus, the contact register digital construction (B3) and its seed register mechanical engineering (C3) are less well distinguished by evaluative features in the 2000s than the other registers.

In summary, the contact registers in particular have clearly undergone a process of diversification, increasing their distance from the other registers in terms of expressing evaluative meaning over time. Computational linguistics (B1) shows the highest amount of diversification on the macro level, with an increase in distance of approx.

register	70s/80s	2000s
A	37.62	5.65
B1	82.54	10.22
B2	71.99	13.17
B3	69.78	14.46
B4	85.66	11.22
C1	12.79	5.41
C2	25.23	8.26
C3	68.38	22.77
C4	42.16	3.59

A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

Table 5.4: Percentage of misclassifications across both time periods

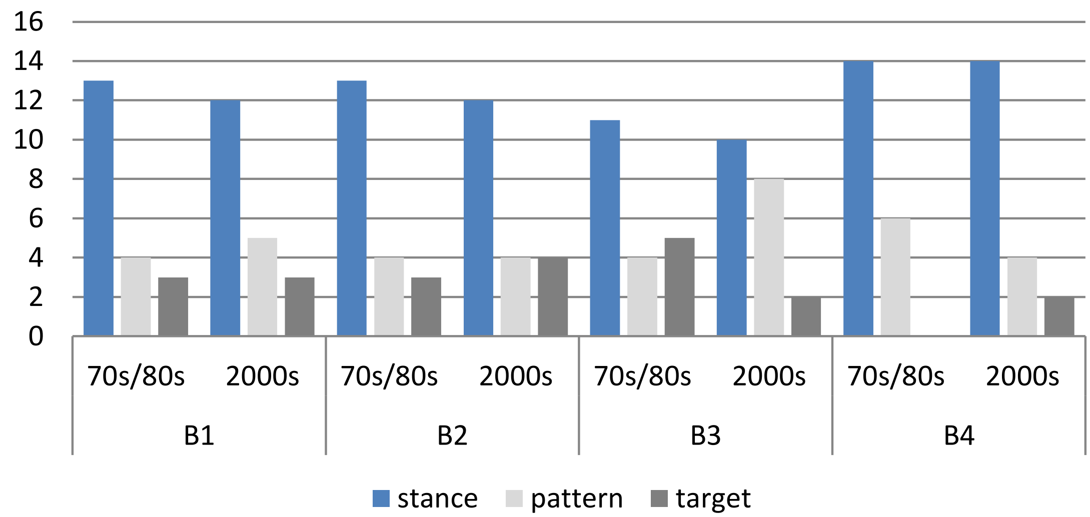
69%, closely followed by microelectronics (B4) with approx. 67%.

5.2.2 Kind of register diversification

From the macro level, we have seen that the contact registers in particular do present a register diversification process in the time period investigated. So far, we have inspected the degree of diversification. To investigate which specific kind of register diversification adheres to the contact and seed registers, we adopt a micro-analytical perspective by investigating more closely the features that contribute to a better distinction over time.

As we have seen from the macro-analysis, for the 70s/80s, the F-measures for the contact registers are quite low (refer again to Table 5.1). Thus, their top-ranking features by SVM weight cannot really be said to be characteristic of that particular register in that time period. However, we can investigate general tendencies in diachronic terms by comparing top-ranking feature types of each time period, i.e. whether the types have changed over time, contributing to a better distinction. For the 2000s, instead, as the classification presents quite high F-measures for the contact registers, pointing to a distinct use of evaluative features, we investigate the top-ranking features that contribute to a distinct usage and consider them to be characteristic for these registers.

For the seed registers, in contrast, the F-measure is relatively high in both time periods, especially for linguistics (C1), biology (C2) and computer science (A) (see, again, Table 5.1). Thus, we inspect these three seed registers diachronically in terms of the kind of register diversification trends they show with respect to features contributing to a better distinction. This will allow us to investigate more detailed



B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics

Figure 5.5: Number of feature types across contact registers for both time periods

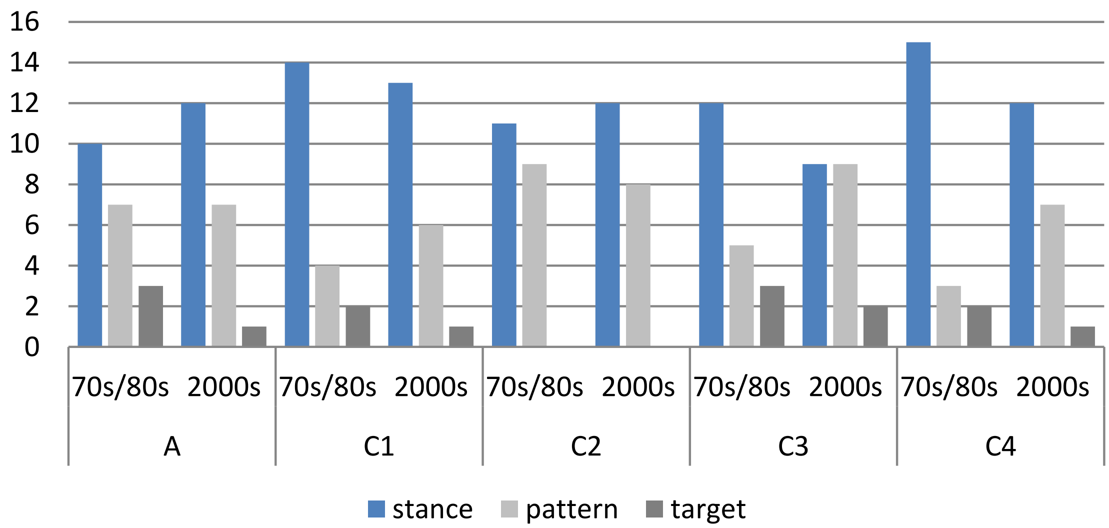
diachronic diversification trends of quite established registers.

5.2.2.1 General diachronic tendencies for contact and seed registers

To inspect diachronic tendencies of the contact registers, first, we look at the feature types (stance, pattern, target) of the top 20 features ranked by SVM weights for both time periods. Here, we consider the number of features of each type among the top 20 features. As we can see from Figure 5.5, for all contact register the number of stance features predominates with 10–14 features, while the number of pattern and target features among the top 20 features ranges from only 2–8. Diachronically, the distributions of the number of features across types are not significantly different (p-value > 0.05). Thus, the feature types are similarly used across registers and time. Considering the seed registers, we get a very similar picture as shown in Figure 5.6, i.e. stance features prevail, followed by pattern and target features.

Moving toward more fine-grained differences, we look at the feature subtypes. For this we consider:

(1) for stance, the meaning of stance expressions (*S_meaning*, e.g., importance, desirability), the parts-of-speech of stance expressions (*S_pos*, e.g., common singular noun (NN), adjective (JJ)), the type of stance expressions (*S_type*, e.g., self-mention, booster), and the document section in which stance expressions occur (*S_doc*, e.g., Abstract, Introduction),



A: computer science; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

Figure 5.6: Number of feature types across seed registers for both time periods

(2) for patterns, the pattern type (*P_type*, e.g., eval-adj_target-np) and the stance meaning expressed within patterns (*P_meaning*, e.g., obviousness, importance), and

(3) for targets, the lemma of the target (*T_lemma*, e.g., algorithm, function) as well as its part-of-speech (*T_pos*, e.g., common singular noun (NN), proper singular noun (NP)).

Table 5.5[1] shows the number of top 5 features according to the feature subtypes across contact registers and for both time periods.[2] While for the 70s/80s, most features belong to the stance meaning (*S_meaning*), the top 5 features for the 2000s show more variation in terms of feature subtypes. This implies that the contact registers develop, over time, a distinct usage of evaluative meaning using a wider range of feature subtypes, rather than differing only by particular stance meanings. Consider that the contact registers classify much better in the 2000s, which seems to be interrelated with the variation of feature types among the top 5 features for each contact register. By calculating the Pearson's correlation between different

[1]Each black box represents one feature. All boxes in one column add up to 5. Note that the Sent. b. row, in contrast, has to be considered separately, as it shows how many of the top 5 features are used at sentence beginning.

[2]Here we have chosen the top 5 features only, due to the low classification performance of the contact registers in the 70s/80s.

	70s/80s				**2000s**			
	B1	**B2**	**B3**	**B4**	**B1**	**B2**	**B3**	**B4**
S_meaning	■■■	■■■■	■■■	■■■■	■	■	■	■■
S_pos		■		■	■			
S_type			■		■■	■■■	■	
S_doc							■	■■
P_type								
P_meaning							■■	
T_lemma	■■		■		■			
T_pos						■		■
Sent. b.		■	■		■■■	■■	■■■	■■

B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics

Table 5.5: Number of top 5 features according to feature subtypes for the contact registers

feature types and the classification F-measures for the contact registers, we obtain a value of 0.87, which also indicates that there seems to be a correlation between more variation in feature types and better classification performance. By considering the top 5 features of the seed registers (see Table 5.6), we can see that diachronically the distribution across feature subtypes is more homogeneous in comparison to the contact registers, except for C3 (mechanical engineering), whose top 5 features are all stance meanings in the 70s/80s (similarly to the contact registers), while in the 2000s it shows more variation in feature subtypes. Considering again the classification performance, C3 obtains a relatively low F-measure (0.413) in the 70s/80s, while in the 2000s it is higher (0.785) (see, again, Table 5.1). This also enforces the fact that different feature subtypes within the most distinctive features seem to lead to a better classification performance, and thus to a better distinction. Considering that different feature subtypes can be interrelated, i.e. they can be related to the same evaluative phenomenon to different extents, this might point toward a construal of a characteristic feature over time, which culminates in being reflected across several feature subtypes, leading then to a better classification performance. Consider, for example, the stance meaning of obviousness (*S_obviousness*). As this stance meaning reflects epistemic expressions, it is related to the feature of stance meanings of the epistemic type (*S_epistemic*). Additionally, obviousness is a boosting expression, so it is related to stance expression of boosters in general (*S_booster*), and so on. Therefore, while a characteristic usage develops, it might be reflected across different feature subtypes as it becomes more and more distinctive for a register.

An additional observation is made by considering whether the features occur at the beginning of a sentence (see *Sent. b.* in Table 5.5). In the 70s/80s, almost none of the top 5 features are positioned at sentence beginning for the contact registers. In the 2000s, however, most features of the top 5 features are located at the beginning

	70s/80s					2000s				
	A	**C1**	**C2**	**C3**	**C4**	**A**	**C1**	**C2**	**C3**	**C4**
S_meaning	■■	■■■	■	■■■■■	■■■	■	■■	■	■■	■■
S_pos	■		■		■	■	■■	■■		
S_type		■	■			■■	■	■	■■	
S_doc	■	■			■	■			■	
P_type			■							
P_meaning	■		■					■		■■■
T_lemma										
T_pos										
Sent. b.	■■		■■	■	■	■■■■	■	■■	■■	■

A: computer science; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

Table 5.6: Number of top 5 features according to feature subtypes for the seed registers

of a sentence. This diachronic tendency is also observed for the seed registers (refer again to Table 5.6). This clearly indicates a diachronic development of expressing evaluative meaning distinctly at sentence beginning, which is somehow related to information structure, i.e. evaluative meaning is distinctively used at the beginning of a sentence rather than somewhere else, and conventionalization, i.e. diachronically for all registers, evaluative meaning seems to be used more toward the beginning of a sentence.

5.2.2.2 Characteristic features and diachronic tendencies of evaluative meaning for the seed registers

To investigate more detailed diachronic diversification trends on characteristic features of evaluative meaning, we have selected three seed registers which show the best classification performance (computer science, linguistics and biology). Given that these registers show high classification accuracies for both time periods, their distinctive features can be said to be characteristic for the registers for the 70s/80s and the 2000s. Thus, we will be able to exemplify diachronic trends which are related to more established registers. For this, we move toward the more qualitative range of the scale of micro-analysis.

5.2.2.2.1 Computer science For computer science, Tables 5.7 and 5.8 show the top 10 features for the 70s/80s and 2000s, respectively. There are three main trends that can be observed diachronically: (1) the attitudinal meanings of complexity and progress are distinctive for both time periods, while the epistemic meaning

feature	examples	Sent. b.	subtype	S_sem	weight
S_complexity	*easy*, *difficult*		S_m	attitudinal	-1.61
S_progress	*new*, *novel*		S_m	attitudinal	-1.60
P_Abstract	*an important factor*		P_doc		-1.48
Pb_idiosyncrasy	*Specifically*, [...]	x	P_m	attitudinal	-1.31
Sb_PP	*We*	x	S_pos		-1.26
T1_algorithm	*most well-known algorithms*		T_lemma		-0.87
Pb_eval-adj_target-n	*An important operation*	x	Pb_type		-0.83
S_Introduction			S_doc		-0.79
S_PP	*we*, *us*		S_pos		-0.77
Sb_self-mention	*We*, *Our*	x	Sb_type		-0.77

Table 5.7: Top 10 features for computer science (70s/80s)

feature	examples	Sent. b.	subtype	S_sem	weight
Sb_obviousness	*Clearly*, *Obviously*	x	Sb_m	epistemic	-0.92
Sb_self-mention	*We*, *Our*	x	Sb_type		-0.88
Sb_Main		x	Sb_doc		-0.76
Sb_att-marker	*Specifically*, *Unfortunately*	x	Sb_type	attitudinal	-0.71
S_JJS	*best*, *simplest*		S_pos		-0.52
S_progress	*new*, *novel*		S_m	attitudinal	-0.51
S_complexity	*easy*, *difficult*		S_m	attitudinal	-0.51
Pb_complexity	*It is easy to see that*	x	Pb_m	attitudinal	-0.50
S_sufficiency	*this is not sufficient on its own*		S_m	attitudinal	-0.44
P_to_eval-adv_v_target-np	*to correctly capture the characteristics*		P_type		-0.42

Table 5.8: Top 10 features for computer science (2000s)

of obviousness becomes distinctive only in the 2000s, (2) self-mention gains distinctness over time, and (3) features at sentence beginning *(Sent. b.)* become more prominently distinctive over time, confirming the general diachronic trend observed in Section 5.2.2.1. To inspect this more closely, we investigate each feature within concordance lines and extract relevant distributional information to obtain a better understanding of the evaluative characteristics of computer science.

Attitudinal meanings While in the 70s/80s, the stance meaning of complexity (*S_complexity*) is the most distinctive feature for A, and thus the most characteristic feature for computer science (rank 1), in the 2000s its distinctness decreases (rank 7). However, we can see from Table 5.8 that the usage of the complexity meaning within patterns at sentence beginning (*Pb_complexity*) becomes distinctive in the

2000s. This points toward a more conventionalized usage of this meaning by writers. By inspecting the patterns that express the complexity meaning in the 2000s, the it-pattern (*it_rel-v_eval-adj_target-clause*) is the most often used one, i.e. the usage of the complexity meaning seems to become relatively conventionalized over time for computer science as it changes from being distinctive in general to being distinctive when used within the it-pattern at sentence beginning (see Example (1)). Moreover, in combination with the complexity meaning, the pattern always evaluates a to-infinitive which mostly consists of *to* plus the verb *see* (in 54.7% of the cases). In this way, writers convey to the reader that what they intend to show is easy to perceive/see. Considering the lemmas and possible negation, in both time periods, approx. 78% of the complexity meaning is used to evaluate things as *easy/simple/not difficult* and approx. 22% is used to evaluate things as *difficult/hard/not easy*. Note that the complexity meaning is reflected in more than one feature in the 2000s, if we consider that it is expressed on the third rank in the *Sb_att-marker* feature. Thus, even though the complexity meaning itself (*S_complexity*) seems to drop in terms of rank, it gains characteristic usage over time, being reflected in several features.

(1) [$_{\textbf{it-pattern}}$ *It is* [$_{\textbf{complexity-adjective}}$ *easy*] *to see that exactly one of those packets arrives at v at each time step*]. (A: Computer science 2000s)

The progress meaning (*S_progress*) has also decreased in distinctness diachronically for A (compare rank 2 in the 70s/80s and rank 6 in the 2000s). The target which is mostly evaluated with this meaning in both time periods is *algorithm* (see Examples (2) and (3)). Interestingly, *algorithm* is also a distinctive feature for the 70s/80s as a target (*T1_algorithm*), mostly evaluated as being *new* or *best*, i.e. with the progress or desirability meaning.

(2) *Our* ***new algorithm*** *can recover the optimal WLCS* [...].
(A: Computer science 70s/80s)

(3) *We give a* ***new algorithm*** *which computes all the derivatives of a polynomial in 3m.* (A: Computer science 2000s)

The meaning that has most prominently gained distinctness, i.e. which has become quite characteristic for computer science in the 2000s, is the obviousness meaning at sentence beginning (*Sb_obviousness* at rank 1). In computer science this meaning is mostly expressed by the evaluative adverb *clearly* with approx. 53.1% and *obviously* with approx. 19.2% (see Example (4)). Note that this feature was not among the top 10 in the 70s/80s. By using the obviousness meaning at sentence beginning, writers present the clause following the adverb in such a way that it is somehow perceived as a factual statement. Here, the question arises whether these adverbs might also adopt a textual function besides their interpersonal one. Moreover, as the usage of the obviousness meaning seems to be relatively conventionalized, the question arises whether the evaluation strength is somehow weaker than for other expressions of obviousness (see Example (5)).

(4) ***Clearly***, *every term can be covered by at most one block from B.* (A: Computer science 2000s)

(5) *This algorithm* ***obviously*** *solves the problem* [...]. (A: Computer science 2000s)

In addition, while in the 70s/80s, the meaning of idiosyncrasy is distinctively used at sentence beginning (*Pb_idiosyncrasy*) within the adverbial pattern, in the 2000s, attitude markers in general are used at sentence beginning (*Sb_att-marker*), realized by different meanings, of which idiosyncrasy is most often used. Thus, there seems to be a tendency toward expressing more attitudinal meanings distinctively at the beginning of a sentence over time.

A further meaning that becomes distinctive in the 2000s, even though it is positioned at rank 9, is the sufficiency meaning (*S_sufficiency*). This meaning is mostly expressed within the it-pattern as shown in Example (6), similarly to the complexity meaning. What is sufficient is mostly related to material and mental processes with the verbs *show* and *prove*, respectively.

(6) *It is* ***sufficient*** *to show that for each C in CG(NG[x]), C is in CG(NG[x]) and NG(C)=NG(C).* (A: Computer science 2000s)

Self-mention Diachronically, self-mention at sentence beginning (*Sb_self-mention*) has also become quite characteristic for computer science, moving from rank 10 to rank 2. In both time periods, the personal pronoun *we* prevails (93% for the 70s/80s and 91% for the 2000s) followed mostly by material processes (e.g., *show*, *use*). The personal pronoun *our* is the second most often used pronoun (with 6.6% in the 70s/80s and 9.1% in the 2000s). Considering the noun lemmas used with *our*, *result* prevails in both time periods (around 15% in both), while *algorithm* is used quite often in the 2000s (1.9% in the 70s/80s, 14.3% in the 2000s).

Document section Considering distinctive features associated with document sections, we can observe that while in the 70s/80s patterns in the Abstract section (*P_Abstract*) and stance expressions in the Introduction section (*S_Introduction*) are quite distinctive, in the 2000s, stance expressions at sentence beginning in the Main part of research articles become distinctive (*Sb_Main*). This is clearly related, first, to the obviousness meaning which is expressed at sentence beginning by evaluative adverbs, and second, to self-mention at sentence beginning, both most often used within the Main part of research articles. Thus, writers of computer science articles intrude the discourse mostly within the Main parts of research articles in combination with material processes, i.e. by presenting what they do (e.g., *we use/show*).

Additionally, with the obviousness meaning at sentence beginning, they present information in the Main part in such a way that it is perceived as a factual statement.

Patterns In the 70s/80s, the quite general pattern *eval-adj_target-np* is distinctively used at sentence beginning, in which the meaning of importance is most often used (see Example (7)), while in the 2000s the to-infinitive adverbial pattern (*P_to_eval-adv_v_target-np*) is distinctively used by computer science. In this case, the writers put forward (1) what has to be done, evaluating the process mostly with the adverb *effectively* (see Example (8)), or (2) what has been done, using the adverbs *accurately* or *correctly* (see Example (9)).

(7) *The* [$_{\textbf{eval-adj}}$ *important*] [$_{\textbf{target-np}}$ *JKY theorems*] *that will be used frequently in what follows are given below.* (A: Computer science 70s/80s)

(8) *Note that in order* [$_{\textbf{to-inf}}$ *to*] [$_{\textbf{eval-adv}}$ *effectively*] [$_{\textbf{verb}}$ *apply*] [$_{\textbf{target}}$ *this idea in an information-theoretic setting*]*, one must ensure that* [...]. (A: Computer science 2000s)

(9) *Both models prove* [$_{\textbf{to-inf}}$ *to*] [$_{\textbf{eval-adv}}$ *correctly*] [$_{\textbf{verb}}$ *capture*] [$_{\textbf{target-np}}$ *the characteristics of asymmetric TCP*]. (A: Computer science 2000s)

In summary, computer science is characterized by the use of the complexity meaning, which seems to have become more conventionalized in the 2000s, as it is used in patterns at sentence beginning conveying a sense of ease about what has to be perceived by the reader. This conventionalization trend is also reflected by the distinctive use of the obviousness meaning and self-mention at sentence beginning in the 2000s. Thus, in general, there seems to be a trend toward putting evaluative meaning first in a rather conventionalized way. In addition, we can observe a slight shift from attitudinal (complexity, progress) to epistemic meaning (obviousness) which is most distinctively used in the 2000s.

5.2.2.2.2 Linguistics For linguistics, the most characteristic features are shown in Tables 5.9 and 5.10. Diachronically, the following observations can be made: (1) the amount of stance meaning features decreases over time, while the amount of stance types increases — for example, specific parts-of-speech of stance expressions become distinctive (*S_NNS*, *S_VVZ*, *Sb_RB*) rather than specific meanings, (2) concrete epistemic meanings (e.g., reasoning, suggestion) become less characteristic, while the use of hedges in general at sentence beginning is very distinctive in the 2000s, and (3) similarly to computer science, there seems to be a tendency to express specific kinds of evaluative meaning at sentence beginning (5 out of the top 10 features). In the following, we investigate these tendencies in more detail.

feature	examples	Sent. b.	subtype	S_sem	weight
S_dispute	*problem*, *issue*		S_m	attitudinal	3.71
S_Abstract			S_doc		1.22
S_idiosyncrasy	*especially*, *specifically*		S_m	attitudinal	1.12
S_self-mention	*we*, *I*		S_type		0.90
S_assumption	*seem*, *appear*		S_m	epistemic	0.88
S_VVZ	*seems, appears*		S_pos		0.81
P_likelihood	*it is extremely likely that*		P_m	epistemic (hedge)	0.72
S_acceptance	*acceptable*, *unacceptable*		S_m	attitudinal	0.66
S_reasoning	*argue*, *imply*		S_m	epistemic (hedge)	0.61
S_suggestion	*should*, *suggest*		S_m	epistemic (hedge)	0.60

Table 5.9: Top 10 features for linguistics (70s/80s)

feature	examples	Sent. b.	subtype	S_sem	weight
Sb_hedge	*perhaps*, *assume*	x	Sb_type	epistemic (hedge)	1.37
S_dispute	*problem*, *issue*		S_m	attitudinal	1.00
S_NNS	*problems*, *assumptions*		S_pos		0.64
S_VVZ	*seems*, *suggests*		S_pos		0.61
Sb_importance	*Crucially*, *Importantly*	x	Sb_m	attitudinal	0.61
Sb_attitudinal	*specifically*, *interestingly*	x	Sb_sem	attitudinal	0.58
P_target_hedge-v	*I suggest that*		P_type	epistemic (hedge)	0.57
Sb_interest	*Interestingly*	x	Sb_m	attitudinal	0.51
P_obviousness	*It is quite clear that*		P_m	epistemic (booster)	0.49
Sb_RB	*specifically*, *perhaps*	x	Sb_pos		0.39

Table 5.10: Top 10 features for linguistics (2000s)

Stance meaning In both time periods, linguistics is characterized by various attitudinal and epistemic stance meanings. In terms of attitudinal meanings, dispute (*S_dispute*) is very characteristic for linguistics and preserved over time, as it is also very distinctive in the 2000s (see Example (10)). However, while in the 70s/80s the use of the noun *problem* prevailed with 77.1%, followed by *issue* with 16.8%, in the 2000s the use of *issue* increases up to 37.8%, while the use of *problem* decreases to 52.4%. Additionally, the use of adjectives such as *problematic* also increases over time (from 3.3% to 8.0%). Thus, while the meaning of dispute is preserved over time, the distribution of the lexical items used varies across time periods.

(10) *Patterson and colleagues (2001), however, reported that patients with varying degrees of severity of semantic dementia had selective* ***problems*** *with*

irregular past-tense formation. (C1: Linguistics 2000s)

In the 2000s, the attitudinal meanings of importance and interest at sentence beginning are distinctive (*Sb_importance* and *Sb_interest*). Importance is here mostly realized by the adverbs *crucially* with 33.3% (see Example (11)) and *importantly* with 21.7% (see Example (12)), while interest is mostly expressed by *interestingly* with 95.9% (see Example (13)). Moreover, while in the 70s/80s the attitudinal meaning of idiosyncrasy is distinctive (*S_idiosyncrasy*; see Example (14)), in the 2000s, attitudinal expressions in general are distinctive at sentence beginning (*Sb_attitudinal*). Note, however, that the most frequently used meaning of attitudinal expression at sentence beginning in the 2000s is the idiosyncrasy meaning.

(11) ***Crucially****, the temporal relation between the embedded clause in 32a and the moment of speech is completely unspecified.* (C1: Linguistics 2000s)

(12) ***Importantly****, this finding does not rest on the frequency of combinations of particular words, but rather on an abstraction to classes of phonological contexts.* (C1: Linguistics 2000s)

(13) ***Interestingly****, these findings are qualitatively similar to results of neologism experiments on purely phonological generalizations* [...]. (C1: Linguistics 2000s)

(14) *Manner adverbs which refer* ***specifically*** *to the energy output of an action* [...] *discriminate quite well between 'effective' and 'agentive'* [...]. (C1: Linguistics 70s/80s)

In terms of epistemic hedge meanings, we can observe that in the 70s/80s single meanings are distinctive: (1) assumption, which is mostly realized by verbs in the 3rd person singular, with the part-of-speech VVZ, which is also a distinctive feature (*S_VVZ*), (2) reasoning, which is mostly realized by adjectives (e.g., *reasonable*, *plausible*) and VVZ verbs (e.g., *argues*, *implies*), and (3) suggestion, which is mostly realized by modal verbs (e.g., *should*, *shall*). Additionally, the epistemic meaning of likelihood is distinctively used within patterns (*P_likelihood*). Mostly, this meaning is expressed either in the adjectival pattern (*eval-adj_target-np*) or in the it-pattern (*it_rel-v_eval-adj_target-clause*) (see Table 5.11). The tendency of individual meanings being characteristic for linguistics changes in the 2000s, as single epistemic meanings are not distinctive anymore. What becomes very distinctive in the 2000s is the general use of hedges at the beginning of a sentence (*Sb_hedges*; see Example (15)) rather than individual epistemic meanings, except for the obviousness meaning within patterns (*P_obviousness*), which is also mostly used within the adjectival or it-pattern (see Table 5.12).

