Stefania Degaetano-Ortlieb, Hannah Kermes, Elke Teich

The notion of importance in academic writing: detection, linguistic properties and targets

Abstract We present a semi-automatic approach to study expressions of evaluation in academic writing as well as targets evaluated. The aim is to uncover the linguistic properties of evaluative expressions used in this genre, i.e. investigate which lexico-grammatical patterns are used to attribute an evaluation towards a target. The approach encompasses pattern detection and the semi-automatic annotation of the patterns in the SciTex Corpus (Teich and Fankhauser, 2010; Degaetano-Ortlieb et al., 2013). We exemplify the procedures by investigating the notion of importance expressed in academic writing. By extracting distributional information provided by the annotation, we analyze how this notion might differ across academic disciplines and sections of research articles.

1 Introduction

While there are many studies on the detection and description of evaluative expressions in computational linguistics, corpus linguistics as well as descriptive linguistics (e.g., Wilson (2008); Hunston (2011); Biber et al. (1999); Hyland (2005); Martin and White (2005)), a comprehensive method of analysis is still missing. This is due to the phenomenon itself, which can be realized in a variety of ways and which is extremely context dependent. Additionally, different genres pose diverse challenges to the (automatic/semi-automatic/manual) detection of sentiments. Thus, in order to detect evaluative expressions, one has to uncover the linguistic properties of these expressions according to situational context. Only then are we able to make generalizations on how evaluation is realized within one or more languages in a particular context.
In the present paper, we present a corpus-based analysis of one particular aspect of evaluative expressions, *the notion of importance*, in the genre of *scientific research articles*. According to Swales (1990), the author of a research article tries to create a research space to locate the research. Nwogu (1997) has elaborated Swales' model to show how research articles are structured. His model shows that the Introduction and Conclusion sections are prone to be the most evaluative sections of a research article. 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 own new research is emphasized. In the Conclusion (1) observations are indicated by hedging (with modal verbs or verbs such as *appear* or *seem* which attenuate the evaluative expression, e.g., *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*), and (4) specific research outcomes are explained by lexical items signaling significance/importance (e.g., *results are important*) and by preparatory statements 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 academic writing (such as *importance/significance*) and (b) possible targets the evaluation is directed towards (such as *previous/own research*, *observations*, *outcomes*, etc.).

According to Hunston and Francis (2000) expressions of evaluation towards a target can be expressed by evaluative patterns, i.e. lexico-grammatical structures that attribute an evaluation to a target. Here, we can have strictly evaluative patterns, such as *it BE ADJ to/that*, where the ADJ position is always filled with an evaluative adjective, or patterns that are possibly evaluative, such as *the importance of linear problem kernels*, where the noun preceding the of-phrase can have an evaluative meaning such as importance in this case. Clearly, it is not possible to detect all instances of evaluative language by structural patterns. However, the pattern approach allows a fairly systematic way of identification of particular evaluative expressions in large corpora, supporting a more comprehensive picture of the linguistic properties involved in evaluation.

To detect these patterns and targets, we rely on a corpus-based approach that involves detection of evaluative patterns and pattern annotation. Having the corpus annotated, we can analyze differences between disciplines in terms of evaluative expressions and explore the linguistic properties of specific evaluative 'modes' in academic writing, such as evaluative meanings (e.g., *importance*, *obviousness*, *complexity*) and evaluative attribution structures, i.e. whether the evaluative expression precedes the target (pre-evaluation, e.g., *the* [$_{\text{eval}}$ *importance*] *of* [$_{\text{target}}$ *linear problem kernels*]) or follows the target (post-evaluation, e.g., [$_{\text{target}}$ *A*] [$_{\text{eval}}$ *fails*] *to be a BPP algorithm*) or whether a relational structure is used to attribute the evaluation to the target (e.g., [$_{\text{eval}}$ *One crucial issue*] [$_{\text{rel}}$ *is*] [$_{\text{target}}$ *that of stability*]). Our main goal is to examine the linguistic properties of evaluative expressions of importance and to see whether they differ across academic disciplines and document sections. We address the following questions: Which lexical units are used to express importance? Which are the linguistic properties of expressions of importance, i.e. which lexico-grammatical patterns are used? Which kinds of targets are evaluated and are there differences across disciplines and sections of research articles?

The paper is structured as follows: In Section 2, we describe the data and the methodology applied. Section 3 presents the analysis of importance in academic writing. Section 4 concludes the paper with a summary and an envoi.

