



## **Discours**

Revue de linguistique, psycholinguistique et informatique. A journal of linguistics, psycholinguistics and computational linguistics

**16 | 2015**  
**Varia**

---

# Evaluative Meaning and Cohesion: The Structuring Function of Evaluative Meaning in Scientific Writing

**Stefania Degaetano-Ortlieb**

---



### **Electronic version**

URL: <http://journals.openedition.org/discours/9053>

DOI: 10.4000/discours.9053

ISSN: 1963-1723

### **Publisher:**

Laboratoire LATTICE, Presses universitaires de Caen

### **Electronic reference**

Stefania Degaetano-Ortlieb, « Evaluative Meaning and Cohesion: The Structuring Function of Evaluative Meaning in Scientific Writing », *Discours* [Online], 16 | 2015, Online since 09 September 2015, connection on 30 April 2019. URL : <http://journals.openedition.org/discours/9053> ; DOI : 10.4000/discours.9053

---



*Discours* est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.





Revue de linguistique, psycholinguistique et informatique

<http://discours.revues.org/>

## **Evaluative Meaning and Cohesion: The Structuring Function of Evaluative Meaning in Scientific Writing**

---

Stefania Degaetano-Ortlieb  
Universität des Saarlandes (UdS)  
[s.degaetano@mx.uni-saarland.de](mailto:s.degaetano@mx.uni-saarland.de)

.....  
Stefania Degaetano-Ortlieb, «Evaluative Meaning and Cohesion: The Structuring Function of Evaluative Meaning in Scientific Writing», *Discours* [En ligne], 16 | 2015, mis en ligne le 9 septembre 2015.

.....  
URL : <http://discours.revues.org/9053>

.....  
Titre du numéro: *Varia*  
Coordination: Lydia-Mai Ho-Dac et Frédéric Landragin

**revues.org**  
CENTRE POUR L'ÉDITION ÉLECTRONIQUE OUVERTE  
CENTRE FOR OPEN ELECTRONIC PUBLISHING

 discours

 Presses  
universitaires  
de Caen



# Evaluative Meaning and Cohesion: The Structuring Function of Evaluative Meaning in Scientific Writing

---

Stefania Degaetano-Ortlieb

Universität des Saarlandes (UdS)

.....  
We present a diachronic study of evaluative meaning in scientific writing, focusing on evaluative expressions that possibly serve the interpersonal as well as the textual metafunction in terms of Systemic Functional Linguistics (SFL). These are expressions such as *importantly* or *obviously* used in sentence-initial position to evaluate what follows but which also establish a cohesive link with the adjacent discourse. For the analysis, the SciTex corpus, comprising nine scientific disciplines, was used. The data were analyzed in macro and micro-analytical steps combining quantitative and qualitative analyses. This allows us to observe generalizable trends as well as fine-grained distinctions.

**Keywords:** evaluative meaning, discourse relations, text organization, corpus linguistics, data mining

## 1. Introduction

1 While evaluative meaning is considered to be an interaction between participants putting forward the feelings, attitudes etc. of a speaker or writer to be perceived by a hearer or reader (see, e.g., Biber & Finegan, 1989; Halliday, 2004; Hunston & Thompson, 2000; Hunston, 2011; Hyland, 2005), it has also been argued that specific evaluative expressions (such as *unfortunately*, especially in sentence-initial position) can also serve text organization purposes (see, e.g., Thompson & Zhou, 2001; Ho-Dac et al., 2012), thereby contributing to textual cohesion. Cohesion is the explicit textual marking that a writer performs to create coherence in a text. One of the cohesive devices that realize cohesion is conjunctive items (such as *and*, *so*, *but*, *however*), used to mark explicitly the type of connection that is set up between propositions (e.g., additive or adversative connection types). Thompson and Zhou (2001) argue that besides this more traditional “propositional coherence” signaled in the text by cohesive items such as conjunctions, there is also what they call “evaluative coherence” signaled by evaluative items (such as disjuncts or modal adjuncts like *certainly*) (cf. Hunston & Thompson, 2000: 122).

2 In Systemic Functional Linguistics (SFL) (Halliday, 2004), cohesion belongs to the textual metafunction, while any kind of interaction between participants – evaluative meaning included – belongs to the interpersonal metafunction. A third metafunction in SFL is the ideational metafunction, concerned with the propositional content. While the metafunctions are separate, a clause realizes all three. What Thompson and Zhou (2001) suggest is that there are some evaluative lexical items that can belong to the interpersonal as well as the textual metafunction.

Other authors similarly suggest that certain devices are multi-functional in that they not only convey content (i.e., ideational meaning) but are also involved in text organization (see, e.g., Ho-Dac et al., 2012 on enumeratives such as *first*, *second*).

3 This multi-functional view is connected to the notion of clause relations pioneered by Winter (1977, 1982, 1994) and Hoey (1983). Clause relations represent a kind of dialogue/interaction between the writer and reader (cf. Hunston & Thompson, 2000: 121). They indicate a cognitive process in which the reader interprets the meaning of a sentence (or group of sentences) in the light of its adjoining sentence (or group of sentences) (cf. Winter, 1994: 49). In this sense, some evaluative lexical items also serve a logical-connection function as they connect the previous and the following discourse in a logical sense contributing to the coherence of the text but preserving their evaluative meaning. These logical connections can be of different relation types. Thompson and Zhou (2001) consider, e.g., concessive relations (realized, e.g., by *certainly* and *plainly*), expectancy relations (realized, e.g., by *unfortunately*), and alternative relations (realized, e.g., by *perhaps* and *maybe*).

4 There has been a considerable body of work on evaluative meaning in academic discourse (e.g., Anderson & Bamford, 2004; Stotesbury, 2003 and 2006; Sanz et al., 2010), and previous studies have shown that scientific disciplines vary in their use of evaluative meaning (see, e.g., Hunston, 2004; Hyland, 2005; Degaetano & Teich, 2011; Degaetano-Ortlieb et al., 2014; Degaetano-Ortlieb, 2015). In this study, we aim to investigate, whether the types of realizations of evaluative meaning used in scientific writing have, in addition to their evaluative function, a textual function as well, contributing to the cohesiveness of a text. Here we will consider two evaluative meanings, obviousness and importance, which differ in their evaluative function as well as in their pragmatic and rhetorical properties (see e.g. Eason, 2001).

5 More specifically, we ask whether scientific disciplines make use of the multi-functionality of specific evaluative items. To answer this question, first, we adopt a macro-analytic perspective to account for general trends. Second, we consider the syntagmatic context by adopting a micro-analytic perspective to account for differences related to evaluative expressions used in sentence-initial position, drawing on Thompson and Zhou's (2001) contention that evaluative lexical items in sentence-initial position can fulfil a cohesive function as well<sup>1</sup>. In this micro-analysis, we consider (1) lexico-grammatical differences, i.e. how the evaluative meaning is realized (as an adverb such as *obviously*, an evaluative pattern such as *it is clear that* or a noun phrase *clear tendencies*), and (2) differences in the use of modification or specific type of logical relations used in combination with these realizations, i.e., for instance, how the evaluative meaning is modified (e.g., by intensifiers such as *most*

---

1. It must be emphasized here that the analysis presented is exploratory. Thus, we restrict ourselves to the examination of evaluative lexical items with a possible cohesive function in sentence-initial position only, as non-initial items are affected by other factors. The latter have been deliberately excluded so as to gain a clear understanding of the multi-functionality of these evaluative lexical items first, moving to more complex cases in the future.

or *more* for adverbs). For both, we will consider the implications these differences have on the multi-functional view of specific evaluative lexical items, i.e. whether there is a preference in the use of multi-functional evaluative items over other alternatives in order to enforce coherence in a text, or whether this multi-functional use is restricted to specific evaluative expressions as pointed out by Thompson and Zhou (2001). From this we can generate the following main hypothesis:

*Evaluative coherence:*

Evaluative lexical items will show a possible multi-functionality, fulfilling an evaluative as well as a textual function. They will possibly show differences on the level of lexico-grammar and usage, which might be related to their semantic type.

6 We will investigate this main hypothesis by adopting corpus-based methods, which consist of the annotation and extraction of evaluative meaning from a corpus to gather appropriate data, which will serve to create distributional information and to observe possible fine-grained differences on the level of lexico-grammar in scientific writing.