Additionally, in the 2000s, the hedged character of linguistics is also seen when inspecting the distinctive features *S_NNS* and *S_VVZ*, which relate to plural common

pattern	example	raw freq.	%
eval-adj_target-n	*possible alternatives to the precyclic hypothesis*	1053	68.78
it_rel-v_eval-adj_target-clause	*it is possible to form*	344	22.47
target-np_rel-v_eval-expr	*this second interpretation is possible*	124	8.10
eval-adv_target-clause	*Probably* [...]	10	0.65

Table 5.11: Likelihood expressed within patterns for linguistics (70s/80s)

pattern	example	raw freq.	%
eval-adj_target-n	*a clear prediction*	281	59.91
it_rel-v_eval-adj_target-clause	*It is clear that*	115	24.52
eval-adv_target-clause	*Clearly, further research is needed*	50	10.66
target-np_rel-v_eval-expr	*The most obvious difference is that*	23	4.90

Table 5.12: Obviousness expressed within patterns for linguistics (2000s)

nouns and 3rd person singular verbs, respectively. Of the plural common nouns, besides dispute, the most often used meaning is assumption (see Example (16)). Of the 3rd person singular verbs, assumption is the most frequently expressed meaning with 37.3% (see Example (17)). Thus, these two parts-of-speech also reflect the hedged character of linguistics by using most prominently epistemic hedge meanings. Furthermore, the *target_hedge-v* pattern, distinctive for the 2000s, also points to the hedged character of linguistics (see Example (18)).

(15) ***Perhaps*** *few of them are amenable to notional variation, either because of their semantics or their typical contexts of occurrence.* (C1: Linguistics 2000s)

(16) *I will merely summarize some* ***speculations*** *that have been advanced to answer these questions and will not attempt to decide among them.* (C1: Linguistics 2000s)

(17) *His scale* ***seems*** *to be one of relative referential specificity which accords quite well with the representations I suggest here.* (C1: Linguistics 2000s)

(18) *But* [$_{\textbf{target-np}}$ *I*] [$_{\textbf{hedge-v}}$ *suggest*] *that an appositional account of nomination structures does not violate a basic understanding of what is involved in apposition.* (C: Linguistics 2000s)

pronoun	**example**	**raw freq.**	**%**
we	*We may begin by looking at*	9630	56.66
I	*The analysis I am proposing*	1638	9.64
me	*It seems to me that*	1552	9.13
our	*Our claim is that*	1528	8.99
my	*My intention is partly*	1226	7.21
us	*enable us to distinguish*	1190	7.00
myself	*I will confine myself to*	172	1.01
ourselves	*We restrict ourselves here to*	42	0.25
ours	*are different from ours*	18	0.11

Table 5.13: Personal pronouns expressing self-mention for linguistics (70s/80s)

Sentence beginning Diachronically, we can observe a tendency toward expressing epistemic as well as attitudinal meanings at sentence beginning. While in the 70s/80s, epistemic meanings (such as assumption or suggestion) were distinctive in general (i.e. not at sentence beginning), in the 2000s hedges are distinctive when used at sentence beginning (*Sb_hedge*). Similarly, the attitudinal meanings of importance (*Sb_importance*) and interest (*Sb_interest*) but also attitudinal meanings in general (*Sb_attitudinal*) at sentence beginning are characteristic for linguistics in the 2000s. Thus, writers in linguistics have moved toward expressing epistemic and attitudinal expressions in a more conventionalized way at sentence beginning and within patterns, the latter also reflected in the distinctive use of the obviousness meaning expressed in patterns (*P_obviousness*) and the distinctive use of the *target_hedge-v* pattern (see, again, Table 5.10 and Example (18)).

Self-mention While self-mention is quite distinctive in the 70s/80s, this is not the case anymore in the 2000s. Still, it is quite interesting to see how self-mention is expressed in the 70s/80s, as it shows quite some variability in the choice of personal pronouns used (see Table 5.13). Besides the plural forms, the singular form is relatively frequent in linguistics, related to a higher proportion of research articles written by single authors. Comparing this to computer science, where around 90% of self-mention is expressed by the personal pronoun *we* for the 70s/80s, we can see how this is clearly different in linguistics in that time period, as *we* is used only approx. 56% of the time.

In summary, linguistics shows a quite strong hedged character, which is realized by specific epistemic meanings in the 70s/80s (e.g., assumption). In the 2000s, instead, it is realized more generally by hedged expressions not confined to a specific meaning and nouns and verbs realizing hedges. Additionally, the dispute meaning is relatively distinctive for both time periods, where things are put forward as being problematic. Moreover, the general diachronic trend of expressing evaluative meaning at sentence beginning is also relatively pronounced in linguistics.

feature	examples	Sent. b.	subtype	S_sem	weight
S_RB	*partially*, *rather*		S_pos		2.35
S_suggestion	*suggest*, *propose*		S_m	epistemic	1.61
Pb_presumption	*Apparently*, *Presumably*	x	Pb_m	attitudinal	1.05
S_booster	*demonstrate*, *reveal*		S_type	epistemic	1.05
Pb_to_eval-v_target-np	*To avoid the experimental complication*	x	Pb_type		1.03
P_Main			P_doc		1.02
S_self-mention	*we*, *our*		S_type		1.01
S_likelihood	*may*, *possible*		S_m	epistemic	1.01
P_it_hedge-v_eval-adj_target-clause	*it seems very likely that*		P_type		0.98
Pb_target-np_rel-v_eval-expr	*We were unable to obtain*	x	Pb_type		0.98

Table 5.14: Top 10 features for biology (70s/80s)

5.2.2.2.3 Biology Considering the top 10 features for biology for both time periods, which are shown in Tables 5.14 and 5.15, some main observations can be made which apply to both time periods, even though different features are involved: (1) for both time periods the feature types are quite varied, especially if we compare it to the other registers we have already inspected (which have a preference for stance meanings), (2) pattern features are quite frequent in the top 10 features for both time periods, and (3) if stance meanings are distinctive, they are mostly of the epistemic type, and if they are of the attitudinal type, they are distinctively used in patterns; again, this applies for both time periods.

Epistemic stance meaning Considering the stance meanings distinctive for both time periods, there seems to be a preference for epistemic meanings. In the 70s/80s, the suggestion meaning (*S_suggestion*) is quite distinctive (rank 2). The verb *suggest* is the most frequently used lexical item that expresses this meaning with 47.2%, followed by the modal verb *should* with 38.3% and the verb *propose* with 10.6% (see Table 5.16). Considering what precedes the verb *suggest*, we see that what suggests are mainly experimental concepts realized by nouns such as *results*, *evidence* and *data*. This reflects the interpretative endeavor of researchers in biology on experimental findings. The likelihood meaning is also distinctive for biology in the 70s/80s, being mostly realized by the use of modal verbs with approx. 68.9% (e.g., *can*, *may*, *would*). From Example (19), we can also deduce the interpretative effort on experimental evidence by biologists within their research articles (consider the whole phrase *the available evidence indicates that there might be*). In addition, boosters are distinctive in the 70s/80s, mostly realized by verbal constructions such

feature	examples	Sent. b.	subtype	S_sem	weight
Sb_epistemic	*Perhaps, Apparently*	x	Sb_sem	epistemic	1.11
S_VVD	*revealed, suggested*		S_pos		1.10
S_PP	*we, our*		S_pos		1.06
Sb_relativity	*Roughly, Nearly*	x	Sb_m	epistemic	1.05
P_capability	*the Pole3 E box is capable to*		P_m	attitudinal	0.97
P_eval-adv_target-clause	*Interestingly, Indeed*		P_type		0.84
S_Abstract			S_doc		0.79
S_self-mention	*we, our*		S_type		0.77
P_Main			P_doc		0.76
P_suitability	*the developed microarray is suitable*		P_m	attitudinal	0.69

Table 5.15: Top 10 features for biology (2000s)

lexical item	example	raw freq.	%
suggest	*These results suggest that*	1228	47.25
should	*both properties should revert simultaneously*	996	38.32
propose	*Firstly, we propose a mechanism*	276	10.62
suggestion	*Based on a suggestion by*	74	2.85
shall	*These fragments shall be referred to as*	21	0.81
ought	*therefore, ought to be heterogeneous*	4	0.15

Table 5.16: Lexical items expressing suggestion for biology (2000s)

as *demonstrate that* or nouns such as *evidence*.

Moreover, quite distinctive in the 70s/80s are stance adverbs (*S_RB*). By inspecting the meanings they realize, epistemic meanings prevail, with the most often used meanings being relativity (approx. 31.4%) and likelihood (15.7%) (see Examples (20) and (21), respectively).

(19) *The available evidence indicates that there* ***might*** *be differences in the biosynthetic response of the two genes to amino acid starvation* [...].
(C2: Biology 70s/80s)

(20) *Cleavage at the resistant sites discussed here can be* ***partially*** *effected by long incubation with excess enzyme.* (C2: Biology 70s/80s)

(21) *This variation may reflect real structural differences found in mouse ribosomal genes or* ***possibly*** *deletion events which occurred during cloning.*
(C2: Biology 70s/80s)

In the 2000s, the stance meaning of relativity at sentence beginning (*Sb_relativity*) is distinctive, again an epistemic meaning which is used to report approximations with respect to experimental work (see Example (22)). Besides the relativity meaning, biology uses epistemic expression in general at sentence beginning (*Sb_epistemic*) quite distinctively in the 2000s. Most prominently, the certainty meaning is used. Considering the expressions which realize this meaning, the phrase *consistent with* is used quite often with 74.8% (see Example (23)). In these cases, the writers make clear that their experimental results are consistent with other results, expressing some degree of certainty. Note that besides this quite conventionalized expression, adverbs are also used (as in Example (24)). Thus, writers in biology use, on the one hand, hedging expressions with the relativity meaning distinctively, and on the other hand, boosting expressions with the certainty meaning, both at sentence initial position, which allows the reader to immediately perceive how the utterance should be understood in epistemic terms.

(22) ***Roughly*** *the same number of unique transcripts appeared in the 0-h* [...]. (C2: Biology 2000s)

(23) ***Consistent with*** *our previous study, NCF2 intron 1 5'-UTR variants were expressed more abundantly compared to the exon 1 variant in each of the human tissues tested and in nearly all cell lines* [...]. (C2: Biology 2000s)

(24) ***Certainly****, there is evidence that Sp1 has multiple transcripts and uses alternate splicing and even trans-splicing to increase the diversity of its transcripts* [...]. (C2: Biology 2000s)

In addition, the part-of-speech of verbs in the past tense (*S_VVD*) also reflects the usage of epistemic meaning, as the most frequently used stance meaning is evidence with 66.9% (e.g., *showed that*, *revealed*), followed by suggestion with 9.35% (e.g., *suggested*, *proposed*). For evidence, either the researchers themselves (with the pronoun *we*) or experimental, analytical or biological concepts (e.g., *analysis*, *study*, *result*, *experiment*, *gene*, *protein*, *cell*) give evidence by expressions such as *showed that*, *revealed*, *demonstrated that*, *supported* and *proved*. For the suggestion meaning, experimental, analytical or biological concepts suggest a possible interpretation of specific experimental findings (see Example (25)).

(25) *These results* ***suggested*** *the possibility that the cytoplasmic caspase 8* [...] *was absent or defective in mtDNA-depleted cells.* (C2: Biology 2000s)

Self-mention Characteristic of biology in both time periods is also self-mention. There is no variation in terms of types of pronouns used, i.e. the same pronouns are used across time with slight distributional differences. The pronouns *we* and *our* prevail with approx. 70% for *we* and 20% for *our* in both time periods, similarly to the other registers we have inspected so far. However, considering the verbs used with the pronoun *we*, mostly mental processes are used such as *find*, *conclude*, or *observe*, which differs from computer science and linguistics. Considering nouns used with *our*, experimental concepts are used, such as *result*, *study*, *data*, and *experiment*. Thus, self-mention is distinctive for biology in both time periods in similar terms.

Patterns In terms of meanings expressed within patterns, most distinctive for the 70s/80s is the attitudinal meaning of presumption used at sentence beginning (*Sb_presumption*). This meaning is mainly realized by the adverbial pattern, as shown in Example (26), by the lexical items *apparently* and *presumably*.

(26) ***Apparently***, *the gene responsible for the modification of the drug has its own promoter for expression when integrated as a lysogen.*
(C2: Biology 70s/80s)

In the 2000s, instead, the meaning of capability is distinctive when used within patterns (*P_capability*) with the *target-np_rel-v_eval-expr* pattern being the most frequently used pattern to express this meaning. Note that this pattern is also distinctive (see, again, Table 5.15). Thus, the target is either able or unable to do something such as in Example (27). Targets in this case are mostly concepts related to biology (such as *cells*, *DNA*, *RNA*, *enzyme*, ect.) or the pronoun *we*, i.e. expressing whether the researchers themselves were able or unable to do/obtain something. In addition, the meaning of suitability within patterns is distinctive for the 2000s, which is mostly realized within the *eval-adj_target-np* pattern as shown in Example (28) with the adjectives *appropriate* (65.8%), *proper* (26.7%), and *suitable* (7.5%).

(27) [...] [$_{\textbf{target-np}}$ *mtDNA-depleted C2C12 cells*] [$_{\textbf{rel-v}}$ *are*] [$_{\textbf{eval-expr}}$ *unable to execute the apoptotic process*]. (C2: Biology 2000s)

(28) *Recombinant clones were sequenced in order to verify that the plasmid contained two T7 polymerase promoters in opposite orientation separated by the* [$_{\textbf{eval-adj}}$ *appropriate*] [$_{\textbf{target-np}}$ *chitin synthase gene fragment*].
(C2: Biology 2000s)

Besides distinctive meanings used within patterns, different pattern types are distinctive in both time periods. Diachronically, we can see that more patterns are distinctive in the 70s/80s vs. the 2000s. For the 70s/80s, three patterns are distinctive: the it-pattern with a hedge verb (*P_it_hedge-v_eval-adj_target-clause*) which is used to express some kind of caution associated with the utterance (see Example (29)), and two patterns at sentence initial position — the *to_eval-v_target-np* pattern used for avoidance (by verbs such as *avoid*, *overcome*, *eliminate*, *prevent*; see Example (30)) and improvement (by verbs such as *facilitate*, *resolve*, *improve*, *achieve*; see Example (31)), and the *target-np_rel-v_eval-expr* pattern, which is used to evaluate mostly with capability (see Example (32)) and importance (see Example (33)).

meaning	example	raw freq.	%
interest	*Interestingly, most groups contained cDNAs encoding*	308	41.79
importance	*Notably, the promoter activity*	102	13.84
prediction	*Surprisingly, none of these protein*	41	5.56
desirability	*Unfortunately, we have not been able to*	31	4.21
idiosyncrasy	*More specifically, this model system*	27	3.66

Table 5.17: Top 5 meanings used in the *eval-adv_target-clause* pattern for biology (2000s)

(29) *Before discussing the regulatory authority of these agencies, however,* ***it seems appropriate to first review NIH's authority*** *to issue rules governing recombinant DNA research.* (C2: Biology 70s/80s)

(30) ***To overcome these problems*** *we must increasingly use* [...]. (C2: Biology 70s/80s)

(31) ***To improve the agreement*** *it is necessary to use* [...]. (C2: Biology 70s/80s)

(32) ***This virus is able*** *to replicate in permissive cells* [...]. (C2: Biology 70s/80s)

(33) ***These residues are important*** *in forming the heme contacts as well as a-3-cooperative dimer associations.* (C2: Biology 70s/80s)

For the 2000s, only the adverbial pattern (*P_eval-adv_target-clause*) is distinctive. Considering the evaluative meanings expressed with this pattern, the interest meaning is by far the most frequently used one with 41.8%, followed by importance with 13.8% (see Table 5.17). By the examples shown in Table 5.17, we can see how the evaluative expressions are targeted toward experimental work and findings, which can be positively (e.g., *interestingly*) or negatively (e.g., *unfortunate*) evaluated.

Document section In terms of document sections, patterns used within the Main part of research articles are characteristic of biology in both time periods. Interestingly, in both periods, the importance meaning is most prominently expressed by the *eval-adj_target-np* pattern (e.g., *important molecules*) with 76.9% in the 70s/80s and 71.2% in the 2000s. Note that the pattern types used within the Main part expressing the importance meaning have increased from 10 types in the 70s/80s to 14 types in the 2000s, thus showing a development toward more variation to express this meaning in terms of patterns. The targets most prominently evaluated relate mostly to biological concepts (e.g., *factor*, *protein*, *gene*). Additionally, the stance expressions within the Abstract section are a distinctive feature for biology in the

meaning	**examples**	**raw freq.**	**%**
evidence	*revealed*, *evidence*	308	41.79
likelihood	*may*, *likely*	102	13.84
importance	*important*, *significant*	41	5.56
suggestion	*suggest*, *propose*	31	4.21
desirability	*good*, *preferentially*	27	3.66

Table 5.18: Top 5 meanings of stance expressions used in the Abstract section for biology (2000s)

2000s (*S_Abstract*). Considering the stance meanings expressed, the evidence meaning is most often used with 16.6%, along with the likelihood meaning with 16.2% (see Table 5.18). Thus, while in the Abstract section evidence is put forward, in the Main section biological concepts are described as important.

In summary, we can say that biology is characterized by using epistemic meaning distinctively and especially within patterns. Most prominently, experimental findings are either evaluated or provide evidence. While there are some changes regarding the distinctive features, the general tendencies mostly apply to both time periods, which might indicate that biology did not undergo big changes in terms of developing a distinct usage of evaluative meaning.

5.2.2.3 Characteristic features of evaluative meaning for contact registers

Due to the low classification performance of the contact registers shown in Section 5.2.1, a diachronic perspective on register diversification is not suitable for the contact registers, as the top ranking features cannot be said to be characteristic. Thus, we adopt a synchronic perspective considering by which features contact registers are characterized in the 2000s. We consider each contact register separately. The aim is to investigate possible register profiles of evaluative meaning for the contact registers by the inspection of the top 10 ranking features as for the seed registers described above.

5.2.2.3.1 B1: Computational linguistics Considering the top 10 features for computational linguistics (B1) in the 2000s (see Table 5.19), most features are related to stance, one is a pattern and two others are target features. Four of the ten features are used at sentence beginning (*Sent. b.*). Note that while most of the features are related to attitudinal meaning, self-mention is also quite distinctive. Thus, in general terms, we can say that computational linguistics is characterized by the use of attitudinal meanings and the expression of self-mention.

feature	**examples**	**Sent. b.**	**subtype**	**S_sem**	**weight**
Sb_PP	*we*, *I*	x	S_pos		1.85
Sb_idiosyncrasy	*especially*, *specifically*	x	S_m	attitudinal	1.02
T1_translation	*good translation*		T_lemma		0.61
S_self-mention	*we*, *our*		S_type		0.60
Sb_attitudinal		x	S_sem		0.51
S_trust	*reliable*, *reliably*		S_m	attitudinal	0.43
P_dispute	*issue*, *problem*		P_m	attitudinal	0.43
S_persuasion	*convincingly*, *tempting*		S_m	attitudinal	0.42
Sb_Main		x	S_doc		0.42
S_progress	*new*, *novel*		S_m	attitudinal	0.41

Table 5.19: Top 10 features for computational linguistics (2000s)

verb	**example**	**raw freq.**	**%**
use	*We always use f and g to represent*	194	6.30
show	*We can show that*	82	2.66
find	*We find that*	59	1.91
see	*We see a similar pattern*	57	1.85
assume	*We assume that*	52	1.60

Table 5.20: Top 5 verbs used with *we* and *I* at sentence beginning (*Sb_PP*) for computational linguistics (2000s)

Self-mention The most distinctive feature is *Sb_PP*, which relates to personal pronouns of self-mention used at the beginning of a sentence. The pronouns used are *we* with 88.63% followed by *I* with 11.37%. Again, we inspect which verbs the pronouns occur with, to see which verbal process the writers of research articles fulfill when self-mentioning themselves at the beginning of the sentence in computational linguistics (see Table 5.20). We can see that most processes by the top 5 verbs used are either material (*use*, *show*) or mental (*find*, *see*, *assume*). Additionally, self-mention in general (*S_self-mention*) is distinctive, with the personal pronouns *we* used most often with 71.5%, followed by *our* with 21.4% (see Table 5.21). If we again consider the verbs following the personal pronouns *we* and *I*, we can see that these are almost the same as the ones used with personal pronouns at sentence beginning, besides the verb *describe*, which is mostly used after a textual Theme (e.g., *In this section*), as in the example in Table 5.22. Additionally, we also inspect the nouns following the pronoun *our*. As we can see from Table 5.23, *system* is the most often used noun, followed by *model*. Note that while *system*, *model* and *experiment* relate to concrete things within computational linguistics, *approach* and *method* are quite generally used in the scientific domain.

Attitudinal meaning Regarding attitudinal meaning, it is in general distinctive at the beginning of a sentence (*Sb_attitudinal*). In these cases, the meanings of

pronoun	**example**	**raw freq.**	**%**
we	*we need only*	11688	71.47
our	*our experiments are*	3504	21.43
us	*This lead us to*	613	3.75
I	*I thought it was*	225	1.38
my	*thanks to my Ph.D. advisor*	135	0.83
me	*Let me argue against*	114	0.70
ours	*similar to ours*	40	0.24
ourselves	*we asked ourselves*	29	0.18
myself	*asking myself*	4	0.02
mine	*a friend of mine*	1	0.01

Table 5.21: Personal pronouns expressing self-mention for computational linguistics (2000s)

verb	**example**	**raw freq.**	**%**
use	*In [...] experiment, I used the most general*	737	6.19
describe	*In this section we first describe*	244	2.05
see	*in which case we may see the error*	223	1.87
find	*alignment Recall, as we find no verb*	220	1.85
show	*this requirement, we show how appropriate*	212	1.78

Table 5.22: Top 5 verbs used with *we* and *I* (*S_self-mention*) for computational linguistics (2000s)

noun	**example**	**raw freq.**	**%**
system	*In other words, our hybrid MT system*	292	8.33
model	*better performance of our translation models*	234	6.68
approach	*we tested our approach with*	215	6.14
experiment	*the case for our experiments*	174	4.97
method	*we suggest that our proposed method*	156	4.45

Table 5.23: Top 5 nouns used with *our* (*S_self-mention*) for computational linguistics (2000s)

desirability (25.6%), importance (15.6%) and idiosyncrasy (10.4%) are most often expressed. For the desirability meaning, *unfortunately* is the most often occurring lexical item with 44.7% (see Example (34)).

(34) ***Unfortunately****, we cannot say whether the significant improvement in effectiveness occurs mainly because the probability of giving at least one good translation* [...] *is higher for QT or indeed because of the query expansion effect.*

Considering the single attitudinal meanings, idiosyncrasy at the beginning of a sentence (*Sb_idiosyncrasy*) is one of the most distinctive features for computational linguistics. This meaning is expressed by the adverbs *specifically* with 81.6%, *especially* with 16.3%, and *exceptionally* with only 2.0%. In these cases, the adverb is positioned at sentence beginning preceding a clause. The attention of the reader is pointed toward the clause following the adverb (see Example (35)).

(35) ***Especially****, we gave the experimental results comparing concept-code and word features which have not been reported before and we proved the superiority of concept-code features to word features in disambiguation performance.*

The attitudinal meaning of trust (*S_trust*) is also distinctive and mostly realized by the lexical items *reliable* (with 57.4%), used e.g. to evaluate the nouns *results* and *translation* (see Examples (36) and (37)), and *reliably* (with 25.2%) which evaluates verbal processes such as *correlate*, *perform* or *identify* (see Examples (38)–(40)).

(36) *Consequently, the need for adaptation is reduced to a minimum and the user is supplied with fast and* ***reliable translations****.*

(37) *This result indicates that rules with a higher priority use detailed structural information as a condition and this confirms their* ***reliable results****.*

(38) *We evaluate our results against paraphrase judgments elicited experimentally from* [...] *and show that the model's ranking of meanings* ***correlates reliably*** *with human intuitions.*

(39) *Furthermore, our model* ***performs reliably*** *better than a naive baseline model* [...].

(40) *We describe our model in detail and present experimental results that show that our model is able to learn to* ***reliably identify*** *word- and phrase-level alignments* [...].

As can be seen from the examples, the trust/reliability meaning is in a general sense related to what computational linguists produce or work with (e.g., models, results, translations).

lexical item	example	raw freq.	%
tentative	*can only be viewed as a tentative approach*	14	31.82
convincing	*provides convincing evidence that*	12	27.27
credible	*should be reasonably credible*	4	9.09
convincingly	*this suggests rather convincingly that*	4	9.09
persuasive	*the improvement is not as persuasive as*	3	6.82
tentatively	*we tentatively conclude that*	3	6.82
tempting	*It is often tempting for an experienced linguist*	2	4.55
compelling	*The most compelling reasons for*	2	4.55

Table 5.24: Lexical items expressing persuasion for computational linguistics (2000s)

Also distinctive is the meaning of persuasion (*S_persuasion*), mostly realized by the lexical items *tentative* and *convincing* (see Table 5.24). Note that while *tentative* is negatively co-notated and used to evaluate nouns such as *translation*, *implementation* or *approach*, *convincing* is used positively to enforce things such as *evidence*, *proof* or the general noun *way* (as in *a very convincing way*).

The attitudinal meaning of progress (*S_progress*) is also distinctive, even though less distinctive than the other stance meanings illustrated above. The adjective *new* is by far the most often used adjective of this meaning with 96.0%, followed by *novel* (2.5), and *modern* (1.4%) (see Examples (41)–(43), respectively). Note that the targets evaluated with this meaning are associated with theoretical or technical concepts (e.g., *algorithm*, *model*, *technique*, *approach*, *concept*).

(41) *Even if the variants are deemed to depart substantially from the original algorithm, we have at least obtained a family of **new bootstrapping algorithms** that are mathematically understood.*

(42) *This paper describes a **novel approach** to morphological tagging for Korean* [...].

(43) *This is not only because some systems combine standard NLG with templates and canned text (Piwek 2003), but also because **modern template-based systems** tend to use syntactically structured templates and allow the gaps in them to be filled recursively* [...].

Additionally, the attitudinal meaning of *dispute* expressed within a evaluative attributive pattern (*P_dispute*) is distinctive for computational linguistics. Table 5.25 shows the patterns used with this meaning. The prepositional pattern (*eval-np_prep_target-np*) is the most often used pattern (71.7%), closely followed by the noun phrase pattern *eval-adj_target-np* (23.2%). The targets in these cases are either relatively abstract nouns such as *approach* or *claim*, or concrete things related to computational linguistics such as *zero anaphora*, as shown in the example in Table 5.25 for

pattern	**example**	**raw freq.**	**%**
eval-np_prep_ target-np	*The problem of this approach*	71	71.72
eval-adj_target-np	*C1 is the most problematic claim*	23	23.23
target-np_rel-v_ eval-expr	*zero anaphora are problematic*	4	4.04
target_v_eval-expr	*termination becomes problematic*	1	1.01

Table 5.25: Evaluative patterns used with the dispute meaning for computational linguistics (2000s)

meaning	**example**	**raw freq.**	**%**
accuracy	*finding the correct translation*	31	31.0
desirability	*leads to a better translation*	23	23.0
acceptability	*The only acceptable translation*	28	14.0
likelihood	*identifying possible translations*	16	12.0
suitability	*The appropriate translation is shown below*	5	4.0

Table 5.26: Top 5 meanings used to evaluate *translation* as a uni-gram for computational linguistics (2000s)

the *target-np_rel-v_eval-expr* pattern.