2 Data and Methods

2.1 Corpus

To investigate evaluation in academic writing, we use SciTex, the English Scientific Text Corpus (Teich and Fankhauser, 2010; Degaetano-Ortlieb et al., 2013), which covers nine academic disciplines (computer science, computational linguistics, bioinformatics, digital construction, microelectronics, linguistics, biology, mechanical engineering and electrical engineering) and contains 34 million words. SciTex comprises two time slices, the 70/80s (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). In this investigation, we consider the early 2000s subcorpus only which amounts to approx. 17.5 million words. 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).

2.2 Pattern detection by text analysis

Inspected subcorpus To detect evaluative lexico-grammatical patterns involved in academic writing, a random sample of DaSciTex, which amounts to approx. 52.000 words, was manually inspected and annotated. This subcorpus is built out of the abstract, introduction and conclusion sections only. The selection was motivated by Nwogu (1997)'s observations that these sections are apt to include a large amount of evaluation in comparison to the main part of research articles and was supported during our own corpus inspection. Taking only these sections of the articles allows us to cover more text that is possibly evaluative and a greater variety of authors. The annotation was performed by one person with the UAM CorpusTool (O'Donnell, 2008), which allows users to annotate text spans manually and to create own annotation schemes that can be adapted during the annotation. Note that the purpose of the annotation was to determine the lexico-grammatical patterns signaling evaluation for later use in larger-scale extraction. In order to do so, a random sample from the corpus was inspected. The purpose was not to create a gold-standard, as is needed, e.g., in tasks of determining positive and negative evaluations, so that in our case annotation by multiple annotators was not necessary. Our procedure allowed us to detect and quantify specific lexico-grammatical patterns of evaluation used in the corpus. The detected patterns are grouped into sets for which annotation rules are created that enable the annotation of much bigger corpora in a consistent and semi-automatic way. More detail on the semi-automatic annotation procedures used to annotate the full version of DaSciTex is provided in the following sections.

Lexico-grammatical patterns The manual text analysis showed that five sets of lexico-grammatical patterns (see Figure 1) are used to express evaluation in academic writing, covering 1740 instances of evaluation in the sample of 52.000 words: two pre-evaluation sets (*eval_target* (40.29%), *eval_relational-v_target* (7.36%)) and three post-evaluation sets (*target_eval* (32.36%), *target_relational-v_eval* (18.10%), *target_v_eval* (4.20%)). Note that different evaluative meanings can be expressed by these patterns (see, e.g., *importance* in Example (1) and *appropriateness* in Example (7)).
The *eval_target* comprises patterns where the evaluative expression precedes the target (see Examples (1)-(2)), whereas in the *target_eval* the evaluative expression follows the target (see Examples (3)-(4)). Two of the pattern sets are used with relational verbs, *eval_relational-v_target* and *target_relational-v_eval*, used also with pre- or post-evaluation, respectively (see Examples (5)-(8)). Additionally, there is one pattern set that involves no relational but other types of verbs, *target_v_eval*, which is only used with post-evaluation (see Examples (9)). Note that in terms of targets, we encounter not only nominal targets but also clausal ones as in Examples (2) and (6).

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

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

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

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

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

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

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

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

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

2.3 Pattern annotation by semi-automatic annotation procedures

To annotate the full 2000s version of SciTex with the patterns discovered by the manual annotation, we use annotation procedures derived from the YAC recursive chunker (Kermes, 2003). We use the Corpus Workbench (CWB, 2010) to annotate patterns by using (1) queries as rules based on PoS tags and structural attributes that search for a defined pattern in the corpus and (2) Perl scripts that allow one to delimit the range of the patterns found and define the attributes to be annotated.
Consider the query in Figure 2 which is used to annotate one prepositional pattern (*eval-np_of_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

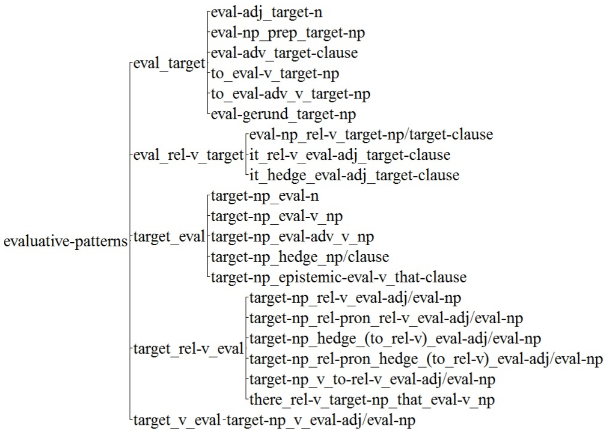

Figure 1: Evaluative patterns identified

by a prepositional phrase, a conjunction or a dash, trying to cover the most common noun phrase dispositions in DaSciTex. These rules were defined manually and results were evaluated for precision in a small version of DaSciTex (one million words). Precision for all patterns amounted to approx. 94.24% to 100%.[1]