7 The paper is structured as follows. First, we will present the descriptive framework by describing evaluative meaning and its relation to cohesion (Section 2). Second, we will present the SciTex corpus as well as the corpus processing and the analytical cycle with its macro- and micro-analytical steps (Section 3). In Section 4, we present the analyses considering diachronic tendencies as well as fine-grained differences for selected realizations of evaluative meaning that possibly show multi-functionality in terms of evaluative and textual function. We conclude with a summary and an envoi (Section 5).

## 2. Descriptive framework

### 2.1. Evaluative meaning

8 There has been considerable work on evaluative meaning, which has pointed to a number of linguistic resources that express evaluative meaning in scientific writing (e.g., Hyland, 2005; Hyland & Tse, 2004; Swales, 1990; Thompson, 2001). Yet, only selective aspects have been investigated and only in the last decade or so have attempts been made to move towards creating models of analysis that try to integrate the whole range of the potential of evaluative meaning (the pioneers here are Hunston & Thompson, 2000; Hyland, 2005; Martin & White, 2005). In Degaetano-Ortlieb (2015) an attempt was made to formulate a model of analysis of evaluative meaning for scientific research articles, adopting some already established ideas and methodologies and combining them with new insights, grounding the model within Systemic Functional Linguistics. For a full description of the framework, see Degaetano-Ortlieb (2015: 35ff.). Here, we introduce the main concepts relevant for the analysis.

9 When an evaluative act is performed the writer expresses an evaluation toward some content with a persuasive intention toward the reader. Clearly, this is related

to the interpersonal metafunction in SFL terms, which relates to the interaction between participants (e.g. writer and reader or hearer and listener) in a discourse and the exchange of meanings during interaction, here evaluative meanings. One important aspect of evaluative acts is the potential for arguing, i.e. an exchange of speech roles among the interactants by which a dialog is formed. With respect to research articles, a dialog is formed between the writer(s) and the reader(s). Evaluative meaning is mostly concerned with an exchange of information rather than an exchange of goods-&-services, because the latter, being either command or offer, is usually not arguable. Thus, when we consider evaluative meaning, we are dealing with an exchange of information that is arguable. The argumentative scope ranges from positive to negative polarity, yet it is not restricted to these two poles as it offers a whole range of semantic meanings, which lie between these two, expressing the writer's stance. The semantic meanings expressed can be either of the epistemic or the attitudinal type. Epistemic meanings express modality (e.g., certainty or likelihood; see example [1] and [2], respectively), while attitudinal meanings can be of various types (expressing, e.g., desirability, importance, benefit; see examples [3]-[5]). Note that both epistemic and attitudinal meanings can range from a positive to a negative pole, as can be seen in example [4], where *unfortunately* lies at the negative pole of the desirability meaning.

- [1] This *certainly* constitutes a limiting factor, since the aim of the controller resides precisely in vibration suppression.  
(SciTex; electrical engineering)
- [2] This variation may reflect real structural differences found in mouse ribosomal genes or *possibly* deletion events which occurred during cloning.  
(SciTex; biology)
- [3] These residues are *important* in forming the heme contacts as well as  $\alpha$ -3-cooperative dimer associations.  
(SciTex; biology)
- [4] *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.  
(SciTex; computational linguistics)
- [5] The primary *advantage* of the CSA was that it provided a simple qualitative picture of temporal variations in the EEG power spectrum.  
(SciTex; bioinformatics)

10 Besides the different kinds of meanings used to evaluate, when analyzing evaluative meaning, one is also interested in differences related to the entities that are involved in an evaluative act. By entities we relate, first of all, to the participants in the discourse, i.e. writer and reader. Moreover, we also consider the entity or process that the



writer evaluates, i.e. the target of the evaluation (cf. Hunston & Sinclair, 2001). In example [6], *the approach* constitutes the target, which is evaluated as being appropriate.

[6] [...] [<sub>target-noun-phrase</sub> the approach] [<sub>relational-verb</sub> is] [<sub>evaluative-adjective</sub> appropriate].

- 11 Whenever an evaluative act is performed, the target as well as the stance expression (epistemic or attitudinal) are present in the discourse. Semantically, target and stance expression are closely related, as the evaluation expressed is always directed toward a target. In SFL terms, the semantic relation is visible in the lexico-grammar, where the realizations of stance expression and target appear together in lexico-grammatical patterns as in example [6], where we have a relational pattern with a target noun phrase, followed by a relational verb and an evaluative adjective. In lexico-grammatical terms, a target can either precede or follow the evaluative expression. In example [6], the target precedes the evaluation, while in example [7] the target follows the evaluation. From example [7], we can also see that a target can also be constituted by a clause, which is evaluated by an evaluative adverb that precedes the clause.

[7] [<sub>evaluative-adverb</sub> Importantly], [<sub>target-clause</sub> it also permits a neat interface] [...].

- 12 Given this framework, we can analyze evaluative meaning, its realizations and lexico-grammatical usage in scientific writing.

## 2.2. Evaluative meaning and cohesion

- 13 As we are also interested in observing evaluative expressions that may fulfil both an interpersonal and a textual function in this paper, we will focus on evaluative expressions in sentence-initial position, such as the ones in examples [7]-[9], to better understand this kind of multi-functionality before moving to more complex cases (such as those in non-initial position) in future work. Note that examples [7]-[9] all express the meaning of importance. However, they are realized differently in terms of lexico-grammar. In example [7] the importance meaning is realized by an adverb (*importantly*), in example [8] by an evaluative noun phrase (*one key output variable*), and in example [9] by an it-pattern (*it is essential* + that-clause). Besides the importance meaning, other meanings can be expressed, such as the obviousness meaning as shown in examples [10]-[12].

[8] [<sub>evaluative-noun-phrase</sub> One key output variable] [<sub>relational-verb</sub> is] [<sub>target-noun-phrase</sub> area A1 in Fig. 17].

[9] [<sub>it</sub> It] [<sub>relational-verb</sub> is] [<sub>evaluative-adjective</sub> essential] [<sub>target-clause</sub> that the train and test set are identical].

[10] [<sub>evaluative-adverb</sub> Clearly], [<sub>target-clause</sub> syntactic information plays a central role in establishing appropriate referents].

[11] [<sub>evaluative-noun-phrase</sub> An obvious aspect] [<sub>relational-verb</sub> is] [<sub>target-clause</sub> that the alternative execution path is shorter than the normal pipeline].

[12] [<sub>it</sub> It] [<sub>relational-verb</sub> is] [<sub>evaluative-adjective</sub> obvious] [<sub>target-clause</sub> that the self-intersection occurs beyond the neighboring facets].

- 14 If we consider more closely the surrounding context in which example [10] is positioned (see example [10a]), we can observe, similarly to Thompson and Zhou (2001), how a concessive relation is built up in the discourse. Two propositions are presented, the first introduced with *clearly*, where the writer explicitly marks a specific stance, the second one is introduced by *nevertheless*, presenting the proposition as an unexpected consequence of the first. Even though both propositions are presented as valid, the second one presents the writer's own view. As Thompson and Zhou (2001: 126) point out, the first proposition is conceded, whereas the second is asserted. If the adverb *clearly* were left out, there would still be some kind of concessive relation marked by *nevertheless*. However, *clearly* gives more weight to the concessive relation, as it is involved in the argumentative structure of the discourse, thus fulfilling also a cohesive function.

[10a] Among the few systematic methods which have been proposed for anaphora resolution [...], most are conceptualized within symbolic frameworks as in the cases with other syntactic processing. **Clearly**, syntactic information plays a central role in establishing appropriate referents. Nevertheless, it has long been recognized that the traditional symbolic syntactic approach, which tries to capture the meaning from antecedent linguistic items in text, cannot be translated into a processing theory [...].

(SciTex; computational linguistics)

- 15 It seems that this kind of cohesive function is not limited to adverbs, as can be seen from looking at a broader context of example [11] where the evaluative noun phrase *an obvious aspect* is used to express obviousness (see example [11a]). We can see how the concessive relation is built up between the evaluative noun phrase and *however*. The first introduces the advantages, the second instead points to the limitations, as was pointed out in the first sentence by the writer. Yet, without the noun phrase introducing the advantages, the structure given to these two counterparts would have been missing.