Besides self-mention and attitudinal meanings, computational linguistics is also distinguished by a target feature: *translation* as a uni-gram target. Inspecting more closely how the target is used, it is most often evaluated by the meanings of desirability, accuracy and likelihood (see Table 5.26). Thus, translations are mostly correct (23%), better/good/poor (16%) or possible (11%) in computational linguistics.

Finally, computational linguistics is distinct from the other registers by the use of stance expressions at the beginning of a sentence within the Main part of research articles (*Sb_Main*). Considering the distribution of the stance types (see Table 5.27), we can see that self-mention and attitude markers are the stance types that make this feature a distinctive one. Thus, the above observations of self-mention, as well as the different meanings used at the beginning of a sentence being distinctive for computational linguistics, occur most distinctively within the Main parts of research articles.

In summary, we can say that computational linguistics in the 2000s is characterized by expressing self-mention as well as idiosyncrasy at the beginning of a sentence. These two most distinctive features are related to information structure, but have different functions: self-mention to introduce what the writers do (material and mental process types) or relate their work to (systems, models, etc.), and idiosyncrasy to point the reader's attention toward what follows in terms of further specification. Moreover, the other distinctive attitudinal meanings, besides idiosyncrasy, reflect how things are evaluated within computational linguistics, especially as reliable

stance type	**raw freq.**	**%**
self-mention	2744	78.38
attitude marker	382	10.91
hedge	198	5.66
booster	177	5.06

Table 5.27: Stance types at the beginning of a sentence within the Main part of research articles for computational linguistics (2000s)

feature	**examples**	**Sent. b.**	**subtype**	**S_sem**	**weight**
S_self-mention	*we*, *I*		S_type		1.53
Sb_self-mention	*we*, *our*	x	Sb_type		1.49
Sb_booster	*obviously*, *clearly*	x	Sb_type		1.48
S_progress	*new*, *novel*		S_m	attitudinal	1.38
T1_NNS	*results*, *genes*		T_pos		1.32
Sb_importance	*importantly*, *essentially*	x	Sb_m	attitudinal	1.16
S_PP	*we*, *I*		S_pos		1.15
S_trust	*reliable*, *reliably*,		S_m	attitudinal	1.05
Sb_RB	*apparently*, *specifically*	x	Sb_pos		1.03
S_assumption	*appear*, *seem*		S_m	epistemic (hedge)	0.97

Table 5.28: Top 10 features for bioinformatics (2000s)

(e.g., translations or results), tentative or convincing (e.g., approach or evidence), and new (e.g., algorithms or systems).

5.2.2.3.2 B2: Bioinformatics The top 10 features for bioinformatics (B2) are shown in Table 5.28. In comparison to B1, for B2 even more features are related to stance, while there is only one target feature (*T1_NNS*) present in the top 10. We can say to have three main groups of features by which bioinformatics is characterized: (1) self-mention (*S_self-mention*, *Sb_self-mention* and *S_PP*), (2) booster (*Sb_booster*, *Sb_RB*), and (3) stance meaning (*S_progress*, *Sb_importance*, *S_trust*, *S_assumption*). Four of the ten features are used at sentence beginning.

Self-mention Self-mention is most distinctive for bioinformatics in general, but also at the beginning of the sentence (*S_self-mention*, *Sb_self-mention*). Considering the pronouns that realize self-mention, *we* is the most frequently occurring one with 74.2%, followed by *our* with 23.6%. Note that the pronoun *I* is quite seldom for bioinformatics (0.05%). Similarly to computational linguistics, material and mental processes are used in combination with self-mention, but the material processes prevail. Thus, writers of research articles in bioinformatics *use*, *show* and *apply* (see Table 5.30), while they also *find* and *consider*. When self-mention is used at

pronoun	example	raw freq.	%
we	*then we might expect*	8785	74.22
our	*may improve our ability*	2797	23.63
us	*approach enables us*	225	1.90
ours	*quicker than ours*	17	0.14
I	*I performed a series of*	6	0.05
ourselves	*we restricted ourselves to*	4	0.03
me	*for sending me the Human-Chimp*	2	0.02

Table 5.29: Personal pronouns expressing self-mention for bioinformatics (2000s)

verb	example	raw freq.	%
use	*we used the time-independent*	823	9.36
find	*We have found clefts in the protein*	340	3.87
consider	*we consider a Markov model*	249	2.83
show	*We will show that*	240	2.74
apply	*we apply the tests to*	214	2.43

Table 5.30: Top 5 verbs used with *we* and *I* (*S_self-mention*) for bioinformatics (2000s)

the beginning of a sentence, the top 5 verbs almost match self-mention expressed in general, besides the use of the verb *compare* (see Table 5.31). Thus, rather than considering, writers *compare* in bioinformatics when they mention themselves at sentence beginning.

Boosters Bioinformatics is also characterized by a distinctive usage of boosters at sentence beginning (*Sb_booster*). Here, obviousness is by far the most frequent epistemic meaning used with 64.3%. Regarding the lexical items used most frequently to express obviousness, the adverbs *obviously* and *clearly* prevail, evaluating the sentence that follows them (see Examples (44) and (45)).

(44) ***Obviously***, *the smaller the M-RFP is, the better the results are.*

verb	example	raw freq.	%
use	*We only use the data from*	255	9.97
find	*We also found that*	104	4.07
compare	*We compared our approach to*	81	3.17
show	*We will show that*	70	2.74
apply	*We applied this algorithm*	69	2.70

Table 5.31: Top 5 verbs used with *we* at sentence beginning (*Sb_self-mention*) for bioinformatics (2000s)

lexical item	example	raw freq.	%
new	*The new method finds more*	794	87.54
novel	*we propose a novel method to find*	100	11.03
modern	*many modern modelling tools allow*	13	1.43

Table 5.32: Lexical items expressing progress for bioinformatics (2000s)

target	example	raw freq.	%
method	*represents a new method for*	80	8.82
algorithm	*Here we report a new algorithm*	45	4.96
approach	*Thus a novel knowledge-based approach*	32	3.53
sequence	*these new protein sequences*	24	2.65
tool	*This new tool provides*	22	2.43

Table 5.33: Top 5 targets evaluated with progress for bioinformatics (2000s)

(45) ***Clearly****, the tests on all pairs gave greater information.*

The distinctive feature of stance adverbs at the beginning of a sentence (*Sb_RB*) is in some way related to boosters, especially, when we consider the most often occurring stance meaning expressed by adverbs, which besides idiosyncrasy (19.5%; see Example (46)), is obviousness (17.6%).

(46) ***Specifically****, we used the Gram-positive bacteria dataset that was present in two forms*: [...].

Stance meanings Besides self-mention and boosters, bioinformatics is distinguished by specific attitudinal meanings as well as one epistemic meaning. Most distinctively, it evaluates with the meaning of progress (*S_progress*), i.e. things are new, novel and modern (see Table 5.32), and these things are related to either methodological concepts (*method*, *approach*), computation (*algorithm*, *tool*) or biology (*sequence*) (see Table 5.33).

Additionally, bioinformatics uses importance at sentence beginning (*Sb_importance*) in a distinct way, the lexical items being mainly adverbs (e.g., *importantly*, *essentially*) attributing the importance meaning to the whole following clause (see Table 5.34).

Furthermore, the attitudinal meaning of trust (*S_trust*) is distinctive, realized by the lexical items shown in Table 5.35. Considering the adjective *reliable*, the targets evaluated by it are quite varied if we consider the different types (120 in 137 occurrences). The most often evaluated targets are *result* (4.42%) and *estimation* (3.87%), i.e. targets related to experimental concepts.

Besides attitudinal meanings, there is also one epistemic meaning which is distinctive for bioinformatics, the assumption meaning (*S_assumption*). Table 5.36 shows the

lexical item	**example**	**raw freq.**	**%**
importantly	*Importantly, our algorithm requires*	16	21.92
essentially	*Essentially, CopasiSE allows*	11	15.07
notably	*Notably, the method may support*	8	10.96
significant	*Significant improvement in agreement*	8	10.96
important	*Important progress has been made*	5	6.85

Table 5.34: Top 5 lexical items realizing importance at the beginning of a sentence for bioinformatics (2000s)

lexical item	**example**	**raw freq.**	**%**
reliable	*provided more reliable results*	137	75.69
reliably	*Essentially, CopasiSE allows*	29	16.02
unreliable	*may yield unreliable results*	15	8.29

Table 5.35: Lexical items realizing trust for bioinformatics (2000s)

top 5 lexical items that realize assumption in bioinformatics. Interestingly, the verb *expect* is quite frequently used (26.77%) to assume which outcomes might occur, as shown in the example where an improvement of accuracy is expected by the classifier.

Targets The only target feature in the top 10 for bioinformatics is *T1_NNS*, i.e. uni-gram targets of the type of common nouns in the plural form (see Table 5.37). The top 5 targets are related either to computational and methodological concepts (*results*, *predictions*, *data*, *methods*) or biology (*genes*), similarly to the targets related to *S_progress*. Furthermore, we can say that results are best (approx. 50.6%), while genes are significant (approx. 48.6%). We obtain these observations by inspecting which meanings are most often used to evaluate the T1_NNS targets (see Table 5.38) as well as the specific targets evaluated with these meanings. Considering the patterns used to evaluate T1_NNS targets (see Table 5.39), the *eval-adj_target-np* pattern clearly prevails with 86.5%, while the other patterns are less often used. Note that the other patterns are all of the type where the target precedes the eval-

lexical item	**example**	**raw freq.**	**%**
expect	*classifier is expected to improve the accuracy*	581	26.77
assume	*phylogenetic data can be assumed to have evolved*	481	22.17
appear	*patients in this study appear to be*	389	17.93
assumption	*based on the assumption that*	302	13.92
seem	*method which seems to produce*	180	8.29

Table 5.36: Lexical items realizing assumption for bioinformatics (2000s)

target	example	raw freq.	%
results	*gD3 gave the best results*	83	5.16
genes	*the list of interesting genes is obtained by*	70	4.35
predictions	*biologically relevant predictions*	39	2.42
data	*fast access to relevant data*	34	2.11
methods	*the new methods exhibit*	30	1.86

Table 5.37: Top 5 uni-gram targets in the plural form (*T1_NNS*) for bioinformatics (2000s)

meaning	example	raw freq.	%
desirability	*Successful results are presented*	385	23.91
importance	*discover important genes*	314	19.50
likelihood	*Possible explanations were that*	240	14.91
progress	*rate of detecting novel polyketides*	122	7.58
complexity	*applied to more complex systems*	95	5.90

Table 5.38: Stance meanings evaluating uni-gram targets in the plural form (*T1_NNS*) for bioinformatics (2000s)

uation.

In summary, bioinformatics is characterized by a distinctive use of self-mention (in general but also at sentence beginning), with a preference of material verbal processes expressed. Also characteristic is the enforcement of concepts by boosters of obviousness. Considering the meanings expressed, attitudinal meanings are most characteristic, yet the assumption meaning as an epistemic meaning is also a distinctive feature of bioinformatics by which expectations are expressed. From the characteristic attitudinal meanings, progress is mostly expressed toward methodological and computational concepts (i.e. *methods/algorithms/tools* are *new/novel*), while trust relates to experimental ones (i.e. *results/estimations* are *reliable*). The meaning of importance, by contrast, mostly expressed by adverbs, is related to a

pattern	example	raw freq.	%
eval-adj_target-np	*remarkable* ***similarities***	1386	87.83
target-np_rel-v_eval-expr	*These* ***genes*** *are good classifiers*	96	6.08
target_eval-v	*the proposed* ***methods*** *outperformed random-order based K2*	72	4.56
target_v_eval-expr	*both* ***algorithms*** *gave rather unstable results*	24	1.52

Table 5.39: Patterns used with uni-gram targets formed of plural nouns (*T1_NNS*) for bioinformatics (2000s)

feature	examples	Sent. b.	subtype	S_sem	weight
P_benefit	*The advantage of this approach*		P_m	attitudinal	1.40
Pb_progress	*The novel stem shapes*	x	Pb_m	attitudinal	1.37
S_complexity	*simple*, *complex*		S_m	attitudinal	1.34
Sb_self-mention	*we*, *our*	x	Sb_type		1.34
S_attitudinal			S_sem		1.23
S_att-marker			S_type		1.17
T1_length	*factors*, *developments*		T_length		0.90
Sb_attitudinal		x	Sb_sem		0.83
Sb_idiosyncrasy	*specifically*, *especially*	x	Sb_m	attitudinal	0.78
S_suggestion	*the Mediative protocol suggested here*		S_m	epistemic (hedge)	0.75

Table 5.40: Top 10 features for digital construction (2000s)

variety of concepts. In terms of targets, uni-gram targets of nouns in the plural form are characteristic, matching the type of targets evaluated by the attitudinal meanings of progress and trust.

5.2.2.3.3 B3: Digital construction Of the top 10 features characteristic for digital construction (see Table 5.40), three groups are mainly involved in distinguishing B3 from the other registers: (1) pattern features, which are quite strong discriminators, (2) stance meanings (two attitudinal and one epistemic), and (3) self-mention. Specific stance meanings are less distinctive for digital construction in comparison to B1 and B2. Considering the target features, there is only one that is characteristic for digital construction, uni-gram targets. From the top 10 features, four are used at sentence beginning.

Patterns Highly characteristic for digital construction is to express benefit in a quite conventionalized way within patterns (*P_benefit*). Mostly, it is expressed within a prepositional pattern (54.89%) such as *the advantage of* (see Table 5.41), with the noun *advantage* being the most often used evaluative noun within this pattern (44.3%; see Table 5.42). Advantageous for digital construction are methodological concepts (e.g., *method* with 20.0%, *approach* with 14.0%) as well as computational ones (e.g., *system* with 10.0%, *algorithm* with 6.0%). Furthermore, benefit is also expressed by verbal processes in digital construction, with the verbs *facilitate* (66.7%), *help* (22.2%) and *assist* (8.9%). What is facilitated are mostly computational concepts (44.4%) such as *computation*, *implementation*, and *simulation*.

Additionally, digital construction also uses the progress meaning, quite conventionalized in patterns occurring at the beginning of a sentence (*Pb_progress*). The *eval-adj_target-np* pattern is the most often occurring one with 95.4% (e.g., *The new*

pattern	example	raw freq.	%
eval-np_prep_target-np	*the advantages of both methods*	113	54.85
to_eval-v_target-np	*to facilitate the estimation process*	45	21.84
target_eval-v	*The ... system facilitates the ... acquisition*	34	16.50
eval-gerund_target	*facilitating the computer-aided collaboration*	5	2.43
target-np_rel-v_eval-expr	*a multiple colony approach is beneficial*	4	1.94
eval-adj_target-n	*the most favorable path*	3	1.46
it_rel-v_eval-adj_target-cl.	*it could be advantageous to employ*	2	0.97

Table 5.41: Patterns expressing benefit (*P_benefit*) for digital construction (2000s)

noun	example	raw freq.	%
advantage	*The advantage of this approach*	50	44.25
help	*a useful help to the user in the process*	17	15.04
benefit	*the benefits of the genetic-based MAs*	13	11.50
disadvantage	*The disadvantage of this approach*	11	9.73
assistance	*with the assistance of the product agent*	4	3.54

Table 5.42: Evaluative nouns of benefit within the *eval-np_prep_target-np* pattern for digital construction (2000s)

approach, *a new procedure*) and concepts evaluated are related to methodological aspects (e.g., *approach*, *problem*, *method*).

Stance meaning Quite characteristic for digital construction is attitudinal meaning, especially the attitudinal meaning of complexity (*S_complexity*), which is used to evaluate things mostly as simple or complex (see Table 5.43). Methodological (e.g., *method*, *approach*) as well as technical concepts (e.g., *geometrical structures*, *amplifier*) are evaluated as simple; computational and also technical concepts (e.g., *agent-based systems*, *engineering tasks*) are evaluated as being complex.

Also characteristic is the use of idiosyncrasy at sentence beginning (*Sb_idiosyncrasy*), similarly to computational linguistics (B1), but for digital construction it is less distinctive (rank 9 for B3 vs. rank 2 for B1). Moreover, digital construction uses the lexical items that realize idiosyncrasy quite evenly (*specifically* with 54.1% and *especially* with 45.9%).

In addition to the two attitudinal meanings of complexity and idiosyncrasy, digital construction uses the epistemic meaning of suggestion (*S_suggestion*) distinctively. In this sense, digital construction proposes and suggests (see Table 5.44). Proposed

lexical item	example	raw freq.	%
simple	*via a simple thresholding process*	522	18.89
complex	*may be very complex*	470	17.01
easily	*has been easily and precisely identified*	254	9.19
complexity	*Despite the complexity involved*	232	8.40
difficult	*that was difficult to achieve previously*	219	7.93

Table 5.43: Lexical items expressing complexity for digital construction (2000s)

lexical item	example	raw freq.	%
propose	*propose an object communication model*	1408	54.89
should	*Such model should regulate interaction*	883	34.42
suggest	*as it is suggested in Fig.*	212	8.27
suggestion	*This program mainly gives suggestions*	32	1.25
shall	*The business logic shall be able to*	27	1.05
ought	*We ought to be able to*	3	0.12

Table 5.44: Lexical items expressing suggestion for digital construction (2000s)

are mostly computational concepts (approx. 16.8%, such as *algorithm*, *system* (e.g., *parallel 3D RKPM system*), and *model* (e.g., *NN-MT model*)) as well as methodological ones (approx. 14.5%, such as *method* and *approach*). In terms of what should be done, digital construction suggests mostly material processes (e.g., *carry out*, *perform*, and *minimize*), but also mental processes (e.g., *note* (which is the most often occurring verb after *should* with 5.7%), *consider*, and *take into account*).

Self-mention Digital construction is also characterized by self-mention at the beginning of a sentence (*Sb_self-mention*). Most prominently, the pronoun *we* is used (83.6%), followed by *our* (15.0%). By considering the verbal processes expressed, we can deduce that writers in research articles on digital construction use, besides material (e.g., *use*, *apply*) and mental processes (e.g., *consider*, *assume*), communication processes (e.g., *propose*, *call*) as well (see top 5 verbs in Table 5.46).

In summary, digital construction is most prominently characterized by pattern fea-

pronoun	example	raw freq.	%
we	*We will use the term component*	761	83.63
our	*Our design follows these guidelines*	136	14.95
I	*I decided to extend HI-RIS*	8	0.88
my	*My first exposure to KBES*	4	0.44
ours	*Ours has a result in three iterations*	1	0.11

Table 5.45: Pronouns used at sentence beginning for self-mention for digital construction (2000s)

verb	**example**	**raw freq.**	**%**
use	*We use dynamic scheduling*	75	9.86
propose	*We propose a modified version*	30	3.94
consider	*We consider another situation*	27	3.55
call	*We call a a lower length limit*	23	3.02
assume	*We assume that the designer*	19	2.50

Table 5.46: Verbs used after self-mention at sentence beginning for digital construction (2000s)

feature	**examples**	**Sent. b.**	**subtype**	**S_sem**	**weight**
Sb_Main		x	Sb_doc		1.17
T1_NP	*SEUs*, *DAC*		T_pos		1.16
S_Conclusion			S_doc		1.16
S_idiosyncrasy	*especially*, *specifically*		S_m	attitudinal	1.08
S_typicality	*usually*, *common*		S_m	attitudinal	1.04
Pb_capability	*capable*, *able*	x	Pb_m	attitudinal	0.96
S_self-mention	*we*, *our*		S_m		0.94
S_usefulness	*useful*, *efficient*		S_m	attitudinal	0.80
Pb_it_rel-v_eval-adj_target-clause	*it is important that*	x	Pb_type		0.78
Sb_booster	*clearly*, *obviously*	x	Sb_type	epistemic	0.78

Table 5.47: Top 10 features for microelectronics (2000s)

tures indicating a quite conventionalized usage of evaluative meaning in comparison to B1 and B2, especially when expressing the advantage of methods and computational concepts and how these are facilitated, as well as the novelty of methodological concepts. The expression of complexity (either simple or complex) is also distinctive and directed toward methods and technical as well as computational concepts. In addition, digital construction suggests and proposes concepts related to computation. Finally, self-mention is distinctive, and writers of research articles in digital construction not only *use/apply* or *consider/assume*, but also *propose/call*, i.e. they put forward their own concepts.

B4: Microelectronics The top 10 features for microelectronics are shown in Table 5.47 and are quite varied if we consider the different feature types covered: (1) expression of evaluative meaning within specific document sections (Main, Conclusion), (2) target feature of proper noun uni-grams, (3) one epistemic and three attitudinal stance meanings, (4) pattern features related to meaning and pattern type, and (5) self-mention. Thus, microelectronics shows a relatively variable profile of its characteristics of evaluative meaning in comparison to the other contact registers we have seen so far.

stance type	example	raw freq.	%
self-mention	***We*** *borrow the idea from*	1727	76.11
att-marker	***Unfortunately****, the complexity*	260	11.46
hedge	***Possible*** *approaches are*	161	7.10
booster	***Obviously****, this methodology has*	121	5.33

Table 5.48: Stance types used within the Main part of research articles at sentence beginning for microelectronics (2000s)

stance type	example	raw freq.	%
att-marker	*The architecture competes* ***favorably*** *with*	1455	45.84
hedge	*We* ***proposed*** *a new method*	836	26.34
self-mention	***We*** *have presented*	547	17.23
booster	*shown that PTL can* ***certainly*** *improve*	336	10.59

Table 5.49: Stance types used within the Conclusion section of research articles for microelectronics (2000s)

Document section The expression of evaluative meaning at sentence beginning within the Main part of research articles (*Sb_Main*) is most characteristic for microelectronics. Considering the stance types expressed, writers most prominently refer to themselves by self-mention within the Main part (see Table 5.48). Considering stance expressions within the Conclusion section (*S_Conclusion*), however, attitude markers are used most often (compare to Table 5.49). Thus, while in the Main part writers of microelectronic articles have a preference for positioning themselves within the discourse, in the Conclusion section they tend to evaluate.

Targets What writers evaluate distinctively within microelectronics are uni-gram targets of proper nouns, which mostly consist of acronyms. These targets are most frequently evaluated by the meanings of desirability and progress. Thus, concepts or tools related to microelectronics are mostly evaluated as *best* (25.7%; see Example (47)) or *modern* (11.4%; see Example (48)).

(47) *Column 4 lists the leakage current for each circuit when the* ***best MLV*** *is applied.*

(48) *Logic and memory resources that efficiently implement specific functionality commonly used in system-sized designs, such as multipliers or large memories, are already a part of* ***modern FPGAs****.*

Attitudinal meaning Microelectronics is also characterized by the use of idiosyncratic attitudinal meaning. While for B1 and B3, idiosyncrasy was distinctive at sentence beginning, for microelectronics it is distinctive in general (*S_idiosyncrasy*).

lexical item	example	raw freq.	%
especially	*especially power simulation systems*	115	28.05
specifically	*Our DFT scheme is specifically targeted*	104	25.37
unique	*A unique feature of our approach*	90	21.95
extensively	*Our approach is extensively evaluated*	33	8.05
uniquely	*equivalence is uniquely characterized by*	8	1.95

Table 5.50: Top 5 lexical items expressing idiosyncrasy for microelectronics (2000s)

lexical item	example	raw freq.	%
usually	*algorithms are usually performed on*	214	46.32
common	*the most common digital systems*	198	42.86
representative	*the input data or a representative subset of it*	45	9.74
unusual	*Traces created on unusual code paths*	4	0.87
unusually	*are chosen (unusually) high to demonstrate*	1	0.22

Table 5.51: Top 5 lexical items expressing typicality for microelectronics (2000s)

This is also why the lexical items realizing this meaning in microelectronics show a greater variation (see Table 5.50) than for B1 and B3. Additionally, microelectronics is also characterized by the stance meaning of typicality (*S_typicality*) used to evaluate mostly either material processes (such as *perform* or *implement*) with *usually* (46.3%), or methodological (e.g., *approach*, *method*) and computational concepts (e.g., *implementation architecture*, *digital systems*) as *common* (42.9%) (see Table 5.51). Stance expression of usefulness (*S_usefulness*) are also characteristic for microelectronics. Table 5.52 shows the top 5 lexical items used to express this meaning. Most things are evaluated as being efficient and represent mostly computational concepts (approx. 18%, e.g., *implementation* and *algorithm*).

Patterns Characteristic for microelectronics is also the meaning of capability expressed within patterns at sentence beginning (*Pb_capability*). This meaning is mostly expressed by (1) the relational pattern *target-np_rel-v_eval-expr* (approx. 58.1%), where the target precedes the evaluative expression (see Example (49)), (2) the *target_eval-v* pattern (approx. 19.4%), where the target is evaluated by a verb

lexical item	example	raw freq.	%
efficient	*provide an efficient model for CAD tools*	407	35.30
effective	*internal memory is an effective method*	260	22.55
useful	*skew routings which are useful in reducing*	123	10.67
effectiveness	*The effectiveness of the K-L heuristic*	105	9.11
effectively	*by effectively exploiting the regularities*	104	9.02

Table 5.52: Top 5 lexical items expressing usefulness for microelectronics (2000s)

meaning	example	raw freq.	%
importance	*It is important to note that the edge set*	46	31.29
likelihood	*It is possible that it will become a preferred approach*	29	19.73
obviousness	*It is clear that the proposed partial-scan method*	28	19.05
interest	*It is particularly interesting to notice that*	16	10.88
complexity	*It is not difficult to see that*	15	10.20

Table 5.53: Top 5 meanings used with the *it_rel-v_eval-adj_target-clause* pattern at sentence beginning for microelectronics (2000s)

(see Example (50)), and (3) the prepositional pattern *eval-np_prep_target-np* (approx. 16.1%; see Example (51)). In addition, microelectronics is also characterized by the use of the *it_rel-v_eval-adj_target-clause* pattern at sentence beginning. In this case, the writer points the reader's attention toward the clause following the evaluative expression. Mostly, the importance meaning is used within this pattern (see Table 5.53), i.e. the writer wants the reader to be aware that what follows is of importance. This is also indicated by the use of the verb *note* after the *to* particle, which clearly prevails (45.65%), followed by *observe* (6.52%), which has a similar function.

(49) [$_{\textbf{target}}$ *The initial approach*] [$_{\textbf{rel-v}}$ *is*] [$_{\textbf{eval-expr}}$ *unable*] *to satisfy any of the TCP constraints* [...].

(50) [$_{\textbf{target}}$ *This approach*] [$_{\textbf{eval-v}}$ *enables*] *the possibility of having three processes running concurrently* [...].

(51) [$_{\textbf{eval-np}}$ *The ability*] [$_{\textbf{prep}}$ *of*] [$_{\textbf{target-np}}$ *the algorithm]* *to determine an efficient memory binding* [...].