Additional information is annotated in form of attributes and comprises: (a) the *evaluation type* described by the pattern sets (e.g., *eval_target*) with the information of having a pre- or post-evaluation, a relational pattern or a verbal one, (b) the *evaluation pattern* (e.g., *eval-adj_target-n*), (c) the *precision of the annotation* derived by the 1

[1] Recall has not been calculated at this stage as it is not a trivial task in a corpus of 34 million words, but we plan to do so by annotating a part of SciTex with the annotation procedure and evaluate the results obtained.

```
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 2: Example of an annotation rule for attributive features

million words subcorpus, and (d) the *evaluation meaning* of the evaluative expression (e.g., *importance*, *obviousness*).

lexical item	pos	notes	lexical item	pos	notes
acute	adj	FN	necessary	adj	corpus/WN (essential)
central	adj	corpus/WN (essential)	necessarily	adv	corpus/WN (essential)
considerable	adj	corpus/FN	necessity	noun	WN (essential)
considerably	adv	WN (considerable)	notable	adj	corpus/WN (significant)
critical	adj	corpus/FN	notably	adv	WN (remarkable)
crucial	adj	corpus/FN	noteworthy	adj	corpus/WN (significant)
crucially	adv	corpus	noticeable	adj	corpus/WN (noteworthy)
decisive	adj	FN	noticeably	adv	WN (noteworthy)
emphasize/se	verb	corpus/WN (important)	outstanding	adj	WN (significant)
essential	adj	corpus/WN (important)	pivotal	adj	FN
essentially	adv	WN (essential)	prominent	adj	corpus/WN (important)
fundamental	adj	corpus/FN	relevant	adj	corpus
fundamentally	adv	WN (essential)	remarkable	adj	corpus/WN (significant)
highlight	verb	corpus/WN (prominent)	salient	adj	WN/FN (prominent)
importance	noun	corpus/FN	serious	adj	corpus/FN
important	adj	corpus/FN	seriously	adv	corpus/FN
importantly	adv	corpus/WN (important)	significance	noun	corpus/FN
indispensable	adj	WN (essential)	significant	adj	corpus/FN
interest	noun	corpus	significantly	adv	corpus/WN (significant)
key	adj	corpus/FN	stress	verb	WN (important)
main	adj	corpus/FN	substantial	adj	corpus/WN (important)
major	adj	corpus/FN	substantially	adv	WN (considerable)
meaningful	adj	corpus	valuable	adj	corpus/WN (worth)
			vital	adj	corpus/FN

Table 1: Lexical items of importance used

To annotate evaluative meanings, we create lists of lexical items expressing these meanings for adjectives, nouns, adverbs and verbs. The procedure applied is exemplified by the importance meaning in the following. Other meanings that we are going to cover are desirability (e.g., *fortunate*, *hopefully*), obviousness (e.g., *clear*, *obvious*), probability (e.g., *probably*, *possibly*), progress (e.g., *improve*, *enhance*), evidence (e.g., *confirm*, *prove*), complexity (e.g, *difficult*, *easy*) and others. Some of these represent assessment types for modal adverbs according to Halliday (2004: 82 and 130), others are related to Hunston (2004) and own previous work on SciTex (Degaetano-Ortlieb et al., 2012; Degaetano, 2010).

To create a list of lexical items expressing importance, (1) we used the lexical items listed in the Frame Index in FrameNet (Ruppenhofer et al., 2010) for the importance meaning (marked with 'FN' in Table 1), (2) we extracted a list of lexical items annotated as being evaluative in our sample corpus and selected those expressing importance (marked in Table 1 with 'corpus'), and (3) used WordNet to find synonyms for the lexical items taken from FrameNet and the own corpus (marked with 'WN' in Table 1). Considering the lexical items in FrameNet for importance, we have a 83% overlap with items found in our sample corpus, i.e. the notion of importance in FrameNet mostly matches the notion found in our sample corpus (besides *acute*, *decisive* and *pivotal* which are not present in the sample corpus, but are used in DaSciTex). Additionally, we added the notions of *essential*, *noteworthy*, *prominent* and *significant* as well as their synonyms from WordNet to our notion of importance (see again Table 1), resulting in a somehow broader definition of importance than FrameNet, which accounts for them separately.