[11a] From a performance point of view, the proposed microarchitecture has both advantages, as well as limitations. **An obvious aspect** is that the alternative execution path is shorter than the normal pipeline. This reduces the mispredict penalty when the next trace is found in the EC, and becomes a definite advantage when executing programs with bad branch predictability. As the current trend is to use deeper pipelines, the benefit is likely to increase. However, although placing the EC deep in the pipeline limits the branch mispredict penalty, there are some caveats associated to it. Each time instructions are issued from the EC, the normal branch predictor from the fetch stage is not used [...].

(SciTex; microelectronics)

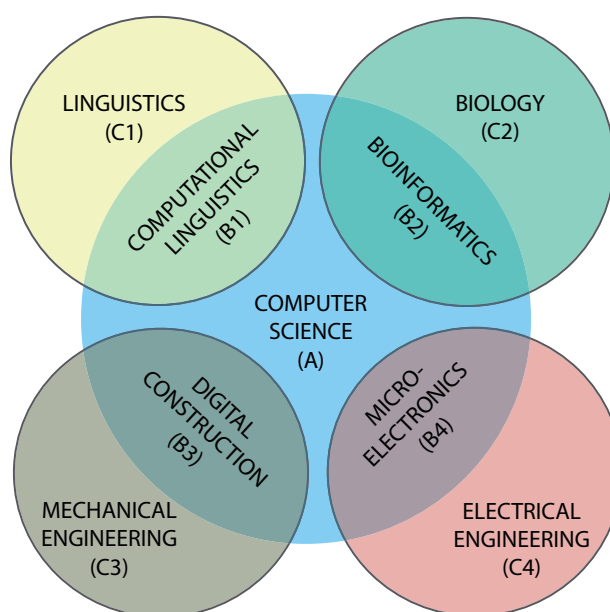


Figure 1. Scientific disciplines in the SciTex corpus

16 In our analysis, we will consider whether there are differences in terms of how this kind of information is presented and whether there are preferences for a particular lexico-grammatical realization.

### 3. Corpus and analytical cycle

#### 3.1. Corpus resource

17 To investigate the hypothesis formulated above, we used the SciTex corpus (“English Scientific Text Corpus”; Teich & Fankhauser, 2010; Degaetano-Ortlieb et al., 2013). The corpus was specifically built to investigate register formation processes in scientific writing, focusing on register contact between computer science and other selected scientific disciplines (see Figure 1). The core idea associated with register is that language use is systemically influenced by contextual factors (Halliday & Hasan, 1985), i.e. the situational context influences the linguistic choices made by a language user and vice versa. According to Halliday (1988), registers are clusters of associated features that have a greater-than-random tendency to co-occur. The corpus covers nine scientific disciplines: computer science (A-subcorpus), computational linguistics, bioinformatics, digital construction and microelectronics (B-subcorpus) as well as linguistics, biology, mechanical engineering and electrical engineering (C-subcorpus). The B-subcorpus represents the *contact disciplines*, while the A- and C-subcorpora represent the disciplines of origin (*seed disciplines*). The corpus contains approximately 34 million words and comprises two time periods,

the 70/80s (SaSciTex) and the early 2000s (DaSciTex), covering a thirty-year time span similarly to the Brown corpus family (Kučera & Francis, 1967; Hundt et al., 1999). SciTex encompasses full English journal articles from at least two different journals per discipline and has been annotated on the level of tokens, lemmas and parts-of-speech using TreeTagger (Schmid, 1994). In addition, each document has been enriched with meta-information (such as author(s), title, scientific journal, academic discipline, and year of publication) and document structure (e.g., abstract, introduction, sentence boundaries). The corpus is encoded in Corpus Query Processor (CQP) format (Evert & Hardie, 2011) and can be queried by using regular expressions in combination with positional (e.g., part-of-speech) and structural (e.g., sentence boundary) attributes. Moreover, we have annotated the corpus using a dedicated annotation procedure (cf. Kermes & Teich, 2012) on the level of evaluative meaning (see Section 3).

### 3.2. Analytical cycle and techniques

#### 3.2.1. Analytical cycle: macro- and micro-analysis

18 Besides the corpus, we need an appropriate methodology to test our hypothesis. For this, we have designed a macro- and micro-analytical cycle with recursive steps, which will allow us to consider generalizable trends, when we look at differences with respect to the realizations used, but also to consider specific trends, when we look at fine-grained differences of each realization type. The concept of macro- and micro-analysis was introduced by Jockers (2013), where macro-analysis is concerned with quantitative findings, whereas micro-analysis is concerned with qualitative findings gained by what Jockers (2013) terms “close-reading”.

19 The cycle starts on the macro-level, where we want to test whether there are differences in terms of the realization types used (e.g., an evaluative noun phrase such as *an obvious fact* vs. an adverb such as *obviously*). To accomplish this, we first employed corpus-based methods for the annotation and extraction of evaluative meaning (see details in Section 3.1.2). In a second step, we undertook a quantitative analysis of the findings looking at the distributional information of the realization types, e.g., whether an evaluative adverb is used more frequently than an evaluative noun phrase.

20 On the micro-level, we conducted a “close-reading” of different realization types, i.e. we inspected specific differences in the use of evaluative meaning in their syntagmatic context. In this paper, we focus on differences regarding the evaluative expressions that may also perform a textual function, i.e. contribute to the cohesion of the text. It is important to note here that as this is an exploratory analysis, the part considering “close-reading” will not be complemented by quantitative observations. This will be pursued in further studies. Based on the observations gained in this study, we hope to be able to design in future work a systematic way of quantifying these observations.

### 3.2.2. Corpus-based methods

- 21 To annotate the full version of SciTex, we used annotation procedures derived from the YAC recursive chunker (Kermes, 2003), in which macros are manually defined for each feature and which are then used to automatically annotate the whole corpus. For the annotation, the “Corpus Workbench” (CWB, 2010) was used. Three feature sets were annotated.
- 22 In the first set, *stance features* were annotated. Appropriate lexicons were built with a threefold procedure: (1) use of lexical items listed in the Frame Index in FrameNet (Ruppenhofer et al., 2010) to capture possible meanings expressed to evaluate; (2) extraction of lexical items from a small annotated version of our corpus (amounting to approximately 52,000 tokens); and (3) use of WordNet (Fellbaum, 1998) to find synonyms for the lexical items obtained from (1) and (2). The lexicons were built for 35 meanings (e.g. assumption, likelihood, obviousness, benefit, importance, suitability). The size of the lexicons ranges from 2 up to approximately 70 lexical items per meaning. In the second set, *evaluative lexico-grammatical patterns* were annotated, such as the ones shown in examples [7]-[12], i.e. patterns where an evaluation is attributed to a target. These patterns were defined within the scope of a manual annotation of an extract of SciTex of approximately 52,000 tokens. Altogether, 17 evaluative patterns were annotated (see Degaetano-Ortlieb, 2015 and Degaetano et al., 2014, for a full list of the patterns). In the third set, *targets of the evaluation* were annotated.
- 23 The annotation procedure involved the use of (1) queries as rules based on part-of-speech tags and structural attributes that search for a defined feature in the corpus, and (2) Perl scripts that delimit the range of the features found, if necessary, and define the attributes to be annotated. The rules were defined manually and were used to automatically annotate the whole corpus. Results were evaluated for precision in the small version of SciTex 2000s (one million words). In the event of low precision, the rules were refined to obtain the best possible results. This procedure was particularly important for multiple-word features (such as the evaluative patterns) in order to obtain satisfactory results (cf. Degaetano-Ortlieb, 2015). Precision for all patterns ranged from 94.24% to 100%. Consider, for example, the query in Figure 2, a query macro that is used to annotate an evaluative adverb followed by a target-clause (e.g., *Importantly, the algorithm...*). In this query, the list of evaluative adverbs (`$eval-adv`) was used to search for lemmas as well as the part-of-speech information (RB for adverbs) within a sentence structure (`<s>` to `</s>`), and possible comparative and superlative adverbs preceding the adverb (such as *more* or *most*). The annotation rule shown in Figure 3 was then used to delimit the range to be annotated. In this case, the end of the structure to be annotated was set to `-1` token to exclude the token of the annotation macro used at the end, which in this case is the sentence ending token (e.g., a full stop). In addition, the rule defines which attributes should be annotated with the pattern (such as the evaluation set, pattern, meaning, etc.). These rules serve then to annotate the corpus automatically. This produces in CQP the XML-structure shown in Figure 4.

```

MACRO eval-adv-target (0)
(
  #eval-ADV + target-clause
  (<s>[pos="RB(R|S)"]? [pos="RB" & lemma=$eval-adv] []* </s>)
)
;