Self-mention The use of self-mention (*S_self-mention*) is also relatively characteristic for microelectronics, the most prominent pronouns being *we* and *our* (see Table 5.54), similarly to B1 and B2. Looking at *we* plus verb, the material verb *use* occurs most often, but communication verbs (*present* and *propose*) are also among the top 5 verbs, followed by mental verbs (*assume* and *consider*). Considering nouns used with *our*, methodological and computational concepts are most often used (*approach*, *technique*, *algorithm*).

In summary, microelectronics has a wide spectrum of types of features that characterize it in terms of evaluative meaning. It is most predominately characterized by self-mention at sentence beginning in the Main part of research articles and the use of attitude markers within the Conclusion section. The evaluation of acronyms

pronoun	example	raw freq.	%
we	*We evaluate the performance*	7471	73.59
our	*combining analysis with our experimental data*	2438	24.01
us	*Let us consider that*	229	2.26
ours	*had aims similar to ours*	9	0.09
ourselves	*We also restrict ourselves to*	5	0.05

Table 5.54: Pronouns expressing self-mention for microelectronics (2000s)

is also distinctive. In terms of stance meanings, the peculiarity of things is characteristic in microelectronics, with the idiosyncrasy meaning as well as the typicality of material process types or methodological and technical concepts. In addition, microelectronics is also characterized by the evaluation of computational concepts with usefulness/efficiency. It is also characterized by a conventionalized use of showing the reader what either the researchers or their approaches are capable of doing, by the expression of the capability meaning as well as pointing the reader toward important parts of the discourse.

5.2.3 Summary and conclusions on register diversification

In terms of register diversification, we wanted to test whether contact and seed registers show a greater distinctness over time, forming their own clusters of lexico-grammatical features of evaluative meaning. Thus, we considered the degree of register diversification, i.e. how well registers are distinguished from each other and how this has evolved in the 30-year time span investigated, as well as the kind of register diversification, i.e. by which lexico-grammatical features they are characterized.

Degree of register diversification By inspecting the SciTex registers in macro-analytical terms, we have seen that there is a clear improvement in the classification performance in the 2000s in comparison to the 70s/80s. Thus, all registers have undergone a diversification process in terms of evaluative meaning. However, this development is especially pronounced in the contact registers, which have shown a great improvement in the classification performance over time of approx. 50–70%. This indicates that while the contact registers were not readily distinguishable from the other registers in the 70s/80s, showing very low classification performance and no characteristic usage of evaluative features, they clearly develop their own character of evaluative meaning in the 2000s with a classification performance comparable to the seed registers, which is relatively high, with F-measures ranging from 0.88 to 0.94. Thus, while the seed registers show a characteristic usage of evaluative meaning already in the 70s/80s, the contact registers develop a characteristic usage in the 30-year time span investigated, achieving a relatively high degree of register diversification in the 2000s.

Kind of register diversification Even though we could not inspect detailed diachronic register diversification trends in terms of the kind of diversification for the contact registers due to their low classification performance in the 70s/80s, we still considered feature types and subtypes used across both time periods, as they might point to general trends involved in forming a characteristic usage of evaluative meaning. Most prominently, we observed a more variate range of feature subtypes in the top 5 most distinctive features, i.e. while features contributing to a quite moderate distinction in the 70s/80s consisted mostly of features of specific stance meanings only, in the 2000s the characteristic features belong to several subtypes (e.g., part-of-speech of stance expressions, pattern types, etc.). Thus, the more varied the features in terms of subtypes, the better the classification performance, i.e. the better they can be distinguished. This is clearly related to the fact that some features are interrelated, and thus point toward the same evaluative phenomenon characteristic of a particular register, i.e. as a particular evaluative phenomenon becomes characteristic for a register, it will be reflected across several feature subtypes. Therefore, as registers develop their own characteristic usage of particular types of evaluative meaning, this will be reflected in more interrelated features being most distinctive for that register. Additionally, there is also a tendency toward expressing evaluative meaning at sentence beginning over time, a tendency that has also been observed for both seed and contact registers.

Both observations were also reflected in the more detailed feature analysis positioned on the micro-analytical level. For the seed registers, we inspected computer science (A), linguistics (C1), and biology (C2) as they show the best classification performance for both time periods, and thus are best suited to investigate diachronic tendencies.

For computer science, we have seen that while the epistemic meaning of obviousness has become quite distinctive over time, the complexity meaning has also evolved to be quite characteristic for computer science as it is reflected in several top 10 features. Moreover, a conventionalization process seems to take place within computer science, especially for these two characteristic stance meanings, as they are distinctively used at sentence beginning. Additionally, this trend is also reflected in the characteristic usage of self-mention at sentence beginning.

Linguistics, on the other hand, has a rather hedged evaluative profile that is even more strongly reflected over time. For the 70s/80s, individual hedge meanings are distinctive, while in the 2000s, several distinctive feature types reflect hedge expressions (e.g., hedges at sentence beginning, parts-of-speech of nouns and verbs that realize hedges, patterns containing hedges). Moreover, a kind of conventionalization trend can also be observed in linguistics in terms of using hedges at sentence beginning rather than other feature, which might point to less variation, and thus towards conventionalization.

Biology has a preference in expressing epistemic expressions by hedges (e.g., relativity, suggestion) and boosters (e.g., certainty), especially within patterns and at sentence beginning. These epistemic meanings are mostly directed toward experimental results in terms of interpretative endeavors, or used to put forward evidence. Thus, diachronically we can say that even though the individual most distinctive features change, biology seems to preserve its epistemic character, which is somehow balanced between expressing hedges on the one hand and boosters on the other.

Regarding the contact registers, we have investigated their characteristic usage of evaluative meaning in detailed micro-analytical terms for the 2000s only, due to the fact that their classification performance in the 70s/80s does not allow for a reliable interpretation of the top-ranking features.

Computational linguistics (B1) shows a distinct usage of self-mention at sentence beginning, where writers introduce what they do (e.g., *use*, *see*) or what they present (e.g., *systems*, *models*). In addition, it uses attitudinal meanings distinctively either at sentence beginning (e.g., desirability, importance) or in general (e.g., trust/ reliability, persuasion). Positioning the evaluative meaning at sentence beginning points again toward some kind of conventionalization process.

Bioinformatics (B2) is also quite distinctively characterized by self-mention, here both in general and at sentence beginning, mostly, combined with material and mental processes (e.g., *use*, *consider*). Additionally, boosters at sentence beginning are quite distinctive expressing obviousness. In terms of meanings, it is characterized by attitudinal (progress, importance, reliability/trust) and epistemic meanings (assumption).

Digital construction (B3) seems to be quite conventionalized in terms of evaluative meaning, as it is characterized by pattern features and self-mention at sentence beginning. Moreover, it is characterized by evaluating methodological and computational concepts (e.g., as being *advantageous*, *complex/easy* or *novel*) but also by proposing/suggesting them.

Microelectronics (B4) shows quite a varied span of characteristic features in comparison to the other contact registers. However, self-mention at sentence beginning is again a quite distinctive feature, especially used in the Main part of research articles, while attitudinal meanings associated with peculiarity (idiosyncrasy), typicality or efficiency as well as importance are mostly put forward in the Conclusion section.

Thus, we have clearly seen on a macro- and micro-analytical level how registers become more diversified over time, being better distinguished from other registers and creating characteristics of their own in terms of evaluative meaning. We have seen that by the micro-analytical inspection in terms of feature analysis, profiles of evaluative meaning of highly specialized registers can be established, indicating what their characteristic usages are. Still, if we compare these characteristics of the seed and contact registers analyzed so far, we can see some overlaps between seed and contact registers (e.g., self-mention in A and the Bs, boosters in C2 and B2). This points to our second hypothesis of registerial imprint, which is tested in the next section.

	A	B1	C1	F-measure	overlap A-B1 (%)	overlap B1-C1 (%)
A	183	16	3	0.897	7.92	
B1	14	100	12	0.797	11.11	9.52
C1	9	9	201	0.924		4.11

A: computer science; B1: computational linguistics; C1: linguistics

Table 5.55: Confusion matrix and overlap for the A-B1-C1 triple in the 70s/80s

	A	B1	C1	F-measure	overlap A-B1 (%)	overlap B1-C1 (%)
A	228	1	1	0.987	0.43	
B1	2	134	1	0.982	1.46	0.73
C1	2	1	108	0.977		0.90

A: computer science; B1: computational linguistics; C1: linguistics

Table 5.56: Confusion matrix and overlap for the A-B1-C1 triple in the 2000s

5.3 Registerial imprint

5.3.1 Degree of registerial imprint

To be able to measure the degree of registerial imprint on the contact registers, we consider, on a macro level of analysis, the confusion matrices of each triple comparison (A-Bs-Cs), looking at the texts misclassified from the contact registers into their respective seed registers.

Consider, e.g., the A-B1-C1 triple. In the 70s/80s, the overall classification accuracy amounts to 88%. Looking at the F-measures (see Table 5.55), C1 classifies best (0.92), followed by A (0.90) and B1 (0.80), which classifies less well than its seed registers. Moreover, computational linguistics has an overlap of 11.11% with A (14 texts are misclassified from B1 into A) and 9.52% overlap with C1 (12 texts are misclassified from B1 into C1). Additionally, 7.92% of texts are wrongly classified from A into B1, yet only 4.11% are misclassified from C1 into B1. Thus, the overlap of A and B1 seems stronger than between B1 and C1. For the 2000s, the overall classification accuracy is quite high (98%). Considering the F-measures (see Table 5.56), A classifies best (0.987), followed by B1 (0.982) and C1 (0.977). Diachronically, all three registers are better distinguished over time. The number of misclassifications has dropped to a minimum. Computational linguistics shows an overlap of 1.46% with A and 0.73% with C1, only. From the seed register A, 0.43% are misclassified into B1, and from C1, 0.90% into B1. Thus, in the 2000s, computational linguistics does not seem to have a particular tendency toward one or the other seed register.

Table 5.57 shows the misclassification in percentages for all contact registers. We can see that the percentage of misclassifications for the contact registers into the

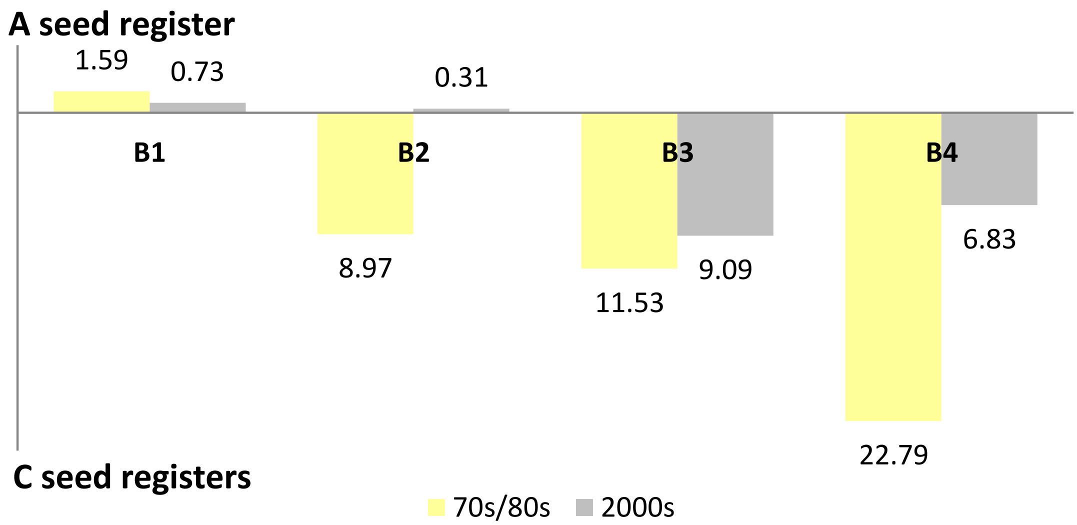
	70s/80s	2000s	tendency
B1 into A	11.11	1.46	-
B1 into C1	9.52	0.73	-
B2 into A	2.41	2.82	+
B2 into C2	11.38	2.51	-
B3 into A	4.4	0	-
B3 into C3	15.93	9.09	-
B4 into A	1.84	0	-
B4 into C4	24.63	6.83	-

A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C1: linguistics; C2: biology; C3: mechanical engineering; C4: electrical engineering

Table 5.57: Percentage of texts misclassified from the contact registers into their seed registers for both time periods

seed registers diminishes over time, except for bioinformatics (B2), where misclassifications into computer science (A) increase slightly. From the numbers, we can also see how B1 and B2 become quite distinct from their seed registers over time (B1 less than 2% misclassified, B2 less than 3%). The engineering contact registers B3 and B4, however, although they show fewer misclassifications diachronically, still show an overlap with their C seed registers of 9% for B3 and 7% for B4. Nevertheless, in general the degree of registerial imprint seems to diminish over time. Furthermore, we can also observe some individual tendencies. Computational linguistics lies, in both time periods, between its two seed registers, showing approx. 10% overlap in the 70s/80s and around 1% in the 2000s. Bioinformatics tends more toward biology in the 70s/80s, while it seems to lie in between A and C2 in the 2000s. Digital construction and microelectronics both tend toward their C seed registers in both time periods. Thus, the degree of registerial imprint varies across contact registers.

We can inspect this further by considering with which seed register each contact register has greater overlap, i.e. shows a greater degree of registerial imprint. Figure 5.7 shows into which seed registers the contact registers misclassify most, i.e. we calculated the difference between misclassifications into A and misclassifications into C. This is shown for both time periods. Considering computational linguistics (B1), we can see how it has more misclassifications into A for both time periods (1.59% more misclassifications into A than C1 for the 70s/80s, and 0.73% in the 2000s). Thus, we can deduce that computational linguistics seems to tend more toward computer science in both time periods, considering the use of evaluative meaning. For bioinformatics (B2), we can see how it shows more overlap with C2 in the 70s/80s, while it shows more overlap with A in the 2000s. Thus, bioinformatics has undergone a shift from being more similar to biology in the 70s/80s (with almost 9% more overlap in terms of more misclassified texts from B2 into C2 than B2

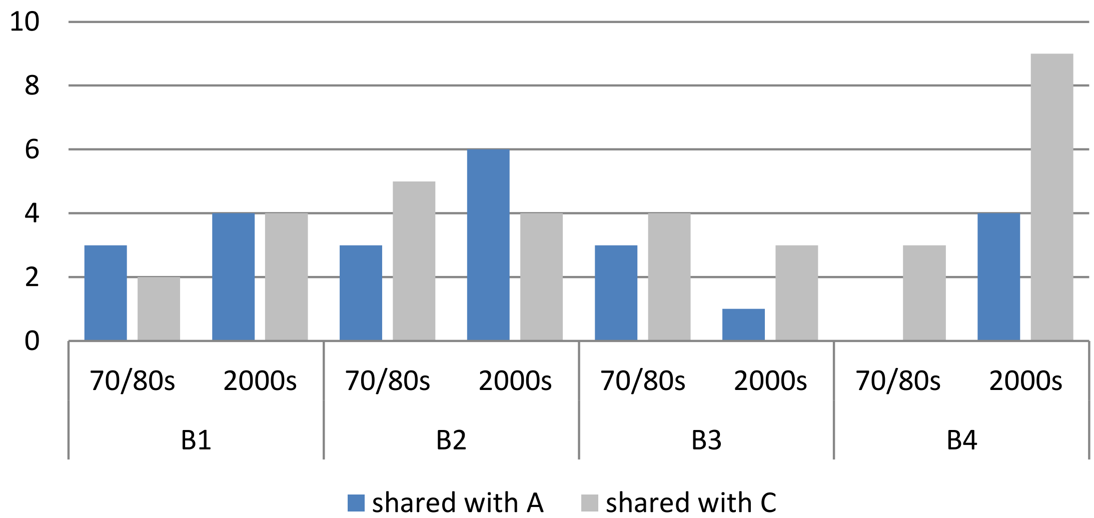


A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C: seed registers

Figure 5.7: Diachronic tendencies of overlaps for contact registers into seed registers

into A) to being more similar to computer science in the 2000s (though it has only 0.31% more overlap with A than C2). Digital construction (B3) shows in both time periods more overlap with its C seed register mechanical engineering (C3), but the overlap diminishes over time. For microelectronics (B4), we see a similar tendency, i.e. it shows, in both time periods, more overlap with electrical engineering (C4) than with computer science. Still, here as well, the overlap diminishes over time.

On a micro level of analysis, which still gives insights on the degree of imprint, we look at the number of features shared with A and Cs by the Bs, i.e. shared among the top 20 features for the Bs and A when compared to the Cs (i.e. features shared with A) as well as the Bs and Cs when compared to A (i.e. features shared with Cs). This is illustrated in Figure 5.8, where the numbers of features shared are shown across contact registers and time periods.
For computational linguistics, in both time periods the number of features shared is very similar, even though it shares slightly more features with A in the 70s/80s (3 vs. 2), while the number of features shared for the 2000s is the same (4 for both). This indicates that computational linguistics does not show a stronger overlap with computer science or linguistics. This is in some way also reflected in the percentage of overlap shown in Figure 5.7, where computational linguistics seems to tend more toward computer science, yet the percentage of overlap is quite small and in terms of features shared, no clear tendency can be shown.



A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C: seed registers

Figure 5.8: Number of features adopted from the seed registers by the contact registers over time

For bioinformatics, we can observe that in the 70s/80s it shares more features with biology (5 vs. 3), while in the 2000s more features are shared with computer science (6 vs. 4). This matches the shift observed in the macro-analysis in Figure 5.7. However, the overlap in terms of misclassifications is less strong (0.31%) than the overlap in terms of features shared (6), as the number of features shared with computer science has increased over time from 4 to 6 (considering the top 20 features). For digital construction, more features are shared with mechanical engineering in both time periods (4 vs. 3 in the 70s/80s and 3 vs. 1 in the 2000s).
Similarly, for microelectronics, more features are shared with electrical engineering in both time periods (3 vs. 0 in the 70s/80s and 9 vs. 4 in the 2000s). Interestingly, the number of features shared between microelectronics and electrical engineering is quite high (almost half of the features are shared in the top 20).

Thus, in terms of the degree of registerial imprint, the contact registers show quite individual tendencies toward their seed registers.

		shared with A			shared with C		
		stance	pattern	target	stance	pattern	target
70s/80s	B1	■		■■	■■	■	
	B2	■■■			■■■■	■	
	B3	■	■■	■	■■	■■	
	B4				■■	■	
2000s	B1	■■	■■		■■■	■	
	B2	■■■	■■	■	■■	■■	
	B3		■		■	■	■
	B4	■		■■	■■■■■■		

A: computer science; B1: computational linguistics; B2: bioinformatics; B3: digital construction; B4: microelectronics; C: seed registers

Table 5.58: Number of features shared according to feature type

5.3.2 Kind of registerial imprint

5.3.2.1 Diachronic tendencies

By considering the features shared among contact and seed registers, besides the degree of registerial imprint, we can inspect more closely the kind of registerial imprint. Table 5.58[3] shows the number of features according to feature types (stance, pattern, target) shared among contact and seed registers over time. There are two main observations which can be made. First, contact registers mostly share stance, especially, but also pattern features with their C register, most prominently in the 70s/80s. Target features, in contrast, are mostly shared with computer science. Second, diachronically, there seems to be a slight trend toward sharing more pattern features with computer science.

5.3.2.2 Shared features

Besides feature types, we can inspect the individual features shared among Bs and A and Bs and Cs when compared to the other seed register. This will point to similarities in the use of evaluative meaning between contact and seed registers. Note that we continue considering the features shared among the top 20 features. Here, we will focus on the contact registers computational linguistics (B1) and bioinformatics (B2), to give two examples of the kind of registerial imprint in terms of shared features.

5.3.2.2.1 Computational linguistics From Table 5.59, we can see that computational linguistics and linguistics have a distinct use of the meaning of assumption

[3]Note that one black square represents one feature.

	70s/80s	2000s
	S_assumption	**S_assumption**
B1 and C1 vs. A	S_importance	S_relativity
		S_RB
		P_target-np_rel-v_eval-expr
	T3_length	P_it_rel-v_eval-adj_target-clause
B1 and A vs. C1	S_usefulness	S_Introduction
	T1_algorithm	S_Main
		P_to_eval-v_target-np

A: computer science; B1: computational linguistics; C1: linguistics

Table 5.59: Features shared across contact and seed registers over time in the top 20 features for the A-B1-C1 triple

when compared to computer science (see Examples (52) and (53)). This is a characteristic which is kept over time. Furthermore, while in the 70s/80s B1 and C1 tend to evaluate by importance, they evaluate with relativity adverbs in the 2000s (see Example (54)). Additionally, they both use the *P_target-np_rel-v_eval-expr* to attribute an evaluation toward a target by a relational pattern (see Example (55)).

(52) *While the initial experiments conducted on a set of manually constructed themes **seemed** promising, the system performance deteriorated significantly when it was applied to automatically constructed themes.* (B1: Computational linguistics 2000s)

(53) *Relative clauses and wh-questions in Gaelic **appear** to involve the same basic structure.* (C1: Linguistics 2000s)

(54) *The **rather** heterogeneous set of heuristics currently adopted within the DArtbio system, for example, will need to be replaced* [...]. (B1: Computational linguistics 2000s)

(55) ***Case-based translation is appropriate**, for example, for handling commonly occurring idioms such as greetings.* (B1: Computational linguistics 2000s)

By looking at features shared among computational linguistics and computer science when compared to linguistics, we can see that target features prevail in the 70s/80s: targets of three-token length (*T3_length*; which mostly consist of nominal phrases; see Example (56)) as well as the single-token target *algorithm* (*T1_algorithm*; see Example (57)). Additionally, the usefulness meaning (*S_usefulness*) is typical for B1 and A when compared to C1. In the 2000s, what is shared among B1 and A changes. Features related to patterns and stance expressions within document sections prevail over target and stance meaning features. More concretely, B1 and A use the it-pattern (*P_it_rel-v_eval-adj_target-clause*) as well as the to-infinitive pattern (*P_to_eval-v_target-np*) similarly and both evaluate mostly within introductions and main parts of research articles (*S_Introduction*, *S_Main*) in comparison to C1. Focusing on the it-pattern, while B1 and A use it similarly often in comparison to C1, it is used differently in terms of meanings expressed: B1 uses the pattern to evaluate mostly with likelihood (e.g., *it is possible/likely*), while A uses it to evaluate mostly with complexity (e.g., *it is hard/difficult/easy*). Similarly for the to-infinitive pattern, B1 and A differ in what meanings they express: B1 uses it to express improvements (see Example (58)), while A uses it to express avoidance (see Example (59)).

(56) [...] *the difficulty of* ***preserving orthographic information*** [...].
(B1: Computational linguistics 70s/80s)

(57) *After the model had converged to a local maximum, a very simple* ***algorithm*** *was used to study the time-alignment* [...].
(B1: Computational linguistics 70s/80s)

(58) *Statistical techniques developed for lexicalized grammars* [...] *readily apply to CCG* ***to improve the average parsing performance in large-scale practical applications*** [...]. (B1: Computational linguistics 2000s)

(59) *In order* ***to avoid a dominating influence of long documents****, we simply computed the arithmetic mean of all error rates obtained for the single documents.* (B1: Computational linguistics 2000s)

5.3.2.2.2 Bioinformatics For bioinformatics, we see from Table 5.60 that it shares mainly stance meanings with biology when compared to computer science in the 70s/80s, while sharing mainly pattern features in the 2000s. B2 and C2 share the assumption meaning, similarly to B1 and C1, as well as the capability, benefit and relativity meaning (see Examples (60)–(62), respectively), i.e. they use these meanings more prominently than A. Furthermore, while in the 70s/80s, they share the prepositional pattern *P_eval-np_prep_target-np*, in the 2000s they share the pattern meanings of improvement (*P_improvement*; see Example (63)) as well as likelihood at sentence beginning (*Pb_likelihood*; see Example (64)) in comparison to A.[4]

When compared to biology, bioinformatics and computer science prominently share the use of modal verbs (*S_MD*) in both time periods, even though they have a lower rank in the 2000s. Note that in both time periods, *can* is most prominently

[4] Note that the use of the *P_eval-np_prep_target-np* in the 70s/80s is related to the use of the capability and benefit meanings, which can be expressed by this pattern (see, again, Examples (60) and (61), *capability of creating* and *advantage of the CSA*).

	70s/80s	**2000s**
B2 and C2 vs. A	S_assumption S_capability P_eval-np_prep_target-np S_benefit S_relativity	S_importance P_improvement Pb_likelihood
B2 and A vs. C2	**S_MD** S_complexity S_Conclusion	P_to-eval-v_target-np Sb_self-mention P_desirability Sb_Main **S_MD** T1_NN

A: computer science; B2: bioinformatics; C2: biology

Table 5.60: Features shared across contact and seed registers over time in the top 20 features for the A-B2-C2 triple

used by A and B2, followed by *may*. Additionally, in the 70s/80s the complexity meaning (*S_complexity*) is shared by B2 and A, when compared to C2, as well as a prominent use of stance expressions within the Conclusion section (*S_Conclusion*). While in the 70s/80s stance features are shared, in the 2000s B2 and A share all three types of features (stance, pattern, target) when compared to C2. The to-infinitive pattern (*P_to_eval-v_target-np*) is shared among B2 and A, similarly to B1 and A. Considering the meanings expressed within this pattern, B2, again similarly to B1, mostly uses the improvement meaning within this pattern (see Example (65)). The desirability meaning expressed within patterns (*P_desirability*) is also shared by B2 and A (mostly by the *eval-adj_target-np* pattern such as *The best performing methods*), as well as the expression of self-mention at the beginning of a sentence, which is related to stance expressions at sentence beginning within the Main part of research articles (*Sb_Main*), i.e. self-mention at sentence beginning mostly appears within the Main part. Additionally, single token targets consisting of nouns in the singular form (*T1_NN*) are shared by B2 and A when compared to C2; these mostly consist of nouns relating to computation (*model*, *solution*, *performance* for B2; *solution*, *algorithm*, *result* for A).

(60) *One particularly important by-product of the HCM model is the **capability** of creating a synthetic EMG signal* [...].
(B2: Bioinformatics 70s/80s)

(61) *The primary **advantage** of the CSA was that it provided a simple qualitative picture of temporal variations in the EEG power spectrum.*
(B2: Bioinformatics 70s/80s)

(62) [...] *whereas, the edge of a tumor is **relatively** diffuse.* (B2: Bioinformatics 70s/80s)

(63) *To **overcome** these difficulties we implemented a decision tree method that* [...]. (B2: Bioinformatics 2000s)

(64) ***It is likely that** histone acetylation and transcription factor binding are mutually dependent steps during gene transcription*[...]. (B2: Bioinformatics 2000s)

(65) *Ong et al. (2002) explicitly included operons* [...] ***to improve the quality of the analysis***. (B2: Bioinformatics 2000s)

In summary, we have seen for computational linguistics and bioinformatics how they use features similarly with their seed registers when compared to the other seed register. Thus, so far, we were able to see similarities of contact and seed registers when compared to the other seed register. Some features are even used similarly over time (e.g., the assumption meaning for B1 and C1 and modal verbs for B2 and A). Focusing on computational linguistics, it seems to share a down-toning character with linguistics when we consider the assumption and relativity meaning, while it shares usefulness and mostly target (both 70s/80s) as well as pattern features (2000s) with computer science. Considering B2, it also shares a down-toning character with C2, similarly to B1 with C1, but it also shares other meanings that change over time. With A, B2 shares the use of modal verbs over time, as well as complexity in the 70s/80s, while it shares a variation of features in the 2000s. Thus, computational linguistics and bioinformatics show some similarities, but mostly a specific kind of registerial imprint when compared to the one or other seed registers. However, what we cannot show is which registers adopt features from which other register. We can only see that features are shared.