```
1 <evaluation>[_.evaluation_meaning="importance"]+</evaluation>;
2 group Last match text_ad;
3 <evaluation>[_.evaluation_pattern="eval-adj_target-n"]+</evaluation>;
4 group Last matchend lemma;
5 <evaluation>[_.evaluation_pattern="eval-np_rel-V_target-np" & pos!="N.*"]{0,3}
  @[_.evaluation_pattern="eval-np_rel-V_target-np" & pos="N.*"]
  [_.evaluation_pattern="eval-np_rel-V_target-np"]+</evaluation>;
6 group Last target lemma;
```

Figure 3: Queries used to extract targets

target	freq
result	810
model	613
solution	603
value	595
role	562
condition	534
system	506
algorithm	470
difference	468

Figure 4: Targets extracted from the eval-adj_target-n pattern

2.4 Extraction of distributional information and targets

Having the patterns and the attributes annotated, we can extract distributional information, i.e. we can, for example, look at how the patterns are distributed across disciplines or how the meaning of importance is used across disciplines and document sections. The query in Figure 3 line 1, for example, is used to extract instances of the meaning of importance. Distributional information across academic disciplines (text_ad) is extracted with the command in line 2. Moreover, we can extract targets from the annotated structures. Depending on where the target is positioned within the evaluative pattern, the complexity of the extraction can vary. For the *eval-adj_target-n* pattern, for example, target extraction is quite simple as the target is located at the end of the annotated pattern. The command for the extraction is shown in Figure 3 line 3. The command in line 4 is executed to extract the targets used in the pattern as well as their frequencies (see Figure 4). For the relational pattern *eval-np_rel-V_target-np* the extraction is a bit more complex as the target might be located in the middle of the pattern (see Example (5) above where the target is *area A1*). Here, CQP allows for the marking of specific positions for extraction with the anchor `@`. Line 5 in Figure 3 shows the extraction of a nominal target marked by the anchor. The command in line 6 is then executed to extract the targets and their frequencies.

pattern type	pattern	eval imp freq	eval imp %
eval_target	**eval-adj__target-n**	**13567**	**64.33**
	eval-np__prep__target__np	**863**	**4.10**
	eval-adv__target-clause	360	1.71
	eval-v__target-np	252	1.19
eval_rel-v__target	**it__rel-v__eval-adj__target-clause**	**1354**	**6.47**
	eval-np__rel-v__target-np/clause	**1043**	**4.95**
	ex-there__rel-v__eval-adj/np__target-np	408	1.93
target__eval	target-n__eval-adv__v__np	295	1.40
	target__np__eval-n__np	242	1.14
	target__np__eval-n	88	0.42
target__rel-v__eval	**target-np__rel-v__eval-adj/np**	**1504**	**7.13**
	target-np__v__to__be__eval-np	106	0.50
target__v__eval	**target-np__v__eval-adv/np**	**128**	**4.74**

Table 2: Evaluation type and patterns for importance in DaSciTex

3 Analysis: The notion of importance in academic writing

In order to obtain evidence of the attribution of importance in the SciTex corpus, we pose the following questions:

- Which are the linguistic properties of expressions of importance, i.e. which lexico-grammatical patterns are used?
- Which are the most evaluative sections in a research article and which sections express more evaluations of importance?
- Are there differences in the use of importance across disciplines and document sections?
- Which targets are evaluated as being important?

First, we want to know which linguistic properties are used to express importance within DaSciTex. This information is obtained by the procedures explained in Section 2.4. Table 2 shows that the *eval-adj__target-n* pattern is the most frequent pattern with 64.33% (realized by expressions as shown in Example (1)). The second most frequent pattern is a relational one, *target-np__rel-v__eval-adj/np* (see Example (7) for a realization), which amounts to approx. 7%. Four other patterns follow: the impersonal *it* construction with 6.47% (see Example (6)), the relational construction *eval-np__rel-v__target-np/clause* with 4.95% (see Example (5)), the verbal construction *target-np__v__eval-adv/np* with 4.74% (see Example (9)), and the prepositional construction with 4.10% (such as *the importance of linear problem kernels*). The other patterns, occur all less than 2.00%. In terms of linguistic properties, the importance meaning is mostly propagated by pre-evaluative structures (84.68% pre-evaluative vs. 15.32% post-evaluative), where the evaluative expression precedes the target.