```

Figure 2. Example of an annotation macro

```

Macro("/eval-adv-target[]", sub{
  $end = $end-1;
  Struc(
    "evaluation" => undef,
    "evaluation_set" => "eval_target",
    "evaluation_pattern" => "eval-adv_target-clause",
    "evaluation_meaning" => "obviousness",
    "evaluation_precision" => "100%"
  );
});

```

Figure 3. Example of an annotation rule

```

<s>
  <evaluation
    set="eval_target"
    pattern="eval-adv_target-clause"
    meaning="obviousness"
    sem="attitudinal"
    precision="100%">
    Clearly, syntactic information plays a
    central role in establishing appropriate
    referents
  </evaluation>
  .
</s>

```

Figure 4. Example of the XML annotation in CQP

24 The annotation was performed for each feature (stance, evaluative patterns and target features) on different annotation layers, each with attributes of their own. This allows one to perform queries on multiple annotation layers, so that one can search, for example, for evaluative patterns with specific meanings (e.g., only the importance meaning) or used in specific document sections (Abstract, Introduction, etc.) (cf. Degaetano-Ortlieb, 2015).

25 To be able to perform a quantitative analysis, we then extracted the annotated features from the corpus along with their distributional information. The xml-tags encoded in CQP were used for this purpose. The adverbial pattern, for example, can be queried by the following command:

```

<evaluation>[_.evaluation_pattern="eval-adv_target-clause"]+
</evaluation>;

```

26 To extract all features simultaneously along with their distributional information, the extraction process was automated in an extraction pipeline, where the extraction commands as well as the attributes on which the distributional information should be performed are defined (e.g., on the attribute document section to obtain information on how often the adverbial pattern occurs in the abstract or introduction; see Kermes & Teich, 2012 for a detailed description of the pipeline).

27 The extracted information was then used to perform, first, quantitative analyses on the macro-level in order to inspect differences in the realization types, and second, qualitative analyses by looking at concordances of the extractions to inspect differences related to the multiple functions (evaluative and cohesive) of the realizations.

#### 4. Analysis

28 In the following sections, we will investigate our hypothesis of *evaluative cohesion* on the macro- and micro-level of analysis. In macro-analytical terms, we want to consider general tendencies in the distribution of different realization types. Here, we also account for diachronic differences between the two time periods, the 70/80s and 2000s, in SciTex in order to investigate whether there is a diachronic change regarding the usage of particular evaluative meanings or realizations. As some studies (e.g., Biber, 2004; Conrad & Biber, 2000; Vis et al., 2012) have, for example, shown an increase in evaluative language over time, we examine whether this is the case for some realization types but not for others. In terms of micro-analysis, we consider differences with respect to the functionality of the realization types, which might be involved in evaluative and cohesive functions simultaneously.

29 In a previous study (Degaetano-Ortlieb, 2015), we compared evaluative meaning across scientific disciplines by using text classification, looking at how well the disciplines in SciTex could be distinguished in terms of evaluative meaning. One main observation was that there is a diachronic tendency to express evaluative meaning distinctively in sentence-initial position across disciplines. Linguistics and computer science were the disciplines that could best be distinguished by evaluative meaning. Among the lexico-grammatical features that contributed to their distinction in the top five for the 2000s was the epistemic meaning of obviousness for computer science and the attitudinal meaning of importance for linguistics, both distinctive in sentence-initial position. Thus, in the following, we examine these two meanings in greater detail in SciTex.

##### 4.1. Lexico-grammatical realizations for the evaluative meanings of obviousness and importance

30 Both meanings can be realized lexico-grammatically in a variety of ways. Those that occurred the most frequently in SciTex were the adverbial pattern (such as *Clearly, ...*), the it-pattern (such as *It is important that...*), and the noun phrase (such as *An important point is that...*).

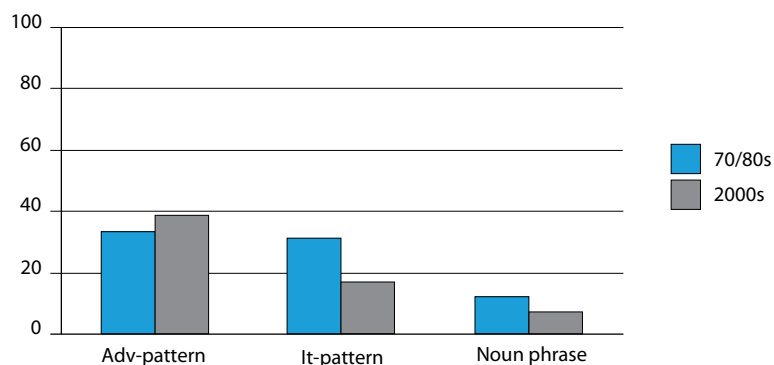


Figure 5. Realization types of the obviousness meaning in sentence-initial position in SciTex (normalized to 1 million)

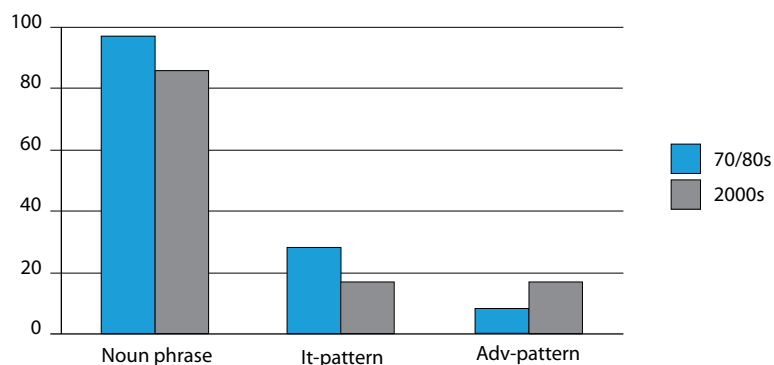


Figure 6. Realization types of the importance meaning in sentence-initial position in SciTex (normalized to 1 million)

31 Figure 5 shows these three major realization types for the obviousness meaning. It can be seen that the adverbial pattern is the most frequent type of realization, followed by the it-pattern and the noun phrase. As shown in Figure 6, the situation is quite different for the importance meaning, which is mostly realized by a noun phrase, while the it-pattern and the adverbial pattern are less frequently used. Note that the noun phrase type realizing the importance meaning has a relatively high frequency in comparison to the other two types. Diachronically, the general distributions of the types used for each meaning remain fairly stable. However, while the it-pattern and the noun phrase either decrease in use or remain stable, the adverbial pattern increases for both meanings. It shows a slight increase of 14.16% for the obviousness meaning and a relatively high increase of 61.03% for the importance meaning.

32 When we consider the evaluative function of these three realization types, they function similarly in that the evaluation is attributed to the discourse following the



evaluative expression (consider examples [13]-[15]). However, while in [13] and [14] the adverb and the it-expression are quite interchangeable with almost no change in meaning, in [15] the evaluative noun phrase could not simply be substituted by either the adverb or the it-expression. The fact that it is a consequence of Theorem 5.3 would be missing. Thus, the noun phrase is a more explicit variant, while the adverb and the it-expression are relatively general but also quite dense in terms of the information they convey. Note that the adverb is also a denser form in comparison to the it-pattern. Thus, the diachronic increase in the adverbial pattern could be related to the fact that as language evolves denser conventions are used to propagate context. According to Levy and Jaeger (2006), to optimize communication, language users reduce less information-dense phrases. Thus, although the adverb and the it-expression are almost equal in the information they convey, the denser form – the adverb – is possibly chosen over the it-expression, which is a more expanded form. This would have to be tested in information-theoretic terms and with appropriate methodologies (see, e.g., Aylett & Turk, 2004; Levy, 2008; Fankhauser et al., 2014a and b).