5.3.2.3 Adopted features

In order to see whether the contact registers adopt features from their seed registers, we have to consider the directionality. For this purpose, we consider the top 10 features typical for the B registers when compared to A and C (i.e. in the triple comparison) and reclassify the three registers by these features. This will show whether features typical for the B register are even more typical for the C register or more typical for A, which would imply that they have been adopted from the one or the other seed register by B.

5.3.2.3.1 Computational linguistics

Features adopted from linguistics Considering the 70s/80s, Table 5.61 shows features typical for C1 in comparison to B1 and features typical for B1 in compar-

typical for C1 vs. B1		typical for B1 vs. A	
feature	weight	feature	weight
S_assumption	3.73	S_capability	2.85
S_relativity	1.58	S_trust	2.07
S_importance	1.15	T1_translation	1.84
S_PP$	1.02	Pb_sufficiency	1.24
P_presumption	0.47	**S_relativity**	1.17
P_target-np_rel-v_eval-expr	0.004	S_improvement	1.0
		P_presumption	0.90
		S_importance	0.81
		P_importance	0.73
		P_suitability	0.34
		S_assumption	0.34
		Pb_desirability	0.30
		P_eval-np_prep_target-np	0.13

A: computer science; B1: computational linguistics; C1: linguistics

Table 5.61: Features adopted from linguistics by computational linguistics (70s/80s)

ison to A. The features marked in bold are the ones adopted from linguistics by computational linguistics, as they are more typical for linguistics when compared to computational linguistics, but also more typical for computational linguistics when compared to computer science. The features adopted from C1 by B1 are the stance meanings of assumption, relativity and importance as well as the presumption meaning used within patterns (see Example (66)). In comparison to the previous analysis, now we are able to say that, e.g., the assumption meaning is not only typical for both C1 and B1 in comparison to A, but it is adopted by B1 from C1 as it is even more typical for C1 than B1. Nonetheless, it is still typical for B1 in comparison to A.

(66) ***Presumably****, if translators are willing to work for this wage, the system must be doing at least half the work* [...].
(B1: Computational linguistics 70s/80s)

Considering the 2000s, we can observe from Table 5.62 that similarly to the 70s/80s, mostly stance meanings are adopted from C1: dispute (e.g., *problem*, *issue*), relativity (e.g., *rather*, *quite*), limitation (see Example (67)), importance at sentence beginning (e.g., *Importantly*), and assumption (e.g., *appear/seem to be*). Diachronically, we can see that assumption and relativity were already adopted in the 70s/80s and have been preserved. Thus, computational linguistics does not only share a downtoning character with linguistics (see Section 5.3.2.2), it has adopted it from linguistics. Additionally, the relational pattern *P_target-np_rel-v_eval-expr* is adopted

typical for C1 vs. B1		typical for B1 vs. A	
feature	weight	feature	weight
Sb_hedge	1.58	Sb_relativity	3.73
Sb_importance	0.87	Sb_evidence	0.80
S_dispute	0.74	**S_dispute**	0.44
S_assumption	0.72	S_suitability	0.43
S_limitation	0.29	S_RB	0.37
S_relativity	0.23	**S_relativity**	0.36
P_Abstract	0.22	**S_limitation**	0.23
P_target-np_rel-v_eval-expr	0.21	**Sb_importance**	0.16
S_Abstract	0.16	**P_target-np_rel-v_eval-expr**	0.15
		S_assumption	0.09

A: computer science; B1: computational linguistics; C1: linguistics

Table 5.62: Features adopted from linguistics by computational linguistics (2000s)

from linguistics as well. Diachronically, we can see that while it was typical for C1 vs. B1, B1 had not adopted it in the 70s/80s, while it has in the 2000s (compare Tables 5.61 and 5.62).

(67) [...] *it is necessary to use analysts who know the scheme about as well as anyone has ever known it - which severely **limits** the candidates available.* (B1: Computational linguistics 2000s)

Features adopted from computer science For the 70s/80s (see Table 5.63), computational linguistics adopts from computer science most prominently the meaning of usefulness (see Example (68)), which makes B1 quite distinctive from C1, being the highest ranking feature. Other meanings adopted are the meanings of typicality at sentence beginning (see Example (69)) as well as the suitability meaning (see Example (70)). Additionally, the feature of three targets length (*T3_length*) is adopted as well. Note that this is a feature shared between B1 and A when compared to C1 in Section 5.3.2.2, i.e. it is not only similarly used between computational linguistics and computer science, it is adopted from computer science by computational linguistics. This is also the case for the usefulness meaning.

(68) [...] *this paradigm feature is extremely **useful**.* (B1: Computational linguistics 70s/80s)

(69) ***Usually**, error analysis is regarded as an aid for didactic methods* [...]. (B1: Computational linguistics 70s/80s)

typical for A vs. B1		typical for B1 vs. C1	
feature	weight	feature	weight
S_PP$	-2.09	**S_usefulness**	-2.96
S_usefulness	-1.29	S_improvement	-1.76
P_target-np_rel-v_eval-expr	-0.47	S_trust	-1.52
Sb_typicality	-0.39	Sb_prediction	-1.21
T3_length	-0.02	**Sb_typicality**	-0.87
S_suitability	-0.02	T1_translation	-0.84
		T3_length	-0.83
		S_suitability	-0.72
		P_suitability	-0.59
		P_eval-np_prep_target-np	-0.52
		Pb_sufficiency	-0.50
		S_capability	-0.48
		Pb_desirability	-0.24
		P_importance	-0.04

A: computer science; B1: computational linguistics; C1: linguistics

Table 5.63: Features adopted from computer science by computational linguistics (70s/80s)

(70) *Recoding input in terms of a more* ***appropriate*** *similarity metric can aid in this type of discrimination.* (B1: Computational linguistics 70s/80s)

In the 2000s (see Table 5.64), the usefulness meaning has been kept over time as being the most typical feature for computational linguistics when compared to linguistics. Note, however, that it is still more typical for computer science when compared to computational linguistics. However, in the 2000s, it is not as typical for A vs. B1 in comparison to the 70s/80s, where the typicality for A is stronger (compare rank 2 with a weight of 1.29 in the 70s/80s and rank 6 with a weight of 0.09 in the 2000s). The other adopted features are two related to stance occurring in document sections (*S_Introduction* and *S_Main*) as well as the it-pattern (*it_rel-v_eval-adj_target-clause*). These three features are not only similarly used between computer science and computational linguistics (as shown in Section 5.3.2.2), but are adopted from A by B1. The *P_eval-np_prep_target-np* cannot be said to have been adopted as it is already typical for B1 vs. C1 in the 70s/80s, while it is not for A vs. B1. Moreover, it is only slightly more typical for A vs. B1 in the 2000s (see Table 5.64, with a weight of -0.05 and positioned at the last rank).

In summary, we can say that computational linguistics adopts from linguistics some kind of down-toning (hedged) epistemic character with the use of the assumption and relativity meanings as early as the 70s/80s. This is even enforced over time by

typical for A vs. B1		typical for B1 vs. C1	
feature	weight	feature	weight
Sb_hedge	-2.26	**S_usefulness**	-1.45
S_Introduction	-0.99	Sb_relativity	-1.37
S_Abstract	-0.92	Sb_evidence	-1.11
P_it_rel-v_eval-adj_target-clause	-0.41	S_suitability	-0.85
S_Main	-0.31	**S_Main**	-0.82
S_usefulness	-0.09	**P_eval-np_prep_target-np**	-0.58
P_Abstract	-0.06	**S_Introduction**	-0.30
P_eval-np_prep_target-np	-0.05	**P_it_rel-v_eval-adj_target-clause**	-0.28
		S_RB	-0.07

A: computer science; B1: computational linguistics; C1: linguistics

Table 5.64: Features adopted from computer science by computational linguistics (2000s)

adopting the epistemic limitation meaning in addition to the other two down-toning meanings in the 2000s. While in the 70s/80s the importance meaning in general was adopted, in the 2000s its use at sentence beginning is adopted from C1 by B1. Thus, the trend toward putting specific kinds of evaluations at sentence beginning is also adopted in some way. In addition, the dispute meaning is adopted in the 2000s, which implies putting forward problems and issues. The relational pattern where a target is attributed with an evaluation by a relational construction (*P_target-np_rel-v_eval-expr*) is only adopted in the 2000s, even though it is typical for C1 in the 70s/80s, but not for B1. From computer science, instead, computational linguistics adopts most prominently the usefulness meaning, a tendency which is kept over time, even though in the 2000s, it is not as typical anymore for computer science. Interestingly, it is mostly algorithms that are evaluated as being useful in both registers (19.5% for A and 9% for B1). While the prepositional pattern is not quite adopted from A by B1, the it-pattern is adopted, even though the use is different in terms of meanings expressed, as shown in Section 5.3.2.2 (complexity for A but likelihood for B1).

In general, while computational linguistics adopts mostly epistemic down-toning stance meanings form linguistics, it adopts usefulness and some kinds of structural properties (patterns, use in document sections) from computer science.

typical for C2 vs. B2		typical for B2 vs. A	
feature	weight	feature	weight
S_assumption	1.59	S_importance	2.42
Sb_PP$	1.30	**S_capability**	2.16
S_PP$	1.11	S_trust	2.16
Pb_typicality	0.83	S_accuracy	2.07
S_capability	0.41	S_risk	1.48
		S_NNS	1.20
		S_assumption	1.06
		Pb_typicality	0.98
		P_to_eval-v_target-np	0.17
		S_progress	0.13

A: computer science; B2: bioinformatics; C2: biology

Table 5.65: Features adopted from biology by bioinformatics (70s/80s)

5.3.2.3.2 Bioinformatics

Features adopted from biology For the 70s/80s, Table 5.65 shows features adopted from biology by bioinformatics. Similarly to computational linguistics, bioinformatics adopts the use of the assumption meaning (*S_assumption*) from its seed register biology. Moreover, the stance meaning of capability (*S_capability*) is also adopted from C2 by B2. Both meanings have already been shown to be shared between C2 and B2 in Section 5.3.2.2, yet they are adopted from biology by bioinformatics. Additionally, the typicality meaning expressed within patterns at sentence beginning (*Pb_typicality*) is adopted by bioinformatics, which is mostly expressed either within an evaluative nominal phrase or as an adverb, as shown in Examples (71) and (72).

(71) *A* ***common*** *criticism of these histories is that* [...].
(B2: Bioinformatics 70s/80s)

(72) ***Usually****, two or more enzymes are used for these degradations.*
(B2: Bioinformatics 70s/80s)

For the 2000s (see Table 5.66), we can see that different features are adopted from biology by bioinformatics in comparison to the 70s/80s. There is no diachronic overlap of adopted features from C2 by B2. What is adopted in the 2000s is mostly the importance meaning in general (*S_importance*) but also at sentence beginning (*Sb_importance*). While at sentence beginning mostly adverbs are used (most often *notably* for C2 and *importantly* for B2), the importance meaning in general is used most prominently within the *eval-adj-target-np* pattern. Interestingly, B2 mostly

typical for C2 vs. B2		**typical for B2 vs. A**	
feature	weight	feature	weight
S_importance	1.20	Sb_opposition	2.47
P_eval-adj_target-n	1.16	P_Abstract	1.72
Sb_importance	0.51	**S_benefit**	1.47
S_benefit	0.37	S_trust	1.24
S_MD	0.20	Sb_improvement	1.17
		S_importance	0.97
		S_capability	0.79
		P_target-np_rel-v_eval-expr	0.72
		P_target-np_hedge-v_rel-v_eval-expr	0.67
		P_improvement	0.44
		Sb_importance	0.33

A: computer science; B2: bioinformatics; C2: biology

Table 5.66: Features adopted from biology by bioinformatics (2000s)

evaluates *genes* with importance, while C2 mostly uses *role* in a conventionalized way in combination with importance adjectives, such as a *major/important/essential role* (see Example (73)). The meaning of benefit (*S_benefit*) is also adopted from biology and is quite typical for bioinformatics in comparison to computer science. For both registers this meaning is mostly expressed by verbs (e.g., *facilitate*, *help*) with approx. 54% for biology and approx. 43% for bioinformatics (see Example (74)).

(73) *To date, RANKL has been shown to play* ***pivotal roles*** *in regulating various biological processes such as bone homeostasis* [...]. (C2: Biology 2000s)

(74) *Genome-wide location data can* ***help*** *us understand how an individual TF regulates its target gene.* (B2: Bioinformatics 2000s)

Features adopted from computer science Considering the features adopted from computer science by bioinformatics in the 70s/80s, the complexity meaning (*S_complexity*) is the most prominent one. From Table 5.67, we can see that it is ranked at the top for both comparisons, and thus is quite typical for A vs. B2, but also for B2 vs. C2. For both registers, things are evaluated as either difficult or simple; there seems to be no clear tendency toward the one or the other complexity pole. Additionally, in terms of meanings, the desirability meaning (*S_desirability*) is adopted. Note that in both registers, results are mostly evaluated as being good. Two other features are related to document sections, stance expressions occurring in the conclusion sections (*S_Conclusion*) as well as patterns occurring within abstracts

typical for A vs. B2		typical for B2 vs. C2	
feature	weight	feature	weight
S_complexity	-3.98	S_accuracy	-2.73
S_MD	-2.26	**S_complexity**	-2.46
S_desirability	-1.44	**S_Conclusion**	-1.96
S_PP$	-1.40	**P_Abstract**	-1.74
Sb_PP$	-1.04	S_NNS	-1.72
S_JJR	-0.52	**S_desirability**	-1.43
P_sufficiency	-0.22	P_to_eval-v_target-np	-1.41
S_Conclusion	-0.20	S_risk	-1.34
P_Abstract	-0.14	S_progress	-1.25
		S_trust	-1.20
		S_importance	-1.01
		P_sufficiency	-0.98
		S_MD	-0.44
		S_JJR	-0.29

A: computer science; B2: bioinformatics; C2: biology

Table 5.67: Features adopted from computer science by bioinformatics (70s/80s)

(*P_Abstract*). Both are relatively typical for bioinformatics. Less typical for B2 but still adopted from A are sufficiency expressed within patterns (*P_sufficiency*; see Example (75)), modal verbs (*S_MD*) and comparative adjectives (*S_JJR*) such as *better* associated with the desirability meaning.

(75) *The results indicate that **this passive mechanism is sufficient for the human kidney**.* (B2: Bioinformatics 70s/80s)

For the 2000s (see Table 5.68), the features adopted from computer science change. Self-mention at sentence beginning (*Sb_self-mention*) is adopted by computer science and is most typical for bioinformatics in comparison to biology. This feature is closely related to stance expressions within the Main part of research articles at sentence beginning (*Sb_Main*), as self-mention is the most frequently occurring stance type (approx. 72.12%). The desirability meaning expressed within patterns (*P_desirability*) is also adopted from computer science by bioinformatics. A less typical feature adopted from computer science is the to-infinitive pattern (*P_to_eval-v_target-np*).

In summary, bioinformatics shows quite different tendencies in terms of features it adopts from either biology or computer science over time. This is clearly related to its registerial shift observed in our macro-analysis (see Section 5.3.1). Thus, bioinformatics adopts the capability and assumption meaning from biology in the 70s/80s, while it adopts benefit and importance in the 2000s. From computer science, in-

typical for A vs. B2		typical for B2 vs. C2	
feature	weight	feature	weight
Sb_Main	-2.85	**Sb_self-mention**	-4.33
S_MD	-2.17	P_Abstract	-3.91
P_desirability	-1.14	Sb_improvement	-1.07
P_eval-adj_target-n	-0.51	**Sb_Main**	-0.89
P_to_eval-v_target-np	-0.28	Sb_opposition	-0.82
Sb_self-mention	-0.20	**P_desirability**	-0.73
		S_trust	-0.48
		P_to_eval-v_target-np	-0.46
		S_capability	-0.25
		P_improvement	-0.04
		P_target-np_rel-v_eval-expr	-0.04

A: computer science; B2: bioinformatics; C2: biology

Table 5.68: Features adopted from computer science by bioinformatics (2000s)

stead, bioinformatics adopts complexity and desirability in the 70s/80s, while it adopts most prominently the use of self-mention at sentence beginning in the 2000s. Furthermore, we can also observe that while it adopts mostly stance meanings from biology, it adopts structural properties from computer science, i.e. the usage of evaluative meaning expressed within patterns, and the position of expressions of evaluative meaning within a research article (document section).

5.3.2.3.3 Digital construction

Features adopted from mechanical engineering In the 70s/80s (see Table 5.69), digital construction adopts from its seed register mechanical engineering most prominently the desirability meaning expressed at sentence beginning within patterns (*Pb_desirability*). Here the adverbial pattern *eval-adv_target-clause* is used most often in combination with a negative evaluation, with the adverb *unfortunately*, at 62.5% (see Example (76)), followed by *fortunately* with 24.3% (see Example (77)). Note, however, that in mechanical engineering this meaning is mostly realized at sentence beginning by a nominal phrase (43.1%; see Example (78)) rather than by an adverb (18.8%). However, if mechanical engineering uses the evaluative adverb at sentence beginning, similarly to digital construction, the negative evaluation with *unfortunately* prevails. In addition, the benefit meaning at sentence beginning (*Sb_benefit*) is adopted by digital construction, even though it is less typical than the desirability feature. In this case, mostly advantages are put forward, for both B3 and C3 (see Example (79)).

typical for C3 vs. B3		typical for B3 vs. A	
feature	weight	feature	weight
Sb_limitation	1.49	S_usefulness	1.91
Pb_desirability	0.94	S_capability	1.87
Sb_benefit	0.41	**Pb_desirability**	1.54
		S_acceptance	1.43
		P_to_eval-adv_v_target-np	1.34
		Sb_Conclusion	1.15
		S_JJS	0.83
		Pb_target-np_rel-v_eval-expr	0.82
		T1_path	0.49
		P_progress	0.40
		Sb_typicality	0.30
		Sb_benefit	0.11

A: computer science; B3: digital construction; C3: mechanical engineering

Table 5.69: Features adopted from mechanical engineering by digital construction (70s/80s)

(76) ***Unfortunately***, *this algorithm introduces some anomalies.* (B3: Digital construction 70s/80s)

(77) ***Fortunately***, *by climbing the causality tree from the roots to the leaves, the timing data of the causing event will always be computed before that of the resulting event.* (B3: Digital construction 70s/80s)

(78) ***Good results*** *were achieved between the model and prototype for the convective process* [...]. (C3: Mechanical engineering 70s/80s)

(79) ***Advantages*** *of explicit specification are that the designer has tighter control over the design process and the compiler becomes simpler.* (B3: Digital construction 70s/80s)

In the 2000s (see Table 5.70), the importance meaning (*S_importance*) is the only feature adopted from mechanical engineering by digital construction. However, while mechanical engineering evaluates as important *parameters*, *effects* and *factors*, digital construction evaluates *issues*, *information* and *features*.

Features adopted from computer science In the 70s/80s (see Table 5.71), digital construction adopts the use of the adverbial pattern (*P_eval-adv_target-clause*) from computer science. This pattern mostly expresses desirability in digital construction (38.5%), followed by obviousness (20.8%). For computer science, however,

typical for C3 vs. B3		typical for B3 vs. A	
feature	weight	feature	weight
S_importance	2.62	S_benefit	2.04
P_it_rel-v_eval-adj_target-clause	0.67	**S_importance**	2.04
Pb_desirability	0.51	Sb_manner	1.68
S_usefulness	0.36	P_suitability	1.17
T4_length	0.19	P_benefit	1.07
Pb_to_eval-v_target-np	0.17	P_to_eval-v_target-np	0.97
Pb_likelihood	0.02	S_suggestion	0.94
		P_capability	0.80
		S_risk	0.75
		T2_method	0.71
		Pb_capability	0.46

A: computer science; B3: digital construction; C3: mechanical engineering

Table 5.70: Features adopted from mechanical engineering by digital construction (2000s)

obviousness is mostly expressed by this pattern (53.8%). Thus, B3 adopts the pattern usage from A, but the meanings expressed differ: while digital construction evaluates what follows the adverb as being unfortunate (see again Example (76)), i.e. with attitudinal meaning, computer sciences uses it to express obviousness (see Example (80)), i.e. epistemic meaning (e.g., using *clearly*). Additionally, the complexity meaning (*S_complexity*) is adopted from computer science. Note that this is a quite typical feature for digital construction, as has also been shown in Section 5.2, where it is distinctive for B3 vs. all other registers (REST) and is used to evaluate things as either simple or complex. Furthermore, the usage of the possessive pronoun (*S_PP$*) is adopted from computer science, with *our* being the most often used pronoun in both registers. While authors in digital construction refer with the possessive pronoun to *models*, *algorithms*, and *approaches*, computer science refers mostly to *results*, *proofs* and *assumptions*.

(80) ***Clearly**, the effect of these factors, especially the first and second, wears off as N increases* [...]. (A: Computer science 70s/80s)

In the 2000s (see Table 5.72), digital construction adopts the usage of plural noun targets formed of single tokens (*T1_NNS*). These nouns are mostly evaluated by importance (22.0%) in digital construction within the *eval-adj_target-np* pattern, i.e. they are evaluated by evaluative adjectives of importance. The nouns evaluated are things such as *components*, *parameters*, *relations*, *operations*, etc. In computer science, instead, the desirability meaning is mostly used to evaluate these nouns (25.2%), which mostly relate to results, i.e. *best/better results*.

typical for A vs. B3		typical for B3 vs. C3	
feature	weight	feature	weight
S_PP$	-2.84	P_progress	-3.23
S_complexity	-2.77	T1_path	-2.19
P_eval-adv_target-clause	-0.79	Sb_Conclusion	-1.66
		S_capability	-1.58
		Sb_typicality	-1.26
		S_usefulness	-0.74
		Pb_target-np_rel-v_eval-expr	-0.59
		P_to_eval-adv_v_target-np	-0.55
		P_eval-adv_target-clause	-0.43
		S_complexity	-0.39
		S_PP$	-0.38
		S_JJS	-0.38
		S_acceptance	-0.15

A: computer science; B3: digital construction; C3: mechanical engineering

Table 5.71: Features adopted from computer science by digital construction (70s/80s)

typical for A vs. B3		typical for B3 vs. C3	
feature	weight	feature	weight
P_it_rel-v_eval-adj_target-clause	-3.16	**T1_NNS**	-2.48
Pb_desirability	-1.01	Sb_manner	-2.40
Pb_to_eval-v_target-np	-0.69	S_risk	-2.17
T1_NNS	-0.66	T2_method	-1.81
S_usefulness	-0.40	P_to_eval-v_target-np	-1.60
T4_length	-0.30	P_suitability	-1.15
Pb_likelihood	-0.19	P_capability	-0.72
		S_benefit	-0.66
		P_benefit	-0.62
		S_suggestion	-0.38
		Pb_capability	-0.31

A: computer science; B3: digital construction; C3: mechanical engineering

Table 5.72: Features adopted from computer science by digital construction (2000s)

typical for C4 vs. B4		**typical for B4 vs. A**	
feature	weight	feature	weight
P_eval-adj_target-n	2.21	S_improvement	2.36
S_PP$	0.52	S_trust	2.07
S_relativity	0.33	S_benefit	2.01
P_target_eval-v	0.24	S_prevention	1.21
S_importance	0.18	P_capability	1.17
		Sb_prediction	0.91
		Sb_suggestion	0.61
		S_relativity	0.28
		S_acceptance	0.26
		S_importance	0.13
		P_target_v_eval-expr	0.11

A: computer science; B4: microelectronics; C4: electrical engineering

Table 5.73: Features adopted from electrical engineering by microelectronics (70s/80s)

In summary, digital construction mostly adopts meanings from its seed register mechanical engineering. In the 70s/80s, the pattern usage of desirability at sentence beginning is adopted, where things are evaluated as unfortunate, as well as the usage of the benefit meaning at sentence beginning, where advantages are put forward. In the 2000s, the importance meaning is adopted mostly to highlight issues of importance. From computer science, however, digital construction adopts the pattern usage of evaluative adverbs in the 70s/80s. Interestingly, while it adopts the desirability meaning from mechanical engineering with the use of the adverb *unfortunately*, it uses this meaning within a pattern adopted from computer science, i.e. in the *eval-adv_target-clause* pattern. This kind of adoption from each of the seed registers forming the contact register's own usage is also reflected in the 2000s. While the importance meaning is adopted from mechanical engineering, targets constituted by plural single token nouns, a feature adopted from computer science, are evaluated with this meaning.

5.3.2.3.4 Microelectronics

Features adopted from electrical engineering Similarly to digital construction, in the 70s/80s microelectronics adopts from its seed register electrical engineering stance meanings: relativity and importance (see Table 5.73). For the relativity meaning, microelectronics most often uses the adverbs *approximately* (26.3%; see Example (81)) and *relatively* (19.1%; see Example (82)), while electrical engineering uses in addition to *approximately* (20.4%), also *rather* (19.3%; see Example (83)).

typical for C4 vs. B4		typical for B4 vs. A	
feature	weight	feature	weight
P_to_eval-v_target-np	0.26	S_judgement	1.11
S_typicality	0.16	S_risk	0.52
P_Abstract	0.04	S_benefit	0.51
S_idiosyncrasy	0.04	T7_length	0.29
		S_idiosyncrasy	0.21
		Sb_sufficiency	0.18
		Pb_capability	0.18
		S_sufficiency	0.12
		T1_path	0.07

A: computer science; B4: microelectronics; C4: electrical engineering

Table 5.74: Features adopted from electrical engineering by microelectronics (2000s)

Considering the importance meaning, the most often occurring lemma in microelectronics is *important* (12.3%), used mostly within the *eval-adj_target-np* pattern in which targets such as *parameter*, *factor* and *feature* are evaluated (see Example (84)), while electrical engineering uses *necessary* (15.7%), mostly within the *it_rel-v_eval-adj_target-clause* pattern, evaluating mostly mental (e.g., *consider*) and material processes (e.g., *use*) in a to-infinitive clause (see Example (85)).