Second, we look at how much evaluation is expressed by the patterns analyzed and how much of it realizes the meaning of importance across the four document sections marked in SciTex (Abstract, Introduction, Main and Conclusion). Considering evaluation overall, we can see that the Introduction and the Conclusion are the most evaluative sections (both showing approx. 11,300 expressions of evaluation per 1M),

section	section size	eval. freq	eval. per 1M
Introduction	2150390	24343	11320.27
Conclusion	517205	5849	11308.86
Abstract	1501711	15765	10498.03
Main	11196303	85421	7629.39

Table 3: Evaluation across all document sections in DaSciTex

section	eval-imp freq	eval-imp per 1M
Introduction	4362	2028.47
Abstract	2567	1709.38
Conclusion	886	1674.38
Main	13459	1202.09

Table 4: Evaluation of importance across document sections in DaSciTex

which is in line with observations made by (Nwogu, 1997), (Hood, 2005) and others. What follows is the Abstract (approx. 10,500) and the Main section (approx. 7600), the latter showing much less evaluation than the other sections (see Table 3).

Considering the meaning of importance, the amount in SciTex is of approx. 16% (131378 occurrences overall of which 21254 are of importance) and the section in which importance is mostly used is the Introduction section with approx. 2000 importance expressions per 1M (see Table 4). The Abstract and Conclusion sections follow (both approx. 1700) as well as the Main part of research articles with the least amount of importance (approx. 1200). Thus, in comparison to all occurrences of evaluation annotated by our approach, the importance meaning occurs mostly at the beginning of research articles (Introduction and Abstract). Additionally, the Abstract shows almost an equal amount of evaluation of importance as the Conclusion, even though it has less evaluation overall. Comparing the use of evaluation and evaluations of importance in the Introduction and Conclusion sections by chi-square test, we obtain a p-value of 1.905e-06, i.e. the importance meaning is significantly more often used within the Introduction section in DaSciTex.

Third, we analyze the use of importance across academic disciplines and document sections. Figure 5 shows that computer science (A) makes the least use of the importance meaning, linguistics (C1), instead, uses it most frequently and computational linguistics (B1) is somewhere in between. Considering biology (C2) and bioinformatics (B2), they use importance quite similarly in amount. For the engineering disciplines, the newly emerged disciplines, digital construction (B3) and microelectronics (B4), make more use of importance than their seed disciplines, mechanical engineering (C3) and electrical engineering (C4). Considering the distribution across sections for each discipline (see Figure 6), computer science (A) uses importance most frequently in the Abstract and Introduction and less frequently in the Main part and Conclusion section. The comparison of computational linguistics (B1) and linguistics (C1) by chi-square shows significant differences (p-value of 7.862e-11) due to a higher use of importance in the Abstract for computational linguistics (B1). In comparison to the other disciplines, bioinformatics (B2) and biology (C2) use importance evaluations more frequently within the Conclusion section. The engineering disciplines are relatively similar in their use of

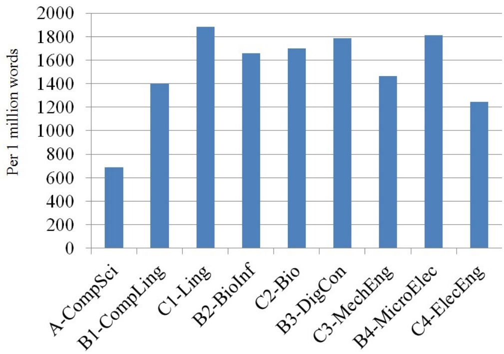

Figure 5: Importance across academic disciplines in DaSciTex

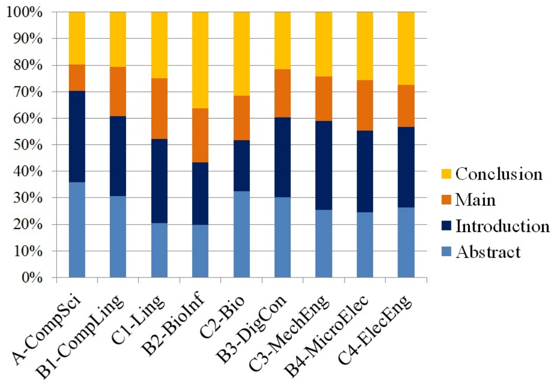

Figure 6: Importance across academic disciplines by document sections in DaSciTex

the importance meaning across sections in comparison to the other disciplines.