- [13] [<sub>evaluative adverb</sub> **Clearly**] [<sub>target-clause</sub> our technologies for storing and delivering language resources fall far short of our need for easy reuse].
- [14] [<sub>evaluative it-expression</sub> **It is clear that**] [<sub>target-clause</sub> our algorithm succeeds when the graph has some geometric structure like in “randomG4” [...]].
- [15] [<sub>evaluative noun phrase</sub> **An obvious consequence of Theorem 5.3**] [<sub>relational verb</sub> is] [<sub>target-clause</sub> that if N is any class of numerical predicates such that [...], then the CBC is false for FO[N]].

33 However, there might also be other reasons why the adverbial pattern increases over time, while the it-pattern decreases. To approach this, we looked at the syntagmatic context of both meanings for the adverb and the it-pattern.

## 4.2. Evaluative meanings and their possible cohesive function

### 4.2.1. The obviousness meaning

34 We start with the obviousness meaning realized by the adverbial pattern. Possible adverbs in this pattern are *clearly*, *conceivably*, *intuitively*, *logically*, *obviously*, *straightforwardly*. Consider example [16], where the adverb *obviously* is used. Within the sentence, the adverb is used to express evaluative meaning toward the clause that follows. The writer makes a statement that is perceived as factual by the reader because of the adverb *obviously*, namely that algorithms that produce approximators to the target  $f$  cannot be used for the task intended.

- [16] Consider the following circuit design problem: you are given a representation of some Boolean function  $f$ , and you have in mind a target class  $C$  of function representations (say DNF, or sum-of-products, representations). You would like to efficiently find a

reasonably small (by some measure) [...]. Can learning algorithms be applied to this problem? **Obviously**, algorithms that produce only approximators to the target  $f$ , such as probably approximately correct (PAC) learning algorithms [16], cannot be used for this task. On the other hand, in the traditional model of Exact learning from an equivalence oracle [1], the learning algorithm is presented with adversarially chosen counterexamples to its intermediate hypotheses, which seems to be a “harder” model of learning than is required for our problem. [...] Thus, we consider a model (introduced by Bshouty) that lies between the PAC and Exact models.

(SciTex; A-computer science, 2000s)

35 However, besides this evaluative function, by considering the wider context, we can see how the adverb also fulfils a text structuring function within the discourse. Because of the question *Can learning algorithms be applied to this problem?*, one assumes that an argumentation will follow. The adverb introduces the first argument, answering the question with a no. The conjunctive adverb *on the other hand* introduces the counterargument, even though the writers explain that this would also not be an option for their purposes. The *thus*, finally introduces the chosen option of the writers. We can see how the discourse is organized by these introductory items in sentence-initial position, i.e. in Theme position (cf. Halliday, 2004). The argumentative structure could be read as follows: “*obviously* we have  $X \rightarrow$  *on the other hand* there is  $Y \rightarrow$  which is still not an option  $\rightarrow$  *thus* we use something in between  $X$  and  $Y$ ”. Note that one could leave out the *obviously*, but then something would be missing that makes it easier for the reader to follow the argumentative structure.

36 The same holds for example [17], where the adverb *clearly* introduces something that cannot be executed, while the concessive *instead* introduces the solution the writer opted for.

[17] A final difficulty is that we are interested in a sublinear algorithm. **Clearly** we cannot fully execute such a recursive procedure (not to mention find the exact median values desired in the different levels of the recursion). Instead, we work only with estimated values, and we execute only a small random number of the recursion paths.

(SciTex; A-computer science, 2000s)

37 So far, we have considered examples from computer science. Example [18] is taken from linguistics, showing a quite similar pattern. The expectation of a two-fold argument is set up in the first sentence by claiming that there are satisfying conclusions for a distinction between semantic and onymic reference. The *clearly* introduces the first argument, while the *but* introduces a counterargument.

[18] There are satisfying conclusions to be drawn from accepting a distinction between semantic and onymic reference. **Clearly**, reference can be achieved through paying attention to the meaning of the words in a referring expression – that is, ordinary semantic reference. But if it is accepted that expressions may be used to refer unmediated by sense, even where they contain overt sense-bearing elements,

then it must be conceded that the same expressions may be used with different processing costs.

(SciTex; CI-linguistics, 2000s)

- 38 According to Thompson and Zhou (2001), in all these examples, we can see how the evaluative adverb fulfils a cohesive function, which serves to structure the discourse in terms of a particular argumentative scope (argument vs. counterargument). However, this kind of function is not the only one that we found for the obviousness meaning realized by an adverb in the SciTex corpus. There can be various cohesive functions that these adverbs can fulfil. In example [19], the *clearly* seems to be involved in a consecutive relation with *hence*, i.e. what is mentioned in the first sentence is further elaborated on in the second. In example [20], instead, the *clearly* can be said to have a deductive function, i.e. something is inferred from the first sentence in the second one. We could also rephrase the adverb as shown in example [20a]. Note that it could be argued that this rephrased structure still allows the reader to perceive both the evaluative and cohesive function. The evaluative function of putting forward something as factual can be said to be realized by *shows that*, while the cohesive function is realized by *percentages*, which relates back to the percentages mentioned in the preceding discourse. However, the adverb is a much denser version when comparing both options.

[19] **Clearly**, a solution that is not locally optimal can be improved by having misplaced vertices change sides. [...] **Hence** the approximation ratio can be improved by adding a local correction step to the algorithm that moves vertices from side to side until the solution becomes locally optimal.

(SciTex; A-computer science, 2000s)

[20] For British speakers the percentage of such choices was 9.2%; for American speakers it was 8.5%. The wide majority of these choices reflected plural attraction, 95.5% for British speakers and 93.8% for Americans. **Clearly**, the ability to recognize prescriptively correct agreement does not differ between the groups.

(SciTex; CI-linguistics, 2000s)

[20a] **A comparison of the percentages shows that** the ability to recognize prescriptively correct agreement does not differ between the groups.

- 39 Thus, while the evaluative function seems to be the same in all the examples, i.e. putting forward a statement that is perceived by the reader as being factual, the cohesive function is much more subtle and highly dependent on the preceding and/or following discourse context. Moreover, while we could rephrase the multi-functionality of these adverbs by moving towards greater explicitness with a more extended expression, the adverb allows the writer to densify the information put forward. Both the adverb's multi-functionality as well as its denser form in comparison to more explicit and longer counterparts might be the reason for their increased use over time.

40 However, it should also be noted that the adverbs do not always have a multiple function. For computer science in SciTex, the multi-functionality was mostly observed in the introduction section, while in the body of the article the adverbs seem to be part of a more formalized kind of language used within computer science (see example [21]).

[21] For simplicity, we say that the representation of  $I$  is of length  $\log k A$ . **Obviously**,  $C$  conjunctive reduces to  $C$ . Let  $\text{pdf} = 1/2 (s + 1) A (\log s + p(r))$ .

41 In the following, we consider the it-pattern expressing obviousness, which can be realized inter alia by the adjectives *apparent*, *clear*, *evident*, *obvious*, *straightforward* (the complete list comprises 10 adjectives). Interestingly, it does not show the kind of multi-functionality illustrated in the examples above. Consider example [22], where the it-pattern seems to express only the evaluative meaning of obviousness. Moreover, the it-pattern expressing the obviousness meaning in computer science is mostly found in the body of research articles. This differs from the use of the adverb, which is also found quite frequently in the introduction section, where it shows its multi-functionality contributing also to the cohesiveness of the introduction. If a cohesive function is combined with the it-pattern, this is accomplished by inserting a textual conjunct such as *however* (see example [23]).

[22] Claim 2. If  $X$  contains  $\{0,1,2,3\}$  then it contains  $u$ . Consider the self-map  $\hat{I}_{\pm}$  of  $G$  defined by  $\hat{I}_{\pm}(i) = i$  for all  $i = 0,1,2,3$  and  $\hat{I}_{\pm}(x) = x$  otherwise. **It is clear that**  $\hat{I}_{\pm}$  is a homomorphism, and one verifies easily that  $\hat{I}_{\pm}$  is adjacent to  $\text{id}$  in  $\text{Hom}(G, G)$ . Now define a 4-ary idempotent operation  $\check{I}$  on  $G$  as follows: [...].