(81) *For a given implant species, these effects all become significant at **approximately** the same dose* [...]. (B4: microelectronics 70s/80s)

(82) *For a resist to have high sensitivity, the volume of resist modified by each incident particle must be **relatively** large and the resolution of the resist **relatively** low.* (B4: microelectronics 70s/80s)

(83) *Although the expressions are **rather** complex, good results are obtained.* (C4: electrical engineering 70s/80s)

(84) *This paper outlines the factors that require control and presents test data relating to the **important parameters**.* (B4: microelectronics 70s/80s)

(85) *Finally, **it is necessary to consider** the possibility of random frequency variations of the modulating signal.* (C4: electrical engineering 70s/80s)

In the 2000s (see Table 5.74), microelectronics adopts from electrical engineering only the idiosyncrasy meaning. In this case, B4 uses the adverbs *especially* and *specifically* most often (see Example (86)), similarly to electrical engineering.

typical for A vs. B4		**typical for B4 vs. C4**	
feature	weight	feature	weight
P_eval-adj_target-n	-3.76	P_capability	-2.26
S_PP$	-3.39	**P_Introduction**	-1.67
P_target_eval-v	-0.43	S_improvement	-1.64
P_Introduction	-0.14	S_benefit	-1.64
		S_trust	-1.40
		Sb_prediction	-1.00
		Sb_suggestion	-0.96
		S_acceptance	-0.82
		S_prevention	-0.62
		P_target_v_eval-expr	-0.18

A: computer science; B4: microelectronics; C4: electrical engineering

Table 5.75: Features adopted from computer science by microelectronics (70s/80s)

(86) *We also observe that double-bit errors also occur, **especially** in the case of drowsy cache lines that operate at a lower voltage.*
(B4: microelectronics 2000s)

Features adopted from computer science In the 70s/80s, microelectronics adopts from computer science the use of patterns within the Introduction of research articles (see Table 5.75). Both registers most often use the *eval-adj_target-np* pattern, followed by the *it_rel-v_eval-adj_target-clause* pattern; both also use the meanings of importance and desirability most frequently within the Introduction.

In the 2000s (see Table 5.76), microelectronics adopts more features from computer science, which are not related to meanings: one target feature (*T1_NP*), self-mention, and one feature related to document structure (*S_Conclusion*). Most prominently, the single-token targets constituted of proper nouns are adopted (e.g., *the best MLV*), which have shown to be quite typical for B4 vs. all other registers in our diversification analysis (see Section 5.2). Self-mention is also adopted from computer science, with *we* and *our* being the most frequently used pronouns. Again, this feature is quite typical for B4, as we have seen in Section 5.2. However, A and B4 differ in the distribution of these two pronouns, compare 85.6% for *we* and 11.2 for *our* in computer science with 74.6% for *we* and 24.0% for *our* in microelectronics, i.e. B4 makes more use of the pronoun *our*. Nevertheless, both most frequently use communication and mental verbal processes with *we*, and computational concepts with *our* (e.g., *algorithm*, *result*). Additionally, the use of stance expressions within the Conclusion section is adopted from computer science. Again, the adopted use is quite similar, as the likelihood meaning is the most often used meaning in the Conclusion section for both B4 and A.

typical for A vs. B4		typical for B4 vs. C4	
feature	weight	feature	weight
S_self-mention	-2.66	Sb_sufficiency	-4.09
S_Conclusion	-0.60	T1_path	-2.56
S_typicality	-0.51	T7_length	-1.68
P_to_eval-v_target-np	-0.46	**T1_NP**	-0.94
T1_NP	-0.25	S_risk	-0.45
P_Abstract	-0.24	**S_self-mention**	-0.19
		Pb_capability	-0.17
		S_sufficiency	-0.07
		S_benefit	-0.04
		S_Conclusion	-0.02
		S_judgement	-0.01

A: computer science; B4: microelectronics; C4: electrical engineering

Table 5.76: Features adopted from computer science by microelectronics (2000s)

In summary, microelectronics adopts, similarly to digital construction, stance meanings from its seed register electrical engineering, while it adopts features related to document section, stance type and targets from computer science. Furthermore, B4 shows some differences when comparing the realizations of the meanings with electrical engineering, while it is quite similar to computer science in the realizations of the features adopted.

5.3.3 Summary and conclusions on registerial imprint

In terms of registerial imprint, we wanted to test whether contact registers reflect some linguistic characteristics of their seed registers, i.e. we looked into the degree of registerial imprint on the contact registers by their seed registers as well as on the kind of registerial imprint.

Degree of registerial imprint By inspecting the contact register in SciTex in terms of a macro-analysis, we have seen that some registers seem to be in between their seed registers, showing no clear tendency toward the one or the other in both time periods, as for computational linguistics (B1), which shows a quite similar amount of overlap with its two seed registers in the 70s/80s and the 2000s. Other contact registers, however, show a clear tendency toward one of the two seed registers over time. This is the case for the two engineering contact registers digital construction (B3) and microelectronics (B4), which tend to their seed registers mechanical engineering (C3) and microelectronics (C4), respectively. The contact register of bioinformatics (B2), however, has exhibited a shift in terms of registerial imprint.

While it had more overlap with biology (C2) in the 70s/80s, it shifted to have an almost equal amount of overlap between computer science (A) and biology (C2), with a slight preference for computer science (A). In a first micro-analytical step, these tendencies have been also observed in terms of the number of features shared between a given contact and its seed registers.

Kind of registerial imprint The first feature analysis on the micro level has shown as an example which concrete features out of the top 20 most distinctive features are shared between two contact registers (computational linguistics and bioinformatics) and their seed registers. This allowed us to inspect first trends in terms of the kind of registerial imprint. However, as we were interested in whether the contact registers not only share but have adopted the features from their seed registers, with an additional classification and feature analysis of the top 10 distinctive features for the contact registers vs. computer science and the other seed register we were able to see which features are adopted from the seed registers by the contact register. It has been shown that the contact registers differ in terms of what they adopt from computer science or their other seed register.
For computational linguistics, we have seen that it adopts epistemic, rather downtoning evaluative meaning from linguistics, while it uses structural properties similarly to computer science.
Bioinformatics adopts different meanings from biology, with some change over time, while it moves toward adopting fewer stance meaning related features from computer science in the 2000s and instead more structural features, such as evaluative meaning expressed within patterns and document sections.
Digital construction, quite interestingly, adopts meanings from mechanical engineering, which it uses within patterns adopted from computer science.
Finally, microelectronics also adopts particular meanings from electrical engineering and more structural features from computer science such as the use of evaluative meaning in document sections.

From these observations, we can also deduce some general trends. One is that most contact registers adopt the importance meaning from their C seed registers (B1, B2 and B3). More interestingly, however, we observe that while the contact registers seem to adopt from their C seed register mostly epistemic and attitudinal meanings, they adopt more structural properties from computer science, such as patterns, the use of evaluative meaning in document sections, and stance type as well as parts-of-speech of stance expressions and targets.

Chapter 6

Summary and conclusions

6.1 Summary

In this thesis, we considered two main hypotheses: register diversification and registerial imprint.
In terms of register diversification, we aimed at providing answers to whether disciplines show evaluative characteristics of their own reflected linguistically in registers and whether these characteristics have changed over time. In particular, we have considered disciplines that recently emerged out of contact between two other disciplines, i.e. contact disciplines such as computational linguistics emerged from contact between computer science and linguistics, as these disciplines are prone to having a greater incentive toward diversification to become registers of their own.
In terms of registerial imprint, we pursued answers for whether newly emerged disciplines show a linguistic imprint of evaluative meaning that is reflected in their register and that they have adopted from their disciplines of origin (seed disciplines), and additionally whether the registerial imprint has changed over time.
To approach these hypotheses, we have considered several theoretical and methodological concepts.

In theoretical terms, we have relied on the linguistic theory of Systemic Functional Linguistics (SFL), as it provides a basis for analyzing evaluative meaning linguistically, adopting a sociosemiotic perspective. This is necessary, since we consider scientific disciplines positioned on a socio-cultural stratum, as well as their semantic and linguistic choices in terms of evaluative meaning positioned on the strata of semantics and lexico-grammar (Section 3.2). Based on this sociosemiotic interpretation of language, we have created a model of analysis for evaluative meaning, where the semantic choices available for the expression of evaluative meaning are realized in the lexico-grammar (Section 3.4). Thus, by inspecting the lexico-grammatical choices, one can trace the semantic choices made. In particular, we investigated the choices made in different disciplines by looking at particular lexico-grammatical features that realize evaluative meaning. By comparing the lexico-grammatical choices

made across different disciplines, we were able to inspect differences and similarities between disciplines in terms of register diversification and registerial imprint.
The notion of *register* is a central concept in this study, as it provides the ground for our analysis of functional variation across highly specialized scientific registers, i.e. variation according to situational context (Section 2.3). In our case, we have looked at nine contexts of situation reflected in the nine disciplines we investigated: computer science, computational linguistics, linguistics, bioinformatics, biology, digital construction, mechanical engineering, microelectronics, and electrical engineering (Section 4.3). Furthermore, due to modernization processes the need arises for particular disciplines to adopt concepts, techniques etc. from other disciplines. This gives rise to the formation of new registers. In our case, we have looked at four such newly emerged disciplines, with the aim to investigate whether they have formed registers of their own over the time span of approx. thirty years. Here, we have particularly focused on the formation of registerial differences in terms of evaluative meaning.
Evaluative meaning can be positioned within the participatory function of language, the interpersonal metafunction within SFL, as it provides the resources to assign discursive roles to participants and to express evaluations and attitudes (Section 3.3). Our theoretical approach to evaluative meaning differs from others, as it tries to consider the full meaning potential of evaluative meaning, which should account for the participants involved, possible choices the participants can make from the options given to express evaluative meaning, and the target the evaluative expression is directed toward. For this, we presented a semantic system of choices of evaluative meaning that allowed us to trace the semantic choices made by looking at the lexico-grammatical features that realize these choices (Section 3.4). This enabled us to trace relatively fine-grained registerial differences across highly specialized disciplines in terms of evaluative meaning.

In methodological terms, we have adopted a macro- and micro-analytical corpus-based approach, designing an analytical cycle that allows for recursive quantitative as well as qualitative analytical steps (Section 4.4). The corpus we used (SciTex) permitted us to inspect several scientific disciplines diachronically, and the iterative process enabled us to find the most appropriate balance between generalizable trends and fine-grained insights on register diversification and registerial imprint based on empirical findings. Our methodological approach differs from previous work, as it is not biased toward one extreme or the other (i.e. detailed analyses vs. aiming at generalization); it takes into account both perspectives, considering how much detail is needed to make the most appropriate generalizations by combining macro- and micro-analysis.
On the level of macro-analysis, we used text classification with Support Vector Machines (SVM) as a technique to empirically observe how well text productions of one discipline can be correctly identified as texts of that particular discipline in terms of evaluative meaning. This gave insights into (1) the degree of register diversification, i.e. how well each discipline is registerially differentiated from the other disciplines

and whether this has changed over time (Section 5.2.1), and (2) the degree of registerial imprint, i.e. how strongly the contact disciplines show registerial overlaps with their seed disciplines, and again, whether this changes over time (Section 5.3.1). Given that a process of diversification is a process inherently associated with the temporal dimension, i.e. one which develops over time, a diachronic comparison was crucial to gain insights into register diversification processes. Additionally, in terms of registerial imprint, as the contact disciplines undergo a process of diversification, we had to consider the diachronic perspective as well to capture possible changes related to overlaps with one or the other seed discipline.
On the level of micro-analysis, we employed feature analysis by observing (1) the kind of register diversification, i.e. looking at features characteristic of a given register (Section 5.2.2), and (2) the kind of registerial imprint, i.e. looking at features adopted by the contact registers from the seed registers (Section 5.3.2). In our feature analysis, we moved from a more general to a more specific view, i.e. we considered the contribution of feature types, the contribution of individual features as well as the usage of features within concordance lines and their distributional information on different levels. Here, we inspected the top-ranking features of the support vector machine output. Furthermore, taking a detailed micro-analytical perspective allowed us to incorporate a step within our analytical cycle that enhanced data quality in a manageable way, as we were able to perform detailed cleaning procedures on extraction results for top-ranking features (by SVM weights).

The analysis of register diversification in terms of evaluative meaning showed that with respect to the degree of diversification, contact as well as seed registers have undergone a process of diversification over the 30-year time span investigated. However, the contact registers went through a much more pronounced diversification process than their seed registers. In the 70s/80s, the contact registers were not distinguishable from other registers by evaluative meaning, i.e. in that time period they do not have characteristics of their own regarding the expression of evaluative meaning. In the 2000s, in contrast, they are clearly distinguished from the other registers, i.e. over a time period of 30 years or so, they have developed their own usage of evaluative meaning (Section 5.2.1).
Considering the kind of register diversification, even though the contact registers did not have characteristics of their own in the 70s/80s, the analysis of the feature types and subtypes revealed some general diachronic trends in both contact and seed registers: while in the 70s/80s the contact registers could barely be distinguished by individual stance meanings, in the 2000s more feature subtypes (parts-of-speech of stance expressions, stance meanings, stance types, etc.) contributed to a distinction. This is also reflected in the seed registers, even though this tendency was less pronounced, except in mechanical engineering. This variation among the feature subtypes is related to the fact that an evaluative phenomenon is reflected across different feature (sub)types, i.e. the more characteristic an evaluative phenomenon becomes, the more it will be reflected in different types. One example is the complexity meaning in computer science. In the 70s/80s, it is characteristic in terms of

one feature only, while in the 2000s, it seems to become even more typical or more conventionalized as it is reflected in three different features (the complexity meaning itself, its expression within patterns at sentence beginning and as attitude markers at sentence beginning).

A further general diachronic tendency is the relocation of evaluative meaning toward sentence beginnings. Again, this has been observed for both contact and seed registers, pointing toward a more conventionalized usage of evaluative meaning.

In terms of specific characteristics across contact and seed registers, we have seen how they clearly differ from one another: computer science is positioned more toward boosting expressions in terms of epistemic meaning than linguistics, which has clearly shown a rather hedged evaluative character. Biology, however, tends toward using both boosters and hedges, with a strong focus on evaluating experimental findings. The contact registers differ especially in the stance meanings they use. Computational linguistics is distinguished by attitudinal meanings both at sentence beginning and in general, while bioinformatics is characterized by both epistemic as well as attitudinal expressions. Digital construction has a quite conventionalized character as it is mostly distinguished by expressing evaluations within pattern constructions, yet it also uses distinctively attitudinal as well as hedging expressions. Microelectronics is mostly characterized by attitudinal meanings expressed within the Conclusion section.

Furthermore, some features, while clearly characteristic for a register, also seem to be shared across registers. The contact registers, e.g., all distinctively use self-mention, which is also quite characteristic for computer science, even though they differ in the process types (what they do) and concepts (what they possess) used with self-mention, or in which section it most prominently appears. This is also somewhat reflected when considering biology and bioinformatics, as both tend to use boosting as well as hedging expressions, which is not the case for the other registers (Section 5.2.2). This observation led us to investigate our second main hypothesis of registerial imprint, as it seemed that the seed registers had in some way left an imprint on the contact registers.

The analysis of registerial imprint showed different tendencies in terms of the degree of registerial imprint, i.e. the amount of overlap between contact and seed registers. Some contact registers appear to be located in between their seed registers (computational linguistics), a tendency that is preserved over time. The engineering contact registers (digital construction and microelectronics) show a clear tendency toward their engineering seed registers (mechanical engineering and electrical engineering, respectively). Bioinformatics, in contrast, seems to have performed a shift, showing more overlap with biology in the 70s/80s, and more with computer science in the 2000s (Section 5.3.1).

By inspecting the kind of registerial imprint, we obtained quite interesting results on which features the contact registers adopt from which of the two seed registers (i.e. either from computer science or the other seed register). Computational linguistics adopts the hedged character from linguistics, while it adopts not only the use of evaluative patterns but also where to position evaluative meaning within a research article (i.e. the document section in which to express evaluative meaning) from computer science. Similarly, bioinformatics adopts different attitudinal and epistemic meanings from biology, while it adopts the usage of evaluative meaning in specific document sections from computer science, but also pattern usage and the complexity meaning (which is quite characteristic for computer science, as shown in our analysis of register diversification). Digital construction, most interestingly, adopts specific attitudinal meanings from mechanical engineering and uses these meanings within patterns adopted from computer science. Microelectronics, similarly to the other contact registers, adopts attitudinal meanings from electrical engineering, while it adopts structural properties from computer science. Thus, we observed a general tendency of adopting epistemic or attitudinal meanings from the seed registers and more structural properties from computer science (Section 5.3.2).

By inspecting hypotheses of diversification and imprint on the macro and micro levels, we have gained a quite deep understanding of how scientific disciplines vary in their use of evaluative meaning, the degree to which they vary and the kinds of differences and similarities they show in terms of lexico-grammatical features. The study has also shown that on the one hand, disciplines create registers of their own over time in terms of evaluative meaning, i.e. they undergo a diversification process, while on the other hand, (and this relates specifically to the contact disciplines), they still show an imprint left over or adopted from the seed disciplines that can be traced in the lexico-grammar.

6.2 Assessment of the approach

Our approach to the investigation of evaluative meaning across scientific writing comes with several general and beneficial properties related to theory and methodology.

6.2.1 An approach grounded in linguistic theory

The approach relies on several sources of knowledge on evaluative meaning and scientific writing, yet it is grounded in one theory of language that accounts for both: Systemic Functional Linguistics (SFL), which makes it possible to adopt a sociosemiotic perspective on language use. The commitment to the linguistic theory of SFL has been made as it has at its core a number of concepts extremely relevant and useful for this study.

First, its functional perspective on language allowed us to position evaluative meaning within a linguistic theory. It has been proposed that in terms of the semantic system, evaluative meaning is part of the interpersonal metafunction, i.e. the participatory function of language that provides resources for participants to enact a social and intersubjective relationship. The most important aspect in favor of positioning evaluative meaning within this metafunction is the potential for arguing inherent to evaluative as well as interpersonal acts (though interpersonal acts involve more than just arguing). What is argued about is the information exchanged between participants (rather than the exchange of goods-&-services).

Second, SFL allows one to trace choices made in the semantic system from the meaning potential by considering choices on the level of lexico-grammar that can be identified in the text. In our specific case, choices on the semantic level are taken from the potential of evaluative meaning to instantiate an evaluative text. These choices are reflected in the lexico-grammatical system, i.e. specific lexico-grammatical features are available to language users for performing an evaluative act instantiated in the wordings used. Therefore, by inspecting lexico-grammatical choices, one can trace the choices made in the semantic system in terms of evaluative meaning.

A third concept of SFL that is extremely relevant to our study is the notion of register, concerned with functional variation, i.e. variation according to language use. Given that in our case writers of scientific research articles belong to different social contexts according to the different disciplines, they have specific options available in the lexico-grammatical system to express evaluative meaning which may differ across disciplines. Differences in the constellation of lexico-grammatical features give rise to different registers. This can then also change diachronically as new registers develop. Thus, the choices made on the stratum of lexico-grammar define the social context and vice versa, but can change over time. By tracing the lexico-grammatical choices made in two different time periods, we were able to look at the diachronic formation of new registers in terms of evaluative meaning.

This also implies a further benefit of considering SFL as a theoretical basis: comparativeness, i.e. the usage of different constellations of lexico-grammatical features can be compared across different social contexts but also across different time periods. In our study, we considered both comparative dimensions.

Furthermore, the most important aspect that a theoretically-grounded approach offers is that it allows for the interpretation of the findings, i.e. if explanations are sought, as in our case, a theory is needed that allows for the kinds of interpretations pursued in the study. In the present context, we tested whether disciplines show diversification trends that are reflected in their registers, for which SFL was perfectly suited as it allowed us to interpret the findings obtained from the comparison of lexico-grammatical features on the semantic level as well as on the level of the social context. Moreover, the delicacy of focus that the notion of register permits was also crucial to inspect fine-grained differences in terms of register diversification and to consider a possible imprint left over in the contact registers by the seed registers in terms of registerial imprint.

In general terms, a theory has to be chosen that allows for the appropriate level of interpretation aimed at the hypotheses for which one pursues answers. As we were interested in functional variation, SFL was the optimal choice. However, especially in relation to the diachronic perspective, limitations in terms of interpretations were encountered as to why specific diachronic changes or shifts have taken place for some of the disciplines. Such interpretations would have to rely more on a theory geared toward historical linguistics rather than sociolinguistics.

6.2.2 An approach combining macro- and micro-analysis with comparative corpus-based methods

The macro- and micro-analytical perspectives adopted in our approach have proved to be quite valuable for testing our hypotheses. While macro-analysis mostly gives insights into the degree of difference or similarity, micro-analysis is geared toward unfolding what the differences or similarities are, i.e. it relates to the kind of differences or similarities observed. Just one of the two perspectives alone would not be sufficient to answer our hypotheses in a satisfactory manner since we are aiming at a comprehensive account.

If we had considered only macro-analysis, we clearly would have observed that registers undergo a process of diversification over time, as the classification performance drastically improves in the second time period investigated, especially for the contact registers. However, we would not have been able to see which characteristics associated with evaluative meaning contributed to a better distinction over time. Moreover, we would have missed the bigger picture. For example, in terms of register diversification, we would have missed observing the tendency toward a more conventionalized usage of evaluative meaning (as it is predominately used at sentence beginning in the 2000s). In terms of registerial imprint, e.g., on the one hand, we would have not observed the preference of the contact registers for the adoption of stance meanings from the seed registers of the C subcorpus, and on the other hand, the preference for adopting structural properties from computer science would have been missed.

If we had considered only micro-analysis, the detection of particular characteristic features would have been a challenge not feasible to undertake due to time constraints and available resources, considering the different dimensions of comparisons made (registers, time, triples) and the amount of features analyzed (approx. 270). For this, text classification proved to be quite useful as it allowed us to consider, on the macro level, the degree of register diversification and registerial imprint, and on the micro level the kind of diversification or imprint by feature analysis of top-ranking features. Thus, the macro-analysis did not allow only for generalizable findings, but also permitted us to limit the micro-analysis to features characteristic of a given register, i.e. to perform a feature analysis on relevant features. Additionally, the relevant features could be analyzed most carefully, giving a rather detailed account of their usage within a given register by means of a closer inspection and using distributional information across different levels (registers, document sections, evaluative patterns, stance meanings, stance types, parts-of-speech, etc.).

Furthermore, the implementation of an iterative process allowed us to enhance data quality in recursive steps in a cost-effective manner, i.e. with a special focus on relevant features. More specifically, while macro-analysis provided us with a ranking of lexico-grammatical features, in micro-analysis the top-ranking features (i.e. the most characteristic ones) were inspected in detail. Due to non-evaluative discipline-specific usages of possible evaluative lexical items, we were able to exclude non-evaluative usages by taking this detailed micro-analytical view of the data. This, however, also points to some limitations of the approach, as it is mostly geared toward precision rather than recall, i.e. we might have still missed some amount of evaluative meaning, which might then not be represented among the top-ranking features that have been shown to be characteristic of a register. Nevertheless, the insights gained on characteristic and adopted features can still be said to be representative, especially regarding the account of precision applied.

In terms of methods, we also had to consider appropriate dimensions of comparisons that had to be considered to test our two main hypotheses. The registerial dimension was one of the dimensions most important to our study for both hypotheses, yet the delicacy of focus had to be adapted accordingly.
To test for register diversification, we considered a given register against all other registers taken together. This allowed us to obtain distinctive features characteristic of a given register. However, if the classification performance of a register was too low in terms of F-measure (e.g., as for computational linguistics, which achieved only around 20% in the 70s/80s), we did not consider the features to be characteristic of that register.
For registerial imprint, we considered a contact register against each seed registers. This allowed us to inspect features characteristic of a contact register when compared to its seed registers. Moreover, we did an additional classification to inspect whether the features that turned out to be characteristic of a given contact register were adopted from computer science or the other seed register. This was quite insightful. If some features were characteristic for computer science when compared to one of the contact registers, while they were characteristic of the contact register when compared to its other seed register, the features were adopted from computer science by the contact register. If, instead, some features were characteristic of the other seed register in comparison to the contact register, while they were characteristic of the contact register when compared to computer science, the features were adopted from the other seed register by the contact register. This is something we have also observed in other studies on SciTex which looked at other feature sets (Degaetano-Ortlieb et al., 2014a).

In summary, the approach presented in this thesis draws its strength from the fact that it is grounded in linguistic theory. This was extremely useful in defining and testing hypotheses as well as for the interpretation of results. Moreover, by combining macro- and micro-analysis and corpus-based methods, we were able to perform an empirical analysis which allowed making generalizations as well as detailed observations.

6.3 Envoi and future work

Since the methodology designed in this study is not confined to a linguistic phenomenon but is quite general in nature, it can be adopted in other contexts that seek to combine an empirical analysis with a more detailed account on the findings, allowing interpretable results, and in other scenarios in which different dimensions of contrast are of interest.

6.3.1 Considerations on feature selection and interpretation

One of the biggest issues in analyzing linguistic phenomena is which linguistic features to look at (*feature selection*). This is a question that is always quite hard to answer in a comprehensive manner and for which we do not have a ready-made solution. Some features can be derived from previous studies, others from introspection, and others from observation. But even then, one does not know whether the features chosen appropriately reflect the phenomenon to be analyzed for the population under investigation. What we proposed here is to use a quite wide range of features taken from previous studies and our own observations and to decide based on a macro-analysis which features to investigate further. Thus, one can decide which features to analyze in detail by how well they represent the population investigated, in our case each register according to the ranking results of the classification. This method can be used in a variety of comparative studies, such as comparative studies of registers or time periods as in our case, but also comparisons of modes of discourse (e.g., spoken vs. written), different languages, originals and translations, translation modes (e.g., human vs. automatic), etc.
What is also important in terms of feature selection is which kinds of features to consider. Using shallow features (bag-of-words) for classification tasks, such as document classification or in stylometric studies, has been shown to be quite efficient (cf. Joachims (1998); Koppel et al. (2002); Rybicki (2006); Argamon et al. (2008); Fox et al. (2012)). However, if a deeper understanding of the mechanisms involved in a particular variation is sought, features are needed that reflect these mechanisms. Thus, what one has to be aware of is that using relatively shallow features (such as word/lexical items or function words) to gain insights into the distinctness between registers, documents, etc., will also be quite shallow in terms of the insights the features can offer. Feature sets made up of bags-of-words can only show something about the differences in topics. In more recent studies on diversification trends related to registers, for example, features are taken that reflect possible diversification processes (cf. Argamon et al. (2007); Teich et al. (2013); Degaetano-Ortlieb et al.

(2014a)) driven by linguistic theory, which allows for an appropriate interpretation of the findings. Thus, besides using a quite varied range of features, it is also important to select features that might be relevant for the task or the insights sought.

6.3.2 Further contexts of application

6.3.2.1 Other languages

One possible context closely related to ours would be to apply the approach on other languages, i.e. analyze the expression of evaluative meaning in other languages.
Here, what can be adopted is the analytical cycle (Section 4.4) as well as the model of analysis proposed for how evaluative meaning is construed semantically (Section 3.4). As we have seen in the previous section, the analytical cycle can be adopted for a variety of analyses. One would just have to define the levels of macro- and micro-analysis needed. As for the model of analysis of evaluative meaning, it can be said to be quite culture- and language-independent, as it is related to human interaction, i.e. within an evaluative act there will always be a speaker or writer who performs the act and a hearer or reader who perceives the act, as well as the evaluation itself and the target evaluated.
What would have to be adapted is the realization of the semantic meaning in the lexico-grammar. One possible option that might still profit from the insights gained in our study on realizations on the level of lexico-grammar is to consider the context of translations, i.e. annotate the English part of a parallel corpus with the annotation procedure designed for this study to inspect differences and similarities in the lexico-grammatical system of expressing evaluative meaning between English and some other language. Clearly, the translation approach would also bear some issues, which then would have to be dealt with (e.g., correct word and phrase alignment to detect lexical realizations across languages, possible omissions and explicitations in the translation, etc.). Nevertheless, it would give some basic insights into what would have to be changed in terms of lexico-grammatical features when considering another language. However, one would also have to consider that in our study, even though the model of analysis is language- and register-independent, the lexico-grammatical realizations found by the annotation procedure might be relatively register dependent, as they were designed on the basis of scientific writing. Clearly, the more distant a register to be investigated is from written academic English, the more adaptations will be needed. Nevertheless, apart from the challenges to face, it would be quite interesting to investigate how languages differ in terms of evaluative meaning.