Fourth, we inspect which targets are evaluated with importance across the SciTex disciplines. As previously mentioned, targets might be realized as nominal phrases or clauses (e.g., that-clauses). Here, we focus on nominal targets used with the two most frequent patterns that evaluate a nominal target (*eval-adj_target-n* and *target-np_rel-v_eval-adj/np*; see again Table 2). Considering the top 10 to 20 targets across disciplines (see Table 5 for five disciplines), the following observations can be made: (1) we observe domain-specific variation across disciplines (e.g., A-CompSci: *function*, *variable*; B1-CompLing: *word*, *document*; B2-BioInf: *gene*, *residue*), (2) some targets are shared across disciplines being more general in nature (e.g., *difference* and *role* in

A-CompSci		B1-CompLing		C1-Ling		B2-BioInf		C2-Bio	
target	per 1M	target	per 1M	target	per 1M	target	per 1M	target	per 1M
result	52.53	difference	32.00	difference	38.88	gene	53.08	role	90.77
problem	12.24	word	25.48	role	33.78	difference	40.17	difference	39.68
property	12.24	information	21.92	factor	29.32	role	28.69	factor	22.32
idea	11.45	feature	17.78	effect	29.32	feature	27.26	protein	16.86
role	8.29	role	17.78	property	24.86	residue	22.24	gene	16.37
application	6.71	document	17.18	question	22.95	improvement	19.37	component	15.38
question	6.71	problem	16.59	point	17.85	information	18.65	effect	12.90
difference	6.32	component	15.41	feature	16.57	problem	15.06	increase	12.90
improvement	6.32	issue	14.81	aspect	14.66	issue	13.63	similarity	12.40
amount	5.92	point	14.81	issue	14.02	change	12.91	feature	10.91
variable	5.53	part	14.22	part	13.39	result	12.19	region	10.42
contribution	5.53	improvement	13.63	claim	10.20	component	12.19	change	9.92
observation	5.13	question	13.04	discussion	10.20	step	11.48	band	9.42
function	4.74	factor	12.44	argument	9.56	number	10.76	amount	8.93
packet	4.74	advantage	11.85	number	8.92	part	10.76	step	7.94
class	4.34	idea	10.67	way	8.92	pathway	10.76	source	7.94
step	4.34	type	10.07	position	8.29	cluster	9.32	function	7.44
part	3.95	context	10.07	problem	8.29	idea	9.32	level	7.44
point	3.55	property	9.48	constraint	8.29	aspect	8.61	regulator	6.45
way	3.16	contribution	8.89	exception	7.65	effect	7.89	decrease	6.45

Table 5: Targets evaluated with importance across five disciplines

the top 10 and *improvement*, *point*, *problem* and *question* in the top 20).

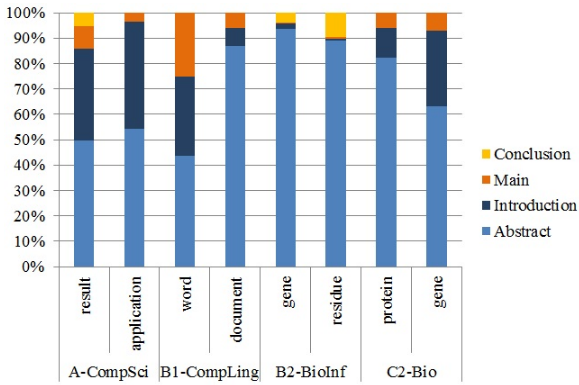

Figure 7: Domain-specific targets of four disciplines across document sections in DaSciTex

If we look at the domain-specific targets across sections evaluated with importance, we observe that they occur most often either in the Introduction or the Abstract (see Figure 7). According to previous studies on SciTex (Degaetano-Ortlieb et al., 2013), these targets mostly form keywords in the specific discipline, which indicates that nominal targets evaluated by importance patterns in most disciplines seem to be topic indicators. Note that this does not mean that they are absent from the Conclusion, they can be evaluated with other meanings (e.g., *application* with *complex*) or the targets change into hyponyms becoming more specific (e.g., in the case of specific genes).

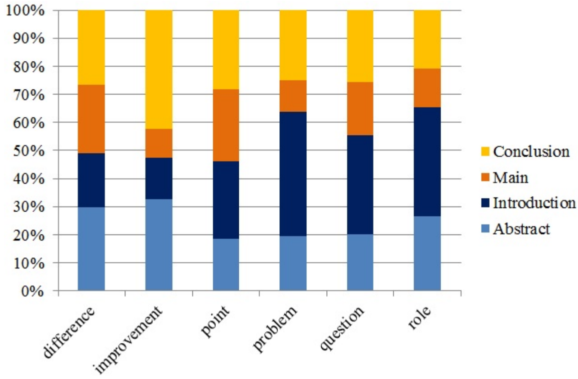

Figure 8: General targets across document sections in DaSciTex

The more general targets shared across disciplines, instead, show some individual tendencies (see Figure 8). The targets *difference* and *point* are distributed relatively evenly across the sections; *problem*, *question* and *role*, instead, are most frequently used within the Introduction, and *improvement* is most frequently used in the Conclusion but also in the Abstract. What these general targets have in common is that they relate to a more specific target. In the case of an *improvement*, the noun itself bears also an evaluation that is attributed to a target, as in Example (10) where the actual target is *combinatorial algorithms*. When we consider *question*, which is most often used in the Introduction (similarly to *problem* and *role*), it relates mostly to research questions authors of research articles pose and emphasize to be important for their study (see Example (11) and (12)). The general target *point*, which is quite evenly distributed across sections (similarly to *difference*), makes this even more clear, as *point* itself is somehow an 'empty' target. The actual target of the evaluation is what follows the relational verb, which is either a clause or a nominal phrase (see Example (13) and (14), respectively).