(SciTex; A-computer science, 2000s)

[23] Thus the minimum degree of any representation has to be within a factor  $p_1$  of the degree of this polynomial. Over  $Z_m$  when  $m$  is not a prime power, however, things are very different. If  $P(X)$   $0-1$  represents the OR function over  $Z_6$ , it can be shown using the Chinese Remainder Theorem (CRT) that  $P(X)$  has degree  $(n)$ . **However, it is not clear that** one can obtain a  $0-1$  representation from a strong or a weak representation since  $Z_6$  is not a field. In fact, the degree of a strong or a weak representation can be very different from that of a  $0-1$  representation.

(SciTex; A-computer science, 2000s)

42 Another difference between the adverb and the it-pattern is related to their modification patterns and their involvement in logical relations. Figure 7 shows the percentage of modification and logical relations used for the obviousness meaning. While the adverb is almost always used in its pure form (0.7% in the 70/80s and 0.5% in the 2000s), i.e. it is neither modified nor combined with logical relations, the it-pattern shows a diachronic increase (from 9.8% to 25.2%) in combination with modification or logical relations.

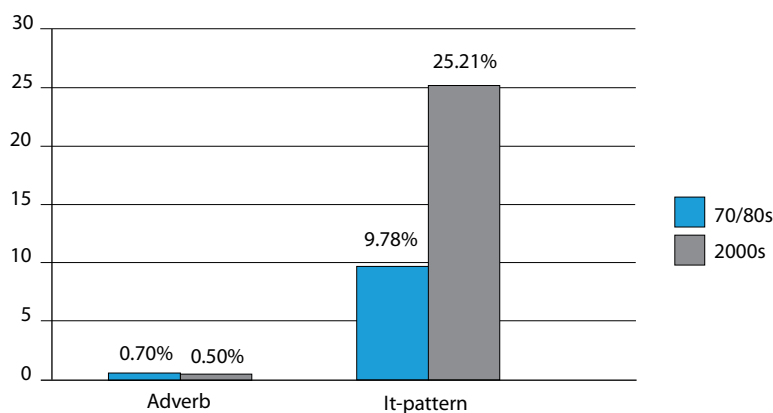


Figure 7. Percentage of modification/relations used with adverb and it-pattern in sentence-initial position in SciTex for the obviousness meaning

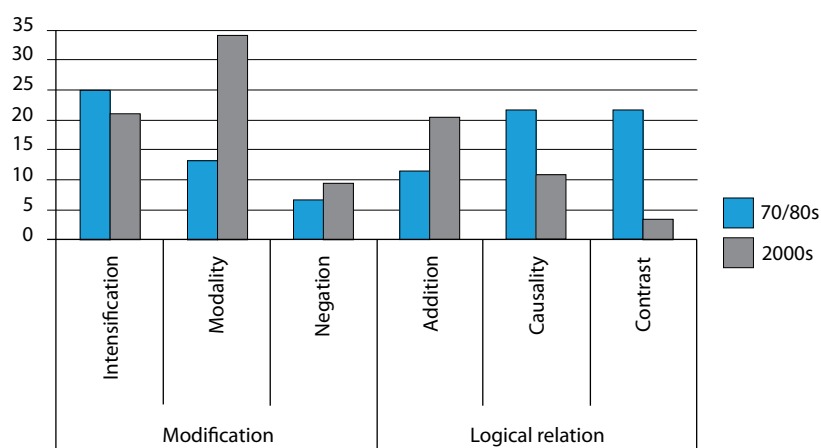


Figure 8. Percentage of modification/relations used with the it-pattern in sentence-initial position in SciTex for the obviousness meaning

43 The modification patterns we have encountered with the it-pattern are intensification (e.g., *It is quite clear that*), modality (e.g., *It should be clear that*), and negation (e.g., *It is not clear that*). The logical relations found are contrast (e.g., *It is, however, clear that*), causality (e.g., *It is, therefore, clear that*), and addition (e.g., *It is also clear that*). Figure 8 shows the percentage of each type for the obviousness meaning in both time periods. While addition, modality and negation increase over time, causality, intensification and especially contrast decrease.

44 In summary, we have seen that for the obviousness meaning the adverb can be multi-functional as it expresses both evaluative and textual meaning, thereby

contributing to the cohesiveness of the discourse. While there are other options to express obviousness in sentence-initial position (such as the *it*-pattern and noun phrase), they seem to be used either when a more explicit variant is needed, as in the case of the noun phrase, or when particular additional meanings (such as modality or negation) have to be expressed. To build up cohesion around these variants (*it*-pattern and noun phrase), respective conjuncts are used in addition (such as *however*, *therefore*, etc.). The adverb, instead, is a very dense way of expressing evaluative and textual meaning at the same time, and thus, might be the preferred option when no explicitness or particular modification or relation type is needed.

#### 4.2.2. *The importance meaning*

45 In the following, we focus on the importance meaning used in sentence-initial position. Here, the aim is to investigate whether, similarly to the evaluative meaning of obviousness, it is also multi-functional.

46 We first consider the importance meaning realized by the adverbial pattern. The adverbs here are i.a. *essentially*, *fundamentally*, *importantly*, *necessarily*, *notably* (the complete list comprises 19 adverbs). We can see from example [24] that this meaning seems to be less strongly involved in contributing to discourse cohesion than the obviousness meaning described above. The relation is much harder to pin down to a specific conjunctive relation.

[24] A contextual-frequency effect on segment deletion that abstracts away from specific word sequences is discussed by Bybee (2002), who shows that the phonological environments a word occurs in frequently affect the word's readiness to undergo reduction. [...] **Importantly**, this finding does not rest on the frequency of combinations of particular words, but rather on an abstraction to classes of phonological contexts. Viewed in this way, the findings in Bybee 2002 lend motivation to an investigation of probabilistic effects at other abstract levels of linguistic information, including probabilities that are not word-to-word or word-to-sound.

(SciTex; C1-linguistics, 2000s)

47 According to Thompson and Zhou (2001: 136), these kinds of adverbs behave more or less in the role traditionally ascribed to them, conveying evaluative meaning. But they also argue that even though in these cases we cannot exactly pin down a cohesive function to these adverbs that matches the cohesive function of a conjunct, as we have done with adverbs of the obviousness meaning, they may still play a role in interpersonal coherence, which has been insufficiently investigated so far and which is therefore quite difficult to label or explicate (cf. Thompson & Zhou, 2001: 137). What we can see from example [24] is that the writer uses the adverb in order to propagate to the reader that the proposition following the adverb is significant, not on its own but in relation to the previous and following discourse. We have observed that importance adverbs mostly introduce a specification of

the preceding discourse, which is presented to the reader as being of importance. The following sentence is then somehow related to the previous more specific and important proposition. In some way a consecutive relation is built up that ranges from a relatively general proposition (first sentence) to a more specific proposition signaled as being important (second sentence introduced by *importantly*) to a further elaboration of this particular proposition (third sentence introduced by *viewed in this way*).

- 48 The relation built up in example [25] is quite similar. Although the use of *however* appears to indicate that a concessive relation is being expressed, a closer look shows that the proposition following the *however* is a further elaboration of what was signaled as being important in the previous sentence. Thus, we could read: “*adjectives decline according to syntactic gender* → *for adjectives this gender is taken up from the noun*”. Again, the consecutive relation is quite subtle and might also be perceived without the adverb. However, it is the adverb that signals the reader to pay attention to the following proposition, helping to establish a logical connection between the previous and the following discourse with this important proposition.

[25] We argue that Russian adjectives fall into two major declension classes, depending on the choice of their theme suffix in the relevant cells of the paradigm (the direct Cases, for the most part). Variation within the classes is therefore mostly phonological, although some readjustment rules (Chomsky and Halle 1968, Halle 1990) also apply. **Importantly**, adjectives decline according to syntactic gender. However, unlike nouns, adjectives have no inherent gender, but are assigned the gender of some noun in the sentence or discourse (agreement or concord).  
(SciTex; C1-linguistics, 2000s)

- 49 Besides linguistics, this cohesive function can also be seen in other disciplines. Example [26] is taken from biology. Again, we can detect a consecutive relation from the more general to the more specific proposition.

[26] Discussion Interactions between hormones and growth factors have been detected in many cells and tissue systems. **Importantly**, these events seem physiologically relevant in the skeleton where selective changes in growth factor expression and activity offer a cogent explanation of hormonal action. For example, stimulatory effects by growth hormone on longitudinal bone growth, and by parathyroid hormone during bone remodeling each depend on increases in new IGF-I synthesis.  
(SciTex; C2-biology, 2000s)

- 50 While the obviousness meaning was involved in different kinds of cohesive relations, the importance meaning realized by an adverb seems to be mainly involved cohesively in a consecutive relation (as seen from the examples above). If, for instance, a concessive relation has to be expressed, a conjunct is used in addition to the adverb (see *however* in example [27]).