6.3.2.2 Inform classification tasks

A further context of application is to inform other classification tasks with more linguistically driven features.

Considering the notion of register and given that the features analyzed in this thesis are located within the tenor of discourse, they could inform approaches that rely solely on features from the field of discourse, such as bag-of-words approaches which are mostly related to topicality. In terms of SFL, this would lead to a more comprehensive account of the data. To show how this might improve the performance of a classifier, we have performed a small pilot study on the SciTex disciplines for both time periods. Here, we carried out a bag-of-words classification with SVM based on the 500 most distinctive nouns selected by Information Gain on the nine classes of the SciTex corpus, i.e. texts in SciTex had to be classified into one of the nine registers based on the 500 most distinctive nouns used in the corpus. Classification by these shallow features, i.e. in a bag-of-words fashion, achieves relatively high F-measures for both time periods (0.898 for the 70s/80s and 0.930 for the 2000s; see Table 6.1). Considering features of evaluative meaning on their own, for the 70s/80s classification performance is quite low (F-measure of 0.662), while in the 2000s it is competitive with the shallow feature classification (0.930 for both). By combining the feature sets, classification can be improved up to an F-measure of 0.927 for the 70s/80s and up to 0.941 in the 2000s. Clearly, the fact that the classification performance for the 70s/80s on SciTex regarding features of evaluative meaning is quite low is related to the contact disciplines. As we have shown in this study, the contact disciplines do not have a distinctive evaluative character in the 70s/80s, and thus cannot be distinguished by it, while they can already be distinguished much better in terms of the 500 most distinctive nouns, i.e. by topics. Nevertheless, this small study already shows how more linguistically driven features might improve classification performance.

6.3.2.3 Inform feature-based sentiment analysis

As we have seen in Section 2.4.3.1.2, while the approach of feature-based sentiment analysis accounts for the assignment of opinions to the target (object) that is evaluated, which from all computationally based approaches shows the greatest resemblance to our approach, its focus lies on determining the polarity of the evaluative expression rather than the meaning of the evaluation. Thus, the approach

	feature set	**accuracy**	**precision**	**recall**	**F-measure**
	shallow	89.82	0.900	0.898	0.898
70s/80s	evaluation	66.32	0.663	0.663	0.662
	combined	92.70	0.927	0.927	0.927
	shallow	92.92	0.931	0.929	0.930
2000s	evaluation	92.93	0.931	0.929	0.930
	combined	94.10	0.941	0.941	0.941

Table 6.1: Classification accuracies obtained by shallow and evaluative feature sets on SciTex

of defining more fine-grained classes of stance meanings might enlarge the scope of applications of this sentiment analysis approach, as it would not be limited to determining the polarity alone, but polarity related to some kind of meaning. This would be a similar task to what at the moment is known as emotion detection (Strapparava and Mihalcea, 2008, 2010; Lei et al., 2014). However, emotions and stance meanings are not exactly the same, even though they may have overlaps. If we consider the stance meaning of importance, this would not be understood as an emotion (it is an evaluation), while frustration is clearly associated with an emotion. Thus, there is more to be accounted for with respect to evaluative meaning, i.e. while it is useful to detect emotions in text, facial emotion, or emotion expressed multi-modally, it is also relevant to consider other kinds of epistemic and attitudinal meanings such as assumption, suggestion, importance, complexity, etc. It depends on the task of what kind of knowledge is being sought, i.e. how products are reviewed (positively vs. negatively), which emotions are triggered for example by political actions on twitter feeds (anger, fear, happiness, etc.), how concepts are evaluated in science (attitudinal/epistemic, important/unimportant, etc.). However, it is quite a challenge to define epistemic and attitudinal meanings in a comprehensive manner, especially for different situational contexts, registers, etc. Thus, as meaning is strongly associated with the context, methods such as topic models (cf. Hofmann (2001); Blei et al. (2003); Blei (2012)) might be quite useful in this endeavor.

6.3.3 Future work on the subject

Apart from determining the meaning of the evaluation, the polarity of the evaluative expression could be taken into account as an additional feature that might be differently used across disciplines. Thus, even though two disciplines might use the complexity meaning similarly often, they might differ in whether they use it more toward the easiness pole or the difficulty pole. We have deliberately not chosen the wordings *positive* and *negative* as this polarity categorization might be too restrictive and does not always apply to all stance meanings. Consider again the importance meaning: if something is not important it is not strictly related to being negative. Thus, in the case of stance meanings, it might be more appropriate to use positive and negative as two poles with no concrete positive or negative connotation at first (similarly to the values of modality; see also Section 3.3) and to determine the connotation in an additional step. Defining whether something is positioned on one or the other extreme of a pole for a given meaning could be accounted for automatically or semi-automatically on the basis of the lexico-grammatical realizations (lexical items, negation, etc.). Determining whether something has a positive or negative connotation, on the other hand, can hardly be operationalized automatically, especially in different contexts of situation. Here, more human effort would be needed, as is being done in polarity detection tasks, where polarity lexicons are built (Wiebe et al., 1999; Jijkoun et al., 2010; Clematide and Klenner, 2010; Taboada et al., 2011). Nevertheless, defining whether a stance expression is positioned toward one or the other pole of the stance meaning could already give additional insights on register diversification. Moreover, instead of relying on lexicons and semantic databases to create classes of meanings, one might rely on approaches such as topic models (Blei, 2012) and the like to create contextually relevant classes.

An additional factor that could be considered is reader engagement, which we considered in the semantic system but did not approach in our analysis. This was for several reasons, the main ones being (1) its rareness in scientific writing as opposed to other situation types such as advertisements, political speech etc. where the reader is much more strongly engaged within the discourse, and (2) its lexico-grammatical realizations, which are relatively fuzzy and still not well understood by linguistic means apart from reader mentioning (e.g., by personal pronouns) and thus quite difficult to detect semi-automatically.

In terms of register diversification, we have considered a time span of approximately 30 years or so, as well as newly emerged disciplines, to trace possible diversification trends in terms of recent diachronic trends. It would also be quite interesting to consider a more remote diachronic approach that reaches back to the point where research articles started to emerge and trace the diachronic development of evaluative meaning in research articles from the start. Additionally, other registers could also be considered or would have to be considered when reaching back in time. For this, the proposed analytical cycle as well as the model of analysis could be adopted, while the realizations in the lexico-grammar might have changed quite strongly. Consider this small excerpt of Newton's Opticks paper from 1730, which has been taken from a concluding remark of a previous section of his paper:

> ***I*** *have now given in Axioms and their Explications the sum of what hath hitherto been treated of in[Pg 20] Opticks. For what hath been* ***generally agreed on I content my self to assume*** *under the notion of Principles, in order to what* ***I*** *have farther to write. And this* ***may suffice*** *for an Introduction to Readers of* ***quick Wit*** *and* ***good Understanding not yet versed*** *in Opticks: Although those who are* ***already acquainted*** *with this Science, and have handled Glasses, will* ***more readily apprehend*** *what followeth.*

Clearly, some lexico-grammatical features observed in contemporary scientific articles can also be spotted in this excerpt (e.g., personal pronouns of self-mentioning, epistemic and attitudinal expressions). However, in terms of stance meanings expressed and lexical items used such as *versed*, adaptation would be needed, i.e. new stance meanings might be encountered and other lexical items might have been used back then to express particular meanings. Furthermore, we can also see how the reader is relatively strongly engaged within the discourse (*Readers*, *those who are already acquainted with*[...]) and presents the target of the evaluation, which differs from the writing style of research articles we are now familiar with, especially in a 'hard' science discipline such as physics.

A further concept that deserves deeper investigation is the conventionalization process that seems to be going on in research articles in general as one of the major motifs of diversification, and which also seems to be reflected in evaluative meaning. In our study, we have noted that there seems to be some trend in using evaluative meaning toward the beginning of the sentence, which might be related to a conventionalized usage, but also points toward what Hunston and Thompson (2000) describe as an additional function of evaluation toward organizing the discourse. Moreover, pattern features become more characteristic over time. This might have an influence on the status of the evaluative expression, i.e. if it moves toward a more conventionalized usage, does it still have the same impact as a less conventionalized expression on the reader, or do we somehow expect it to be there? Would this mean that we are less surprised to find an evaluation? And does it diminish its value or strength somehow? Here, clearly a different approach is needed from ours, as we would have to consider the probabilities of an expression occurring given what precedes it. Answers to these questions would have to be pursued by considering linguistic encoding, i.e. the more predictable an expression is, the denser its linguistic encoding becomes and the more conventionalized the expression will be. This will lead in the direction of work conducted based on measures such as relative entropy, looking at the correlation between linguistic encoding and information density, conducted e.g. in psycholinguistic and sociolinguistic studies (Aylett and Turk, 2004; Levy, 2008; Fankhauser et al., 2014a,b).

In summary, this thesis has presented an empirical corpus-based approach to investigate linguistic properties of disciplines regarding evaluative meaning, which can be reused in other application contexts. The analytical cycle proposed and based on Jockers (2013)'s macro- and micro-analytical perspective is general in nature and can be reused for a variety of other studies that seek both qualitative as well as quantitative insights. The model of analysis on evaluative meaning might be employed in the analysis of evaluative meaning in other languages, situational contexts, time periods, etc. with some adaptation. Apart from this, the insights gained might be relevant to other approaches dealing with evaluative meaning (e.g., when looking at differences across registers or fine-grained distinctions), but also to text classification tasks and other studies that aim at distinguishing between more fine-grained classes. Clearly, there are many open questions worth pursuing, only some of which we have put forward in this section.

Chapter 7

German summary

In der vorliegenden Dissertation wurden zwei Hypothesen untersucht: Registerdiversifikation und registerspezifische Prägung.
Bei der Registerdiversifikation wurden die Fragen ergründet, ob Disziplinen eigene Bewertungscharakteristiken aufzeigen, die linguistisch in Registern nachzuverfolgen sind, und ob sich diese über die Zeit verändert haben. Ein besonderes Augenmerk wurde auf relativ neu entstandene Disziplinen gelegt, die durch Kontakt zwischen zwei anderen Disziplinen entstanden sind (z. B. die Computerlinguistik, die sich aus Kontakt zwischen der Informatik und der Linguistik herausgebildet hat), da diese Kontaktdisziplinen dazu neigen, sich von anderen Disziplinen abzugrenzen, d.h. einen Diversifikationsprozess vollziehen, um als eigenständige Register zu gelten.
Bei der registerspezifischen Prägung wurde auf die Fragen eingegangen, ob neu entstandene Disziplinen eine linguistische Prägung in Bezug auf Bewertungen zeigen, die in ihrem Register nachzuverfolgen ist und die von den Ursprungsdisziplinen übernommen worden ist. Darüber hinaus wurde auch untersucht, ob sich die linguistische Prägung diachron verändert hat.
Um diesen Fragen nachzugehen, wurden verschiedene theoretische und methodische Konzepte angewandt.

Theoretisch ist die Arbeit in der Systemisch Funktionalen Linguistik (SFL) begründet, da diese linguistische Theorie eine Grundlage bietet für die Analyse von Bewertungen und aufgrund ihrer soziosemiotischen Perspektive auf die Sprache. Diese erweist sich als besonders nützlich, da sowohl die Wissenschaftsdisziplinen auf der sozio-kulturellen Ebene als auch ihre semantischen und linguistischen Entscheidungen in Bezug auf Bewertungen auf der semantischen und lexico-grammatischen Ebene Gegenstand der Analysen sind (Kapitel 3.2). Basierend auf dieser soziosemiotischen Interpretation von Sprache wurde ein Analysemodel zur Analyse von Bewertungen entwickelt, um die Auswahl aus dem semantischen System auf der lexico-grammatischen Ebene nachzuverfolgen (Kapitel 3.4). Beim Untersuchen der lexico-grammatischen Entscheidungen können somit Rückschlüsse auf die semantischen Entscheidungen gezogen werden. Insbesondere sind anhand bestimmter lexico-

grammatischer Merkmale, die Bewertungen realisieren, die getroffenen Entscheidungen in verschiedenen Wissenschaftsdisziplinen untersucht worden. Beim Vergleich der Entscheidungen auf der lexico-grammatischen Ebene in den verschiedenen Disziplinen konnten Unterschiede und Gemeinsamkeiten dieser in Bezug auf Registerdiversifikation und registerspezifische Prägung analysiert werden.
Das *Registerkonzept* aus der SFL ist in dieser Arbeit von zentraler Bedeutung, da es als Grundlage für unsere Analyse von funktionaler Variation dient, die zwischen hochspezialisierten Disziplinen besteht, d.h. Variation in Bezug auf den Situationskontext (Kapitel 2.3). In unserem Fall wurden neun Situationskontexte analysiert, die sich in den folgenden neun Wissenschaftsdisziplinen widerspiegeln: Informatik, Computerlinguistik, Linguistik, Bioinformatik, Biologie, Bauinformatik, Maschinenbau, Mikroelektronik und Elektrotechnik (Kapitel 4.3). Darüber hinaus entsteht aufgrund von Modernisierungsprozessen die Notwendigkeit, seitens bestimmter Disziplinen, Konzepte, Techniken etc. von anderen Disziplinen zu übernehmen, was zur Entstehung neuer Register führt. In dieser Arbeit berücksichtigen wir vier solcher neu entstandener Disziplinen, um zu untersuchen, ob sie über eine Zeitperiode von 30 Jahren eigene Register gebildet haben, insbesondere in Bezug auf Bewertungsausdrücken.
Das semantische Konzept der Bewertung kann in die interpersonelle Metafunktion der SFL positioniert werden, da diese die Ressourcen bietet für die Zuweisung von Diskursrollen an Teilnehmer und das Ausdrücken von Bewertungen und Stellungsnahmen (Kapitel 3.3). Unser theoretischer Ansatz zum Konzept der Bewertung unterscheidet sich von anderen, da versucht wird, das gesamte Bedeutungspotential von Bewertungen zu berücksichtigen: (1) die beteiligten Teilnehmer, (2) die möglichen Entscheidungen, die dem Teilnehmer aus den Auswahlmöglichkeiten vom Bedeutungspotential zur Verfügung stehen, um Bewertungen auszudrücken, und (3) das Zielobjekt der Bewertung. Hierfür wurde ein semantisches System von Auswahlmöglichkeiten an Bewertungen präsentiert, das es erlaubt, die semantischen Entscheidungen nachzuverfolgen, indem die lexico-grammatischen Merkmale untersucht werden, die bestimmte Entscheidungen realisieren (Kapitel 3.4). Dadurch konnten sehr feine registerspezifische Unterschiede in den verschiedenen hochspezialisierten Disziplinen in Bezug auf Bewertungen ausgemacht werden.

Methodisch wurde ein makro- und mikro-analytischer, korpus-basierter Ansatz angewandt. In diesem Zusammenhang ist ein analytischer Zyklus konzipiert worden, der es ermöglicht, rekursive quantitative und qualitative analytische Schritte vorzunehmen (Kapitel 4.4). Das verwendete Korpus (SciTex) erlaubt es, verschiedene Wissenschaftsdisziplinen diachron zu untersuchen. Durch den iterativen Prozess konnte die bestmögliche Balance zwischen generalisierbarer Tendenzen und detaillierten Einblicken auf Registerdiversifikation und registerspezifische Prägung erlangt werden und das basierend auf empirische Erkenntnisse. Unser methodischer Ansatz unterscheidet sich von anderen, da er nicht nur auf generalisierbare Tendenzen oder detaillierte Analysen fokussiert ist, denn beide Perspektiven werden miteinbezogen. Die Kombination aus Makro- und Mikro-Analyse erlaubt es, den benötigten Detail-

grad auszumachen, um die bestmöglichen Generalisierungen machen zu können.
Auf der makro-analytischen Ebene wurde die Textklassifikation mit Support Vector Machines (SVM) als Technik verwendet, um empirisch zu untersuchen, wie gut Textproduktionen korrekt einer Disziplin in Bezug auf Bewertungen zugeordnet werden können. Daraus konnten Erkenntnisse bezüglich (1) des Grades der Registerdiversifikation erlangt werden, d.h. wie gut sich die Disziplinen auf der Registerebene von anderen unterscheiden und ob es hierbei eine diachrone Veränderung gegeben hat (Kapitel 5.2.1), und (2) den Grad der registerspezifischen Prägung, d.h. wie stark die Kontaktdisziplinen registerbezogene Überlappungen mit ihren Ursprungsdisziplinen aufzeigen und ob sich dies diachron verändert hat (Kapitel 5.3.1). Da der Diversifikationsprozess mit der Zeitdimension zusammenhängt, d.h. er sich über die Zeit entwickelt, ist ein diachroner Vergleich besonders wichtig, um Erkenntnisse über registerbezogene Diversifikationsprozesse zu gewinnen. Darüber hinaus, da die Kontaktregister einen Diversifikationsprozess vollziehen, wurde für die registerspezifische Prägung ebenfalls die diachrone Perspektive miteinbezogen, um mögliche Veränderungen in Bezug auf die Überlappung mit dem einen oder anderen Ursprungsregister zu berücksichtigen.
Auf der mikro-analytischen Ebene wurde eine Merkmalsanalyse vorgenommen, um (1) die Art der Registerdiversifikation (anhand charakteristischer Merkmale eines Registers; Kapitel 5.2.2) und (2) die Art der registerspezifischen Prägung (anhand übernommener Merkmale vom Kontaktregister aus den Ursprungsregistern; Kapitel 5.3.2) zu untersuchen. In der Merkmalsanalyse wurde von einer eher allgemeinen Betrachtung der Ergebnisse auf eine spezifischere übergegangen, d.h. es wurden der Klassifikationsbeitrag der Merkmalstypen und der einzelnen Merkmale sowie die Merkmalsverwendung innerhalb von Konkordanzen und ihre distributionelle Verteilung auf unterschiedlichen Ebenen mitberücksichtigt. Dabei wurden aus dem SVM Output die am höchsten gerankten Merkmale (anhand der SVM-Gewichte) untersucht. Darüber hinaus wurde in unserem analytischen Zyklus auf der mikroanalytischen Ebene ein weiterer Schritt integriert, der es ermöglicht hat, die Datenqualität zu erhöhen, indem detaillierte Bereinigungsprozeduren auf den Extraktionsergebnissen hinsichtlich der am höchsten gerankten Merkmale durchgeführt wurden.

Die Analyse der Registerdiversifikation in Bezug auf Bewertungen hat gezeigt, dass bezüglich des Diversifikationsgrades sowohl die Kontaktregister als auch die Ursprungsregister über den untersuchten Zeitraum von 30 Jahren einen Diversifikationsprozess vollzogen haben, wobei dieser bei den Kontaktregistern stärker ausgeprägt ist im Gegensatz zu ihren Ursprungsregistern. In der Zeitperiode der 70er/80er Jahren sind die Kontaktregister hinsichtlich Bewertungen nicht distinktiv unterscheidbar, d.h. in dieser Zeitperiode zeigen sie keinen eigenen Bewertungscharakter. In den 2000er dagegen sind sie klar voneinander unterscheidbar, d.h. sie haben über den Zeitraum von 30 Jahren einen eigenen Bewertungscharakter entwickelt (Kapitel 5.2.1).
In Bezug auf die Art des Diversifikationsprozesses hat die Analyse der Merkmalstypen und -subtypen generelle diachrone Tendenzen aufgezeigt: während die Kon-

taktregister in der ersten Zeitperiode (70er/80er) nur sehr schlecht anhand von einzelnen Stance-Bedeutungsgruppen unterschieden werden können und sie deshalb keinen eigenen Bewertungscharackter aufzeigen, tragen in der zweiten Zeitperiode (2000er) unterschiedliche Merkmalssubtypen (Part-of-Speech der Stance-Ausdrücke, Stance-Bedeutungsgruppen, Stance-Typen, etc.) zur Unterscheidung bei. Diese Tendenz zeigt sich auch in den Ursprungsregistern, auch wenn die Tendenz weniger ausgeprägt ist, außer beim Maschinenbau. Die Variation hinsichtlich der Merkmalssubtypen ist darauf zurückzuführen, dass ein Bewertungsphänomen in unterschiedliche Merkmalssubtypen reflektiert wird, d.h. je charakteristischer ein Bewertungsphänomen wird, desto mehr Typen reflektieren das Phänomen. Ein Beispiel dafür ist die semantische Bewertungsgruppe der Komplexität in der Informatik. In der ersten Zeitperiode (70er/80er) ist die Komplexität nur anhand eines Merkmals charakteristisch für die Informatik. In der zweiten Zeitperiode (2000er) dagegen scheint sie noch charakteristischer für die Informatik zu sein, da sie in drei Merkmalen reflektiert wird (der Bewertungsgruppe der Komplexität selbst, ihr Gebrauch innerhalb von Patterns am Satzanfang und als Attitude Marker am Satzanfang).
Eine weitere beobachtete diachrone Tendenz ist die Umlagerung der Bewertungsausdrücke am Satzanfang. Das trifft sowohl auf die Kontakt- als auch auf die Ursprungsregister zu, was auf eine konventionellere Verwendung von Bewertungsausdrücken hinweist.
Hinsichtlich bestimmter Charakteristiken in den Kontakt- und Ursprungsregistern konnte beobachtet werden, dass sich diese klar voneinander unterscheiden: die Informatik geht bezüglich epistemischer Bedeutungen eher in Richtung verstärkender Ausdrucksweisen (boosters), die Linguistik dagegen eher in Richtung abschwächender Ausdruckweisen (hedges) und die Biology tendiert zu beiden mit einem starken Fokus auf das Bewerten von experimentellen Erkenntnissen. Die Kontaktregister unterscheiden sich besonders in Bezug auf die verwendeten Bewertungsgruppen: für die Computerlinguistik sind Verhaltensbedeutungen (attitudinal meanings) sowohl am Satzanfang als auch innerhalb eines Satzes distinktiv, während für die Bioinformatik sowohl epistemische als auch Verhaltensausdrücke charakteristisch sind. Die Bauinformatik zeigt einen recht konventionellen Charakter und ist hauptsächlich anhand der Verwendung von Patterns distinktiv, wobei sie auch Verhaltensausdrücke und abschwächende Ausdrücke distinktiv verwendet. Die Mikroelektronik verwendet hauptsächlich Verhaltensbedeutungen innerhalb der Abschlusskapiteln (conclusion sections) charakteristisch.
Besonders interessant ist, dass sich manche Merkmale zwar deutlich als charakteristisch für ein Register erweisen, sie allerdings auch für mehrere Register charakteristisch sind. Die Kontaktregister zum Beispiel verwenden alle *self-mention*, das sich als stark charakteristisch für die Informatik gezeigt hat, wobei sich die Register unterscheiden in Bezug auf die mit der *self-mention* verwendeten Prozesstypen, d.h. was die Autoren machen (z. B. *we describe*), und Konzepte, d.h. was die Autoren besitzen (z. B. *our model*), sowie auch in welchen Kapiteln die *self-mention* hauptsächlich erscheint. Das wird auch für die Biologie und Bioinformatik deutlich, da beide im

Gegensatz zu den anderen Registern sowohl zur Verwendung von verstärkenden als auch von abschwächenden Ausdrücken neigen (Kapitel 5.2.2). Aufgrund dieser Erkenntnisse sind wir unserer zweiten Hypothese nachgegangen, d.h. der registerspezifischen Prägung, da der Anschein erweckt wurde, dass die Ursprungsregister eine Art Prägung auf die Kontaktregister hinterlassen haben.

Die Analyse der registerspezifischen Prägung hat bezüglich des Prägungsgrades verschiedene Tendenzen aufgezeigt, d.h. in Bezug auf die Überlappungsmenge zwischen Kontakt- und Ursprungsregistern. Manche Kontaktregister scheinen zwischen ihren Ursprungsregistern zu liegen (z. B. die Computerlinguistik), eine Tendenz, die sich über die Zeit konstant gehalten hat. Die Bauinformatik und die Mikroelektronik zeigen eine deutliche Tendenz in Richtung ihrer Ursprungsregister Maschinenbau und Elektrotechnik. Die Bioinformatik dagegen scheint eine Verschiebung vollzogen zu haben, da sie in der ersten Zeitperiode (70er/80er) mehr Überlappung mit der Biologie zeigt und für die zweite Zeitperiode (2000er) mehr Überlappung mit der Informatik (Kapitel 5.3.1).
Beim Untersuchen der Art der registerspezifischen Prägung wurden recht interessante Erkenntnisse hinsichtlich der Merkmale gewonnen, die die Kontaktregister von ihren Ursprungsregistern (d.h. von der Informatik oder dem anderen Ursprungsregister) übernommen haben. Die Computerlinguistik hat den abschwächenden Bewertungscharakter der Linguistik übernommen, während sie von der Informatik die Verwendung von Bewertungspatterns und der Position von Bewertungsausdrücken innerhalb eines wissenschaftlichen Artikels übernommen hat. Die Bioinformatik hat auf ähnliche Art verschiedene epistemische und Verhaltensbedeutungen von der Biologie übernommen, während sie die Verwendung von Bewertungsausdrücken innerhalb bestimmter Kapiteln von wissenschaftlichen Artikeln von der Informatik übernommen hat. Die Bauinformatik hat interessanterweise bestimmte Verhaltensbedeutungen vom Maschinenbau übernommen und verwendet diese Bedeutungen innerhalb von Patterns, die sie von der Informatik übernommen hat. Die Mikroelektronik hat, ähnlich zu den anderen Kontaktregistern, Verhaltensbedeutungen aus der Elektrotechnik übernommen und strukturelle Merkmale dagegen von der Informatik. Demnach wurde eine allgemeine Tendenz festgestellt, in der die Kontaktregister epistemische und Verhaltensbedeutungen von ihren Ursprungsregistern übernehmen und strukturelle Merkmale von der Informatik (Kapitel 5.3.2).

Die Untersuchung unserer beiden Hypothesen (Registerdiversifikation und registerspezifische Prägung) auf der Makro- und Mikro-Ebene hat es ermöglicht, ein recht tiefgründiges Verständnis zu erwerben in Bezug auf den möglichen Bewertungscharakter von Wissenschaftsdisziplinen. Es hat sich gezeigt, dass sich Wissenschaftsdisziplinen in ihrer Verwendung von lexico-grammatischen Merkmalen unterscheiden bezüglich dem Grad und der Art der Unterscheidung, aber auch Gemeinsamkeiten aufzeigen. Die Studie hat auch gezeigt, dass die untersuchten Disziplinen einerseits diachron einen eigenen Bewertungscharakter entwickelt haben, d.h. sie haben einen Diversifikationsprozess durchlaufen, und andererseits (und das speziell hinsichtlich der Kontaktdisziplinen), dass sie dennoch eine linguistische Prägung seitens der Ursprungsdisziplinen übernommen haben, die auf der lexico-grammatischen Ebene nachverfolgt werden kann.