Another general target used similarly to *point* mostly in a relational construction is *role*. More than 70% (319 out of 446) of *role* are used within the fixed expression *to play an important role*, even though the adjective might vary. In this case, the actual target precedes the importance expression (see Example (15)). Thus, *role* has a more standardized structure than *point* which shows more variation.

(10) *From this study we conclude that* $[_{\text{target-np}}$ *the combinatorial algorithms*$]$ [...] $[_{\text{v}}$ *provide*$]$ $[_{\text{eval-np}}$ *significant improvement*$]$.

(11) $[_{\text{eval-np}}$ *Our second major research question*$]$ $[_{\text{rel-v}}$ *is*$]$ *as follows:* $[_{\text{target-np}_{\ldots}}]$.

(12) $[_{\text{eval-np}}$ *The most crucial question*], *in our view*, $[_{\text{rel-v}}$ *is*] $[_{\text{target-clause}}$ *whether a template-based NLG system can ...*].

(13) $[_{\text{eval-np}}$ *The main point*]$[_{\text{rel-v}}$ *is*] $[_{\text{target-clause}}$ *not to dwell on the shortcomings of the individual systems, but to ...*].

(14) $[_{\text{eval-np}}$ *One key point in interoperability*] $[_{\text{rel-v}}$ is] $[_{\text{target-np}}$ *enterprise modeling*].

(15) *Observe that* $[_{\text{target-np}}$ *the meaning of the term Ni (m(j) = i) in G3*] $[_{\text{v}}$ *plays*] $[_{\text{eval-np}}$ *an important role in the algorithm*].

If we consider the distribution of *role* used with importance across disciplines, it is most frequently used in biology (C2) with 90.77 per 1M and least often in computer science (A) with 8 per 1M. However, considering how often the fixed expression *play an imp-ADJ role* is used, computer science (A) uses it most frequently (approx. 81%), while biology (C2) uses it less frequently (approx. 64%). Thus, biology (C2) makes a more varied use of *role*+importance than computer science (A).

In summary, we can say that academic disciplines (a) differ in the amount of evaluations of importance, (b) use different amounts of importance across document sections, and (c) show lexico-grammatical variation in terms of evaluative attribution structures and evaluated targets.

4 Conclusion and Envoi

We have presented a methodology to approach the detection of evaluative expressions and targets evaluated on a semi-automatic basis. The manual annotation led the way to formulate rules for the automatic detection of evaluative expressions and targets. Having the corpus annotated with evaluation patterns and meanings enables further investigations.

In our case, we have focused on the notion of importance in academic research articles. In linguistic terms, we have seen that only particular lexical items and structures are used to express importance. Considering document sections, Introduction and Conclusion are the most evaluative sections, yet the importance meaning is mostly expressed at the beginning of research articles. In terms of nominal targets, we have seen that some general targets are shared across disciplines in SciTex and that they function almost as a placeholder. Nominal domain-specific targets instead are evaluated with importance mostly in the Introduction and Abstract. Thus, we have gained knowledge on how importance is expressed, where it lies and what it evaluates. Furthermore, we have seen how the use of evaluative expressions might vary according to the situational context, i.e. academic disciplines.

In future work, we aim to investigate more closely full nominal targets as well as clausal targets across sections and disciplines and to annotate them into the corpus as well as cover other evaluative meanings.

Knowledge on evaluative patterns may also improve approaches in sentiment analysis, especially the classification approach in which extraction pattern learning algorithms

may profit from additional input.
Knowledge about the contextual configuration of evaluative expressions may provide further useful information. Considering academic writing, different disciplines make use of particular conventions of linguistic feature sets used in that specific situational context. Knowledge on features involved in the formation of these conventions can be extremely valuable in automatic text classification approaches (Teich et al., 2013; Whitelaw and Argamon, 2004). Additionally, the methodology can be adapted for other genres to give similar insights.