[27] The ideas in this article build on his important work in this area. **Importantly**, however, my approach differs in that an individual's knowledge of the elements of his/her language and their usage are given a central role. Thus, in my view, how a listener interprets, or parses, a speech signal is language specific rather than universal, as Ohala assumes.

(SciTex; C1-linguistics, 2000s)

51 In summary, the importance meaning realized by the adverbs is also involved in a cohesive function, but in comparison to the obviousness meaning it seems to be mainly confined to one type of relation – the consecutive one, at least in our corpus.

52 In the following, we consider the it-pattern used with the importance meaning. Possible adjectives used in the pattern are i.a. *central*, *crucial*, *essential*, *important*, *relevant* (the complete list comprises 43 adjectives). Focusing on its possible cohesive function, we can see how it differs in fact from the adverb usage. Consider example [28], where a concessive relation is built up between the it-pattern and the *but* (similarly to the obviousness meaning, which however was realized only by an adverb). Interestingly, approximately 40% of the time the pattern is used with the verb *note* as shown in [28]. Other verbs used are also cognitive verbs such as *observe* (3.3%), *consider* (2.7%), *understand* (2.5%) as well as semiotic verbs such as *point out* (2.7%) and *emphasize* (2.5%). They all function similarly to *note* varying slightly in their semantic notions, i.e. these cognitive and semiotic verbs trigger the engagement of the reader actively when cognitive verbs are used and passively when semiotic verbs are used. Thus, the reader should notice something to be important that he/she is (in a sense passively) pointed towards or should (actively) consider. This allows the writer to use a more explicit way of involving the reader within the discourse (see also example [29], again involved in a concessive relation).

[28] We have touched on just a few of the issues that require deeper investigation for those invested in understanding the processes of language endangerment and maintenance. **It is significant to note** that a rich literature already exists for at least three of them: globalization, literacy, and multilingualism. But few linguists will be exposed to this literature in the course of their graduate study, and few will have the time necessary to gain wide exposure to it.

(SciTex; C1-linguistics, 2000s)

[29] In other words, the dispreference for the sequence CVCCV as opposed to CCVCV (or more accurately: VCC as opposed to VCV) results in the creation of previously unattested consonant sequences at the beginning of the word (CCVCV). **It is important to point out** that this observation is not inconsistent with the claims made here. It does, however, underscore the fact that a listener's sensitivity or bias towards one structure may have consequences that affect the structure of the language more generally.

(SciTex; C1-linguistics, 2000s)



53 However, the it-pattern is also used similarly to the importance adverbs shown above. In example [30], for instance, the it-pattern is used to introduce a more elaborate proposition (introduced by the it-pattern) than that in the previous sentence which is then further developed in the following proposition (introduced by *by this, we mean, roughly speaking*). This builds up a consecutive relation. Again, it is quite hard to say if the it-pattern of importance really builds up the relation, but the fact is that the it-pattern is involved.

[30] In this section, we exploit the layer-structure of a plane graph in order to gain a “nice” separation of the graph. **It is important that** a “yes”-instance  $(G, k)$  (where  $G$  is a plane graph) of the graph problem  $G$  admits a so-called “layerwise separation” of small size. By this, we mean, roughly speaking, a separation of the plane graph  $G$  [...].  
(SciTex; A-computer science, 2000s)

54 When we compare the use of the adverb and the it-pattern to realize the importance meaning, first of all, we have seen that in the 2000s, they are used similarly in terms of frequency (7.54 per 1 million for the adverb and 7.50 for the it-pattern; see Figure 6). This differs from the obviousness meaning where the adverb is more frequently used (38.7% for the adverb and 17.3% for the it-pattern; see Figure 5). Moreover, the adverb and it-pattern of the importance meaning also seem to share the kind of cohesive relations they are involved in. This also differs from the obviousness meaning, which clearly shows a distinctive usage in this respect, i.e. the adverb is used multi-functionally, but not the it-pattern.

55 Moving to the use of modification and logical relations of the adverb and it-pattern for the importance meaning, we can see from Figure 9 that while there is a diachronic decrease, they are both modified or used with other types of relations. Note that for the obviousness meaning, instead, this was mostly encountered with the it-pattern rather than with the adverb (see Figure 7).

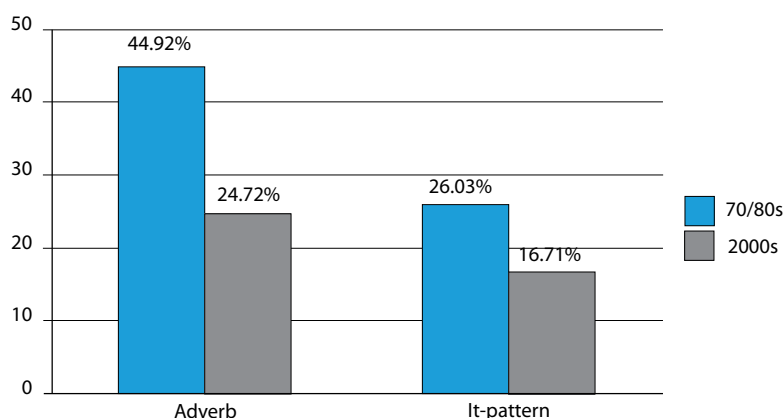


Figure 9. Percentage of modification/relations used with adverb and it-pattern in sentence-initial position in SciTex for the importance meaning

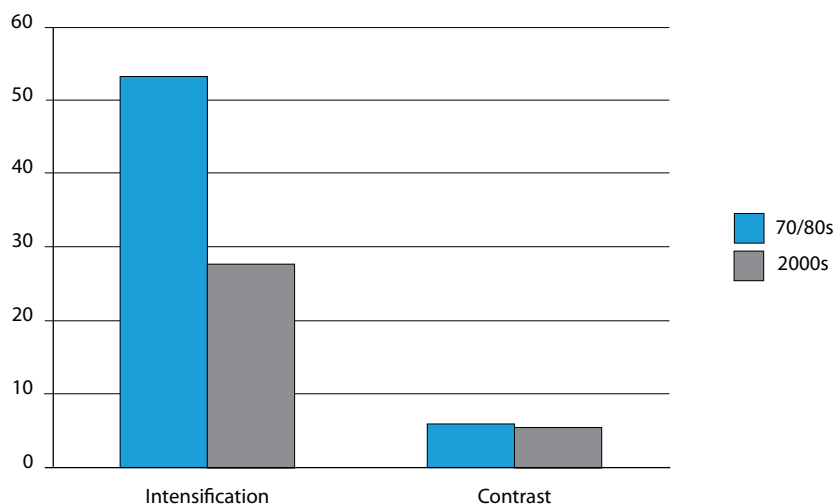


Figure 10. Percentage of modification/relation used with the adverb in sentence-initial position in SciTex for the importance meaning

56 Looking at the types of modifications and logical relations which the adverb appears with, we can see from Figure 10 that it is only modified by intensification (such as *Most importantly*). This type of modification is quite frequent in the 70/80s but less frequent in the 2000s (respectively, 53.4% and 27.8% of occurrences). Additionally, the importance adverbs appear with the logical relation of contrast (such as *Importantly, however*), even though relatively rarely (around 5% in both time periods).

57 Figure 11 shows the different types of modification and logical relations used with the importance meaning for the it-pattern. We can observe that causality (e.g., *It is, therefore, important that*) and addition (e.g., *It is also important to*) are the most frequently used types, having also increased over time, followed closely by intensification (e.g., *It is very important to*) which has remained fairly stable over time. The use of hedges (e.g., *It seems to be important that*), which did not occur with the obviousness meaning, and negation (e.g., *It is not important to*) have decreased over time (especially negation). This is an interesting tendency as hedges diminish the force of the importance and negation even states the opposite (something is unimportant). These usages seem to be avoided over time. In addition, the use of modality with the it-pattern has slightly increased (e.g., *It would be important to*). Finally, contrast is used (e.g., *It is, however, important to*), but again relatively rarely.