Bibliography

Shlomo Argamon, Casey Whitelaw, Paul Chase, Sobhan Raj Hota, Navendu Garg, and Shlomo Levitan. Stylistic Text Classification Using Functional Lexical Features. *Journal of the American Society for Information Science and Technology*, 58(6):802–822, 2007.

Shlomo Argamon, Jeff Dodick, and Paul Chase. Language Use Reflects Scientific Methodology: A Corpus-based Study of Peer-reviewed Journal Articles. *Scientometrics*, 75(2):203–238, 2008.

Matthew Aylett and Alice Turk. The Smooth Signal Redundancy Hypothesis: A Functional Explanation for Relationships between Redundancy, Prosodic Prominence and Duration in Spontaneous Speech. *Language and Speech*, 47(1):31–56, 2004.

R. Harald Baayen. *Analysing Linguistic Data. A Practical Introduction to Statistics using R.* Cambridge University Press, London, 2008.

David Banks. *The Development of Scientific Writing. Linguistic Features and Historical Context.* Equinox, London/Oakville, 2008.

Charles Bazerman. What Written Knowledge Does: Three Examples of Academic Discourse. *Philosophy of the Social Sciences*, 11(3):361–387, 1981.

Vijay K. Bhatia. A Generic View of Academic Discourse. In John Flowerdew, editor, *Academic Discourse*, pages 21–39. Longman, Harlow, 2002.

Douglas Biber. *Variation Across Speech and Writing.* Cambridge University Press, Cambridge, 1988.

Douglas Biber. The Multi-Dimensional Approach to Linguistic Analyses of Genre Variation: An Overview of Methodology and Findings. *Computers and the Humanities*, 26(5-6):331–345, 1993.

Douglas Biber. *Dimensions of Register Variation: A Cross-Linguistic Comparison.* Cambridge University Press, 1995.

Douglas Biber. Stance in Spoken and Written University Registers. *Journal of English for Academic Purposes*, 5(2):97–116, 2006.

Douglas Biber. Register as a Predictor of Linguistic Variation. *Corpus Linguistics and Linguistic Theory*, 8(1):9–37, 2012.

Douglas Biber and Edward Finegan. Styles of Stance in English: Lexical and Grammatical Marking of Evidentiality and Affect. *Text - Interdisciplinary Journal for the Study of Discourse*, 9(1):93–124, 1989.

Douglas Biber and Edward Finegan. *Sociolinguistic Perspectives on Register*. Oxford Studies in Sociolinguistics. Oxford University Press, Oxford, 1994.

Douglas Biber, Stig Johansson, and Geoffrey Leech. *Longman Grammar of Spoken and Written English*. Longman, Harlow, 1999.

David Blei, Andrew Ng, and Michael Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:933–1022, 2003.

David M. Blei. Probabilistic Topic Models. *Communications of the ACM*, 55(4): 77–84, 2012.

Johan Bollen, Huina Mao, and Alberto Pepe. Modeling Public Mood and Emotion: Twitter Sentiment and Socio-economic Phenomena. In *Proceedings of the 5th International AAAI Conference on Webblogs and Social Media (ICWSM)*, pages 450–453, Barcelona, Spain, July 2011.

Wallace L. Chafe. Evidentiality in English Conversation and Academic Writing. In Wallace L. Chafe and Johanna Nichols, editors, *Evidentiality: The Linguistic Coding of Epistemology*, pages 261–273. Norwood, N.J.: Ablex, 1986.

Joanna Channell. Corpus-Based Analysis of Evaluative Lexis. In Susan Hunston and Geoff Thompson, editors, *Evaluation in Text, Authorial Stance and the Construction of Discourse*, pages 38–55. Oxford University Press, Oxford, 2000.

Maggie Charles. "This mystery...": A Corpus-based Study of the Use of Nouns to Construct Stance in Theses from Two Contrasting Disciplines. *Journal of English for Academic Purposes*, 2(4):313–326, 2003.

Simon Clematide and Manfred Klenner. Evaluation and Extension of a Polarity Lexicon for German. In *Proceedings of the 1st Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA) at ECAI2010*, pages 7–13, Lisbon, Portugal, August 2010.

Susan Conrad and Douglas Biber. Adverbial Marking of Stance in Speech and Writing. In Susan Hunston and Geoff Thompson, editors, *Evaluation in Text, Authorial Stance and the Construction of Discourse*, pages 56–73. Oxford University Press, Oxford, 2000.

CWB. The IMS Open Corpus Workbench, 2010. http://www.cwb.sourceforge.net.

Cristian Danescu-Niculescu-Mizil, Michael Gamon, and Susan Dumais. Mark My Words!: Linguistic Style Accommodation in Social Media. In *Proceedings of the 20th International Conference on World Wide Web*, pages 745–754, Hyderabad, India, March 2011.

Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. No Country for Old Members: User Lifecycle and Lingistic Change in Online Communities. In *Proceedings of the 22nd International World Wide Web Conference (IW3C2)*, pages 307–318, Rio de Janeiro, Brazil, May 2013.

Brenda Danet. Language in the Legal Process. *Law and Society Review*, 14(3): 445–564, 1980.

Stefania Degaetano and Elke Teich. The Lexico-Grammar of Stance: An Exploratory Analysis of Scientific Texts. In Stefanie Dipper and Heike Zinsmeister, editors, *Bochumer Linguistische Arbeitsberichte - Beyond Semantics: Corpus-based Investigations of Pragmatic and Discourse Phenomena*, volume 3, pages 57–66, 2011.

Stefania Degaetano-Ortlieb, Hannah Kermes, Ekaterina Lapshinova-Koltunski, and Elke Teich. SciTex - A Diachronic Corpus for Analyzing the Development of Scientific Registers. In Paul Bennett, Martin Durrell, Silke Scheible, and Richard J. Whitt, editors, *New Methods in Historical Corpus Linguistics*, volume 3 of *Corpus Linguistics and Interdisciplinary Perspectives on Language - CLIP*, pages 93–104. Narr, Tübingen, 2013.

Stefania Degaetano-Ortlieb, Peter Fankhauser, Hannah Kermes, Ekaterina Lapshinova-Koltunski, Noam Ordan, and Elke Teich. Data Mining with Shallow vs. Linguistic Features to Study Diversification of Scientific Registers. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC)*, pages 1327–1334, Reykjavik, Iceland, May 2014a.

Stefania Degaetano-Ortlieb, Hannah Kermes, and Elke Teich. The Notion of Importance in Academic Writing: Detection, Linguistic Properties and Targets. *Journal for Language Technology and Computational Linguistics (JLCL)*, 29(1):63–77, 2014b.

Ann Devitt and Khurshid Ahmad. Is There a Language of Sentiment? An Analysis of Lexical Resources for Sentiment Analysis. *Language Resources and Evaluation*, 47(2):475–511, 2013.

Anna Duszak. *Culture and Styles of Academic Discourse*, volume 104. Walter de Gruyter, Berlin, 1997.

Stefan Evert. *The CQP Query Language Tutorial*. IMS Stuttgart, 2005. CWB version 2.2.b90.

Peter Fankhauser, Hannah Kermes, and Elke Teich. Combining Macro- and Microanalysis for Exploring the Construal of Scientific Disciplinarity. In *Proceedings of the Digital Humanities 2014*, Lausanne, Switzerland, July 2014a.

Peter Fankhauser, Jörg Knappen, and Elke Teich. Exploring and Visualizing Variation in Language Resources. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC)*, Reykjavik, Iceland, May 2014b.

Christiane Fellbaum. *WordNet. An Electronic Lexical Database*. The MIT Press, Cambridge, 1998.

Charles A. Ferguson. Baby Talk in Six Languages. *American anthropologist*, 66(6): 103–114, 1964.

Charles A. Ferguson. Dialect, Register, and Genre: Working Assumptions About Conventionalization. In Douglas Biber and Edward Finegan, editors, *Sociolinguistic Perspectives on Register*, pages 15–30. Oxford University Press, New York/Oxford, 1994.

Neal Fox, Omran Ehmoda, and Eugene Charniak. Statistical Stylometrics and the Marlowe - Shakespeare Authorship Debate. In *Proceedings of the Georgetown University Roundtable on Language and Linguistics (GURT)*, Washington, D.C., USA, March 2012.

Murthy Ganapathibhotla and Bing Liu. Mining Opinions in Comparative Sentences. In *Proceedings of the 22nd International Conference on Computational Linguistics (COLING)*, pages 241–248, Manchester, UK, August 2008.

M.A.K. Halliday. Categories of the Theory of Grammar. *Word*, 17:241–292, 1961. Reprinted in abbreviated form in Halliday (1976) edited by Gunther R. Kress, pp. 52–72.

M.A.K. Halliday. *Language as Social Semiotic*. Arnold, London, 1978.

M.A.K. Halliday. On the Language of Physical Science. In Mohsen Ghadessy, editor, *Registers of Written English: Situational Factors and Linguistic Features*, pages 162–177. Pinter, London, 1988.

M.A.K. Halliday. *An Introduction to Functional Grammar*. Arnold, London, 2004.

M.A.K. Halliday and Ruqaiya Hasan. *Language, Context and Text: A Social Semiotic Perspective*. Language and Learning Series. Deakin University Press, Geelong, VIC, Australia, 1985.

M.A.K. Halliday and C.M.I.M. Matthiessen. *Construing Experience Through Meaning. A Language-based Approach to Cognition*. Continuum, London/New York, 2006.

Ruqaiya Hasan. Code, Register and Social Dialect. In Basil Bernstein, editor, *Class, Codes and Control: Applied Studies towards a Sociology of Language*, volume 2, pages 253–292. Routledge and Kegan Paul, London, 1973.

Ruqaiya Hasan. Lending and Borrowing: From Grammar to Lexis. In John E. Clark, editor, *The cultivated Australian*, volume 48, pages 55–67. Helmut Buske, Hamburg, 1985.

Vasileios Hatzivassiloglou and Kathleen R. McKeown. Predicting the Semantic Orientation of Adjectives. In *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics (ACL/EACL)*, pages 174–181, Madrid, Spain, July 1997.

Philip J. Hayes, Laura E. Knecht, and Monica J. Cellio. A News Story Categorization System. In *Proceedings of the 2nd Conference on Applied Natural Language Processing (ACM)*, pages 9–17, Austin, Texas, USA, February 1988.

Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, editors. *The Handbook of Historical Sociolinguistics*. Wiley-Blackwell, Oxford, 2012.

Susan S. Hill, Betty F. Soppelsa, and Gregory K. West. Teaching ESL Students to Read and Write Experimental-Research Papers. *TESOL Quaterly*, 16(3):333–347, 1982.

Thomas Hofmann. Unsupervised Learning by Probabilistic Latent Semantic Analysis. *Machine Learning Journal*, 42(1):177–196, 2001.

Susan Hood. *Appraising Research: Taking a Stance in Academic Writing*. PhD thesis, Faculty of Education, University of Technology, Sydney, 2004.

Susan Hood. Managing Attitude in Undergraduate Academic Writing: A Focus on the Introductions to Research Reports. In Louise J. Ravelli and Robert A. Ellis, editors, *Analysing Academic Writing. Contextualized Frameworks*. Continuum, London/New York, 2005.

Marianne Hundt, Andrea Sand, and Rainer Siemund. *Manual of Information to Accompany The Freiburg LOB Corpus of British English (FLOB)*. Albert-Ludwigs-Universität Freiburg, 1999.

Susan Hunston. *Evaluation in Experimental Research Articles*. PhD thesis, University of Birmingham, 1989.

Susan Hunston. Evaluation and Organization in a Sample of Written Academic Discourse. In Malcolm Coulthard, editor, *Advances in Written Text Analysis*, pages 191–218. Routledge, London, 1994.

Susan Hunston. Counting the Uncountable: Problems of Identifying Evaluation in a Text and in a Corpus. In Alan Partington, John Morley, and Louann Haarman, editors, *Corpora and Discourse*, volume 9 of *Linguistic Insights*, pages 157–188. Peter Lang, Berlin/New York, 2004.

Susan Hunston. *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. Routledge Advances in Corpus Linguistics. Taylor & Francis, New York, 2011.

Susan Hunston and Gill Francis. *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Studies in Corpus Linguistics. John Benjamins Publishing Company, Amsterdam/Philadelphia, 2000.

Susan Hunston and John Sinclair. A Local Grammar of Evaluation. In Susan Hunston and Geoff Thompson, editors, *Evaluation in Text, Authorial Stance and the Construction of Discourse*, pages 74–101. Oxford University Press, Oxford, 2000.

Susan Hunston and Geoff Thompson. *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford University Press, Oxford, 2000.

Ken Hyland. *Hedging in Scientific Research Articles*. John Benjamins Publishing Company, Amsterdam/Philadelphia, 1998.

Ken Hyland. *Disciplinary Discourses: Social Interactions in Academic Writing*. Longman, New York/Toronto, 2000.

Ken Hyland. Bringing in the Reader: Addressee Features in Academic Writing. *Written Communication*, 18:549–574, 2001.

Ken Hyland. Stance and Engagement: A Model of Interaction in Academic Discourse. *Discourse Studies*, 7(2):173–192, 2005.

Ken Hyland. Different Strokes for Different Folks: Disciplinary Variation in Academic Writing. In Kjersti Flottem, editor, *Language and Discipline Perspectives on Academic Discourse*, pages 89–108. Cambridge Scholars Press, Newcastle, 2007.

Ken Hyland. *Academic Discourse*. Continuum Discourse Series. Continuum, London/New York, 2009a.

Ken Hyland. Corpus Informed Discourse Analysis: The Case of Academic Engagement. In *Academic Writing: At the Interface of Corpus and Discourse*, pages 110–128. Continuum, London/New York, 2009b.

Ken Hyland and Polly Tse. Metadiscourse in Academic Writing: A Reappraisal. *Applied Linguistics*, 25(2):156–177, 2004.

Ken Hyland and Polly Tse. 'The Leading Journal in its Field': Evaluation in Journal Descriptions. *Discourse Studies*, 11(6):703–720, 2009.

Valentin Jijkoun, Maarten de Rijke, and Wouter Weerkamp. Generating Focused Topic-specific Sentiment Lexicons. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACM)*, pages 585–594, Uppsala, Sweden, July 2010.

Nitin Jindal and Bing Liu. Identifying Comparative Sentences in Text Documents. In *Proceedings of the 29th ACM Special Interest Group on Information Retrieval (SIGIR)*, pages 244–251, Seattle, WA, USA, August 2006a.

Nitin Jindal and Bing Liu. Mining Comparative Sentences and Relations. In *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI)*, volume 2, pages 1331–1336, Boston, MA, USA, July 2006b.

Thorsten Joachims. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. In Claire Nédellec and Céline Rouveirol, editors, *Machine Learning: ECML-98*, volume 1398 of *Lecture Notes in Computer Science*, pages 137–142. Springer, Berlin/Heidelberg, 1998.

Matthew L. Jockers. *Macroanalysis: Digital Methods and Literary History*. University of Illinois Press, Urbana-Champaign/Chicago/Springfield, 2013.

Hannah Kermes. *Off-line (and On-line) Text Analysis for Computational Lexicography*. PhD thesis, Universität Stuttgart, 2003.

Hannah Kermes and Elke Teich. Formulaic Expressions in Scientific Texts: Corpus Design, Extraction and Exploration. *Lexicographica*, 28(1):99–120, 2012.

Moshe Koppel, Shlomo Argamon, and Anat Rachel Shimoni. Automatically Categorizing Written Texts by Author Gender. *Literary and Linguistic Computing*, 17(4):401–412, 2002.

Henry Kučera and Winthrop N. Francis. *Computational Analysis of Present-Day American English*. Brown University Press, Providence, RI, USA, 1967.

Thomas S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 1 edition, 1962.

John Lafferty, Andrew McCallum, and Fernando Pereira. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In *Proceedings of the 18th International Conference on Machine Learning (ICML)*, pages 282–289, Williamstown, MA, USA, June/July 2001.

Geoffrey N. Leech. *English in Advertising: A Linguistic Study of Advertising in Great Britain*. Longman, London, 1966.

Jingsheng Lei, Yanghui Rao, Qing Li, Xiaojun Quan, and Liu Wenyin. Towards Building a Social Emotion Detection System for Online News. *Future Generation Computer Systems*, 37:438–448, 2014.

Roger Levy. Expectation-based Syntactic Comprehension. *Cognition*, 106(3):1126–1177, 2008.

Bing Liu. Sentiment Analysis and Subjectivity. In Nitin Indurkhya and Fred J. Damerau, editors, *Handbook of Natural Language Processing*, Chapman & Hall/CRC Machine Learning & Pattern Recognition, pages 627–666. CRC Press Taylor & Francis Group, Boca Raton, 2 edition, 2010.

Bing Liu, Minqing Hu, and Junsheng Cheng. Opinion Observer: Analyzing and Comparing Opinions on the Web. In *Proceedings of the 14th International Conference on World Wide Web*, pages 342–351, Chiba, Japan, May 2005.

Christian Mair. *Twentieth-Century English: History, Variation and Standardization.* Cambridge University Press, Cambridge, 2006.

Christian Mair. Corpus Linguistics Meets Sociolinguistics: The Role of Corpus Evidence in the Study of Sociolinguistic Variation and Change. In Antoinette Renouf and Andrew Kehoe, editors, *Corpus Linguistics: Refinements and Reassessments*, number 69 in Language and Computers: Studies in Practical Linguistics, pages 7–32. Rodopi, Amsterdam/New York, 2009.

Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. *Introduction to Information Retrieval*, volume 1. Cambridge University Press, Cambridge, 2008.

Jim R. Martin. Interpersonal Meaning, Persuasion, and Public Discourse: Packing Semiotic Punch. *Australian Journal of Linguistics*, 15:3–67, 1995.

Jim R. Martin. Beyond Exchange: APPRAISAL Systems in English. In Susan Hunston and Geoff Thompson, editors, *Evaluation in Text.* Oxford University Press, Oxford, 2000.

Jim R. Martin and Peter R.R. White. *The Language of Evaluation, Appraisal in English.* Palgrave Macmillan, Hampshire/New York, 2005.

Tony McEnery and Andrew Wilson. *Corpus Linguistics: An Introduction.* Edinburgh University Press, Edinburgh, 2 edition, 2001.

Ramanathan Narayanan, Bing Liu, and Alok Choudhary. Sentiment Analysis of Conditional Sentences. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 180–189, Singapore, August 2009.

Kevin Ngozi Nwogu. The Medical Research Paper: Structure and Functions. *English for Specific Purposes*, 16(2):119–138, 1997.

Elinor Ochs and Bambi Schieffelin. Language Has a Heart. *Text*, 9(1):7–25, 1989.

Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan. Thumbs Up? Sentiment Classification Using Machine Learning Techniques. In *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 79–86, Philadephia, PA, USA, July 2002.

David M.W. Powers. Evaluation: From Precision, Recall and F-Measure Topic ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*, 2(1):37–63, 2011.

Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. *A Comprehensive Grammar of the English Language*. Longman, London, 1985.

Philip Resnik and Jimmy Lin. Evaluation of NLP Systems. In Alexander Clark, Chris Fox, and Shalom Lappin, editors, *The Handbook of Computational Linguistics and Natural Language Processing*, pages 271–296. Wiley-Blackwell, Oxford, 2013.

Ellen Riloff and Janyce Wiebe. Learning Extraction Patterns for Subjective Expressions. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 105–112, Sapporo, Japan, July 2003.

Ellen Riloff, Janyce Wiebe, and Theresa Wilson. Learning Subjective Nouns Using Extraction Pattern Bootstrapping. In *Proceedings of the 7th Conference on Natural Language Learning (CoNLL) at HLT-NAACL*, pages 25–32, Edmonton, Canada, May/June 2003.

Kirk Roberts, Michael A. Roach, Joseph Johnson, Josh Guthrie, and Sanda M. Harabagiu. EmpaTweet: Annotating and Detecting Emotions on Twitter. In *Proceedings of the 8th International Conference on Language Resources and Evaluation Conference (LREC)*, pages 3806–3813, Istanbul, Turkey, May 2012.

Josef Ruppenhofer, Michael Ellsworth, Miriam R.L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. FrameNet II: Extended Theory and Practice. Technical report, ICSI, 2010.

Jan Rybicki. Burrowing into Translation: Character Idiolects in Henryk Sienkiewicz's Trilogy and its Two English Translations. *Literary and Linguistic Computing*, 21(1):91–103, 2006.

Deborah Schiffrin, Deborah Tannen, and Heidi E. Hamilton, editors. *The Handbook of Discourse Analyses*. Blackwell Publishing, Oxford, 2003.

Helmut Schmid. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of the International Conference on New Methods in Language Processing*, pages 44–49, Manchester, UK, September 1994.

Diana E. Adams Smith. Medical Discourse: Aspects of Author's Comment. *English for Specific Purposes*, 3(1):25–36, 1984.

Swapna Somasundaran. *Discourse-level Relations for Opinion Analysis*. PhD thesis, University of Pittsburgh, 2010.

Carlo Strapparava and Rada Mihalcea. Learning to Identify Emotions in Text. In *Proceedings of the 23rd Annual ACM Symposium on Applied Computing*, pages 1556–1560, Fortaleza, Ceará, Brazil, March 2008.

Carlo Strapparava and Rada Mihalcea. Annotating and Identifying Emotions in Text. In Giuliano Armano, Marco de Gemmis, Giovanni Semeraro, and Eloisa Vargiu, editors, *Intelligent Information Access*, volume 301 of *Studies in Computational Intelligence*, pages 21–38. Springer, Berlin/Heidelberg, 2010.

Michael Stubbs. *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*, volume 4 of *Language in Society*. University of Chicago Press, Chicago, 1983.

John Swales. *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press, Cambridge/New York, 1990.

Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, 37 (2):267–307, 2011.

Elke Teich. *Cross-Linguistic Variation in System und Text. A Methodology for the Investigation of Translations and Comparable Texts*. Mouton de Gruyter, Berlin, 2003.

Elke Teich and Peter Fankhauser. Exploring a Corpus of Scientific Texts Using Data Mining. In Stefan Th. Gries, Stefanie Wulff, and Mark Davies, editors, *Corpus-linguistic Applications: Current Studies, New Directions*, pages 233–247. Rodopi, Amsterdam/New York, 2010.

Elke Teich, Stefania Degaetano-Ortlieb, Hannah Kermes, and Ekaterina Lapshinova-Koltunski. Register Contact: An Exploration of Recent Linguistic Trends in the Scientific Domain. In *Proceedings of Historical Corpora 2012*, Frankfurt, Germany, December 2012.

Elke Teich, Stefania Degaetano-Ortlieb, Hannah Kermes, and Ekaterina Lapshinova-Koltunski. Scientific Registers and Disciplinary Diversification: A Comparable Corpus Approach. In *Proceedings of the 6th Workshop on Building and Using*

Comparable Corpora (BUCC) at ACL2013, pages 59–68, Sofia, Bulgaria, August 2013.

Puleng Thetela. Evaluated Entities and Parameters of Value in Academic Research Articles. *English for Specific Purposes*, 16(2):101–118, 1997.

Geoff Thompson. Interaction in Academic Writing: Learning to Argue with the Reader. *Applied Linguistics*, 22(1):58–78, 2001.

Geoff Thompson and Heloisa Collins. Interview with M.A.K. Halliday, Cardiff, July 1998. *DELTA: Documentação de Estudos em Lingüística Teórica e Aplicada*, 17: 131–153, 2001.

Geoff Thompson and Jianglin Zhou. Evaluation and Organization in Text: The Structuring Role of Evaluative Disjuncts. In Susan Hunston and Geoff Thompson, editors, *Evaluation in Text, Authorial Stance and the Construction of Discourse*, pages 121–141. Oxford University Press, Oxford, 2000.

Peter D. Turney. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL)*, pages 417–424, Philadelphia, PA, USA, July 2002.

Jean Ure. Lexical Density and Register Differentiation. In George E. Perren and John L.M. Trim, editors, *Applications of Linguistics: Selected papers of the Second International Congress of Applied Linguistics, Cambridge 1969*, pages 443–452. Cambridge University Press, Cambridge, 1971.

Jean Ure. Introduction: Approaches to the Study of Register Range. *International Journal of the Sociology of Language*, 35:5–23, 1982.

Margaret M. Van Naerssen. Medical Records: One Variation of Physicians Language. *International Journal of the Sociology of Language*, 1985(51):43–74, 1985.

Cornelis J. Van Rijsbergen. *Information Retrieval*. Butterworth, London, 2 edition, 1979.

Vladimir N. Vapnik and A. Ja Chervonenkis. *Theory of Pattern Recognition*. Nauka, Moscow, 1974.

Eija Ventola. Packing and Unpacking of Information in Academic Texts. In Eija Ventola and Anna Mauranen, editors, *Academic Writing. Intercultural and Textual Issues*, Pragmatics & Beyond: New Series (P&B): 41, pages 153–194. John Benjamins Publishing Company, Amsterdam/Philadelphia, 1996.

Gregory K. West. That-Nominal Constructions in Traditional Rhetorical Divisions of Scientific Research Papers. *TESOL Quarterly*, 14(4):483–488, 1980.

Peter R.R. White. Dialogue and Inter-Subjectivity: Reinterpreting the Semantics of Modality and Hedging. In Malcolm Coulthard, Janet Cotterill, and Frances Rock, editors, *Working With Dialog*. Neimeyer, Tübingen, 2000.

Peter R.R. White. Beyond Modality and Hedging: A Dialogic Ciew of the Language of Intersubjective Stance. In Mary Macken-Horarik and James R. Martin, editors, *Text - Negotiating Heteroglossia: Social Perspectives on Evaluation*, volume 23, pages 259–284. Mouton de Gruyter, Berlin/New York, 2003.

Janyce M. Wiebe, Rebecca F. Bruce, and Thomas P. O'Hara. Development and Use of a Gold-standard Data Set for Subjectivity Classifications. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 246–253, Maryland, USA, June 1999.

Theresa Wilson, Janyce Wiebe, and Rebecca Hwa. Just How Mad are You? Finding Strong and Weak Opinion Clauses. In *Proceedings of the 19th National Conference of the Association for the Advancement of Artificial Intelligence (AAAI)*, pages 761–769, San Jose, California, July 2004.

Theresa Wilson, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, Janyce Wiebe, Yejin Choi, Claire Cardie, Ellen Riloff, and Siddharth Patwardhan. OpinionFinder: A System for Subjectivity Analysis. In *Proceedings of Human Language Technologies Conference/the 2005 Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, pages 34–35, Vancouver, Canada, October 2005a.

Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. Recognizing Contextual Polarity in Phrase-level Sentiment Analysis. In *Proceedings of Human Language Technologies Conference/the 2005 Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, pages 347–354, Vancouver, Canada, October 2005b.

Theresa Ann Wilson. *Fine-grained Subjectivity and Sentiment Analysis: Recognizing the Intensity, Polarity, and Attitudes of Private States*. PhD thesis, University of Pittsburgh, 2008.

Ian H. Witten, Eibe Frank, and Mark A. Hall. *Data Mining: Practical Machine Learning Tools and Techniques*. The Morgan Kaufmann Series in Data Management Systems. Elsevier Science, Amsterdam/New York, 3rd edition, 2011.

Ding Xiaowen and Bing Liu. Resolving Object and Attribute Coreference in Opinion Mining. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING)*, Beijing, China, August 2010.