References

Biber, D., Johansson, S., and Leech, G. (1999). *Longman Grammar of Spoken and Written English*. Longman, Harlow.

CWB (2010). The IMS Open Corpus Workbench. http://www.cwb.sourceforge.net.

Degaetano, S. (2010). Evaluation in Academic Research Articles across Scientific Disciplines. Master's thesis, Technische Universität Darmstadt.

Degaetano-Ortlieb, S., Hannah, K., Lapshinova-Koltunski, E., and Elke, T. (2013). SciTex – a diachronic corpus for analyzing the development of scientific registers. In Bennett, P., Durrell, M., Scheible, S., and Whitt, R. J., editors, *New Methods in Historical Corpus Linguistics*, Corpus Linguistics and Interdisciplinary Perspectives on Language (CLIP), Vol. 3. Narr, Tübingen.

Degaetano-Ortlieb, S., Teich, E., and Lapshinova-Koltunski, E. (2012). Domain-specific variation of sentiment expressions: exploring a model of analysis for academic writing. In *1st Workshop on Practice and Theory of Opinion Mining and Sentiment Analysis (PATHOS) at Konvens2012*, Vienna.

Evert, S. (2005). *The CQP Query Language Tutorial*. IMS Stuttgart. CWB version 2.2.b90.

Hood, S. (2005). Managing attitude in undergraduate academic writing: a focus on the introductions to research reports. In Ravelli, L. J. and Ellis, R. A., editors, *Analysing Academic Writing. Contextualized Frameworks*. Continuum, London & New York.

Hundt, M., Sand, A., and Siemund, R. (1999). *Manual of Information to Accompany The Freiburg – LOB Corpus of British English ('FLOB')*. Freiburg: Department of English, Albert-Ludwigs-Universität Freiburg.

Hunston, S. (2004). Counting the uncountable: problems of identifying evaluation in a text and in a corpus. In *Corpora and Discourse*, pages 157–188. Peter Lang.

Hunston, S. (2011). *Corpus approaches to evaluation: phraseology and evaluative language*. Taylor & Francis, London.

Hunston, S. and Francis, G. (2000). *Pattern Grammar: A Corpus-driven Approach to the Lexical Grammar of English*. Studies in Corpus Linguistics. John Benjamins Publishing, Amsterdam/Philadelphia.

Hyland, K. (2005). Stance and engagement: a model of interaction in academic discourse. *Discourse Studies*, 7(2):173–192.

Kermes, H. (2003). *Off-line (and On-line) Text Analysis for Computational Lexicography*. PhD thesis, Universität Stuttgart.

Kučera, H. and Francis, W. N. (1967). *Computational Analysis of Present-Day American English*. Brown University Press, Providence, RI.

Martin, J. R. and White, P. R. (2005). *The Language of Evaluation, Appraisal in English*. Palgrave Macmillan, London & New York.

Nwogu, K. N. (1997). The medical research paper: Structure and functions. *English for Specific Purposes*, 16(2):119–138.

O'Donnell, M. (2008). The UAM CorpusTool: Software for corpus annotation and exploration. In *Proceedings of the XXVI Congreso de AESLA*, Almeria, Spain.

Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R., and Scheffczyk, J. (2010). Framenet II: Extended theory and practice. Technical report, ICSI.

Schmid, H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. In *International Conference on New Methods in Language Processing*, pages 44–49, Manchester, UK.

Swales, J. M. (1990). *Genre Analysis. English in academic and research settings*. Cambridge University Press, Cambridge.

Teich, E., Degaetano-Ortlieb, S., Kermes, H., and Lapshinova-Koltunski, E. (2013). Scientific registers and disciplinary diversification: a comparable corpus approach. In *Proceedings of 6th Workshop on Building and Using Comparable Corpora (BUCC)*, Sofia, Bulgaria.

Teich, E. and Fankhauser, P. (2010). Exploring a corpus of scientific texts using data mining. In Gries, S., Wulff, S., and Davies, M., editors, *Corpus-linguistic applications: Current studies, new directions*, pages 233–247. Rodopi, Amsterdam & New York.

Whitelaw, C. and Argamon, S. (2004). Systemic functional features in stylistic text classification. In *AAAI Fall Symposium Series*, Washington D.C., USA.

Wilson, T. A. (2008). *Fine-grained Subjectivity and Sentiment Analysis: Recognizing the Intensity, Polarity, and Attitudes of Private States*. PhD thesis, University of Pittsburgh.