58 Comparing the adverb and the it-pattern, the adverb is mostly intensified (even though this has decreased over time), whereas the it-pattern shows more variability in the types of modification and logical relations that it can be combined with. We can conclude that if causality or addition has to be expressed, the it-pattern is chosen (see Figure 11), while intensification is used with both the adverb and it-pattern (both accounting for approximately 28% of occurrences in the 2000s).

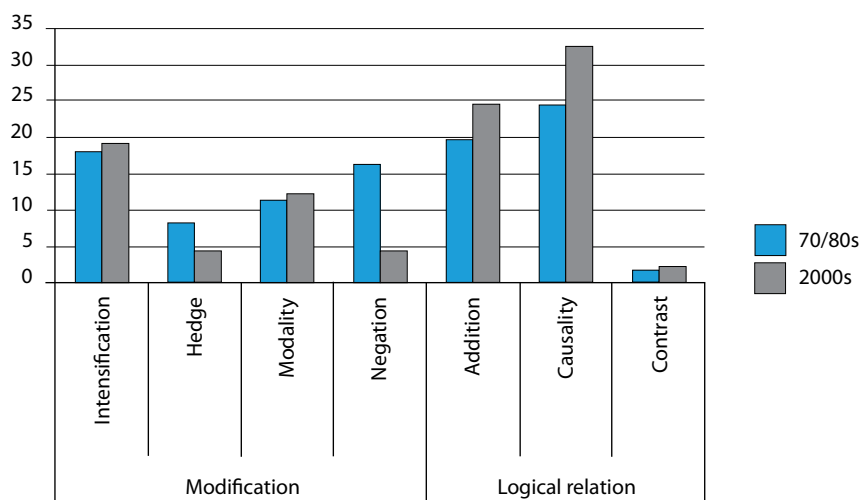


Figure 11. Percentage of modification/relations used with the it-pattern in sentence-initial position in SciTex for the importance meaning

59 Overall, we can conclude that the importance meaning does not show a specific preference for the adverb or the it-pattern, which was the case for the obviousness meaning. While the obviousness meaning realized by adverbs can be used multi-functionally, fulfilling an evaluative as well as cohesive function, the multi-functionality of the importance meaning can be realized in either way (i.e. either with the adverb or the it-pattern). Even though the adverb is a denser form to realize the importance meaning, it is not necessarily the preferred one. What may make the it-pattern attractive is its quite conventionalized usage of importance + cognitive (such as *note*, *consider*) or semiotic (such as *point out*, *emphasize*) process, which helps to enforce the interaction between writer and reader in the discourse, thereby contributing to the cohesiveness of the text.

## 5. Summary and envoi

60 In this paper, we investigated the hypothesis of *evaluative cohesion* (as determined by Thompson and Zhou [2001]) in scientific writing. For this, we looked at how particular types of evaluative meaning expressed in sentence-initial position can also function cohesively, in that they help to structure the discourse by contributing to the cohesiveness of the text. We have analyzed two evaluative meanings, obviousness and importance, which were shown to be used distinctively in sentence-initial position in a previous study (Degaetano-Ortlieb, 2015). In particular, we analyzed these two meanings by considering possible differences in terms of their lexico-grammatical realizations, but also accounting for a possible use of modification or specific types of relations combined with these meanings.

61 Methodologically, we employed corpus-based methods of annotation and  
extraction of evaluative meaning from the SciTex corpus. Moreover, we applied an  
analytical cycle that involves macro- and micro-analytical steps to test our hypothesis  
(Section 3).

62 At the macro-level, we looked at the distribution of the realization types for each  
meaning, considering also diachronic tendencies (Section 4.1). This gave insights into  
possible preferences. The obviousness meaning has been shown to prefer the adverbial  
pattern over the it-pattern and nominal phrase. The importance meaning, instead,  
has a preference for the noun phrase, while the adverbial and the it-patterns are  
used in similar proportions in the 2000s. However, both meanings show an increase  
in the use of the adverbial pattern. We have argued that this might be related to  
information theoretic purposes, i.e. as language users try to communicate in the most  
efficient way, they might opt for denser forms of communication (cf. Levy & Jaeger,  
2006), when other factors (such as modification or explicitation) are not needed.

63 At the micro-level, we have seen that for the obviousness meaning the adverb can  
be multi-functional, mostly involved in concessive relations, while the it-pattern is  
not (Section 4.2.1). For the importance meaning, both the adverb and the it-pattern  
can have a multiple function, mostly involved in consecutive relations (Section 4.2.2).

64 Further investigations could be related to whether the evaluative or the cohesive  
function for these items (adverb or it-pattern) has changed over time, becoming e.g.  
either weaker or stronger. Possible questions might be, for instance, whether the  
importance meaning realized by the it-pattern and used in a quite conventionalized  
way, as we have seen with the verb *note* (*It is important to note that*), has encountered  
a bleaching of its evaluative meaning over time? Or whether the cohesive function  
of these evaluative items has been used constantly over time or emerged at some  
stage? Moreover, so far we have considered the multi-functionality of these evaluative  
items in scientific writing in general (more explicitly in scientific research articles),  
but it would also be interesting to look at whether scientific disciplines differ in  
their use of this multi-functionality.

## References

- ANDERSON, L. & BAMFORD, J. (eds.) 2004. *Evaluation in Oral and Written Academic Discourse*. Roma: Officina Edizioni.
- AYLETT, M. & TURK, A. 2004. 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.
- BIBER, D. 2004. Historical Patterns for the Grammatical Marking of Stance: A Cross-register Comparison. *Journal of Historical Pragmatics* 5 (1): 107-136.
- BIBER, D. & FINEGAN, E. 1989. 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.

- CONRAD, S. & BIBER, D. 2000. Adverbial Marking of Stance in Speech and Writing. In S. HUNSTON & G. THOMPSON (eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press: 56-73.
- CWB 2010. The IMS Open Corpus Workbench. Available online: <http://www.cwb.sourceforge.net>.
- DEGAETANO, S. & TEICH, E. 2011. The Lexico-Grammar of Stance: An Exploratory Analysis of Scientific Texts. In S. DIPPER & H. ZINSMEISTER (eds.), *Beyond Semantics: Corpus-Based Investigations of Pragmatic and Discourse Phenomena. Proceedings of the DGfS Workshop Göttingen, February 23-25, 2011*. Bochumer Linguistische Arbeitsberichte 3: 57-66. Available online: [http://www.linguistics.ruhr-uni-bochum.de/bla/beyondsem2011/bla3\\_online\\_proceedings\\_110301.pdf](http://www.linguistics.ruhr-uni-bochum.de/bla/beyondsem2011/bla3_online_proceedings_110301.pdf).
- DEGAETANO-ORTLIEB, S. 2015. *Evaluative Meaning in Scientific Writing – Macro- and Micro-Analytic Perspectives Using Data Mining*. PhD thesis. Universität des Saarlandes.
- DEGAETANO-ORTLIEB, S. et al. 2013. SciTex – A Diachronic Corpus for Analyzing the Development of Scientific Registers. In P. BENNETT et al. (eds.), *New Methods in Historical Corpus Linguistics. Corpus Linguistics and Interdisciplinary Perspectives on Language (CLIP 3)*. Tübingen: Narr: 93-104.
- DEGAETANO-ORTLIEB, S., KERMES, H. & TEICH, E. 2014. The Notion of Importance in Academic Writing: Detection, Linguistic Properties and Targets. *Journal for Language Technology and Computational Linguistics* 29 (1): 63-77.
- EASON, S. 2001. Stating the Obvious: Modal Adjuncts in NATO Press Briefings. *ASp* 31-33: 39-50. Available online: <http://asp.revues.org/1764>.
- EVERT, S. & HARDIE, A. 2011. Twenty-First Century Corpus Workbench: Updating a Query Architecture for the New Millennium. In *Proceedings of the Corpus Linguistics 2011 Conference (20-22 July 2011, ICC Birmingham)*. 1-21. Available online: <http://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/Paper-153.pdf>.
- FANKHAUSER, P., KERMES, H. & TEICH, E. 2014a. Combining Macro- and Microanalysis for Exploring the Construal of Scientific Disciplinarity. In *Proceedings of the Digital Humanities Conference 2014 (7-12 July 2014, Lausanne, Switzerland)*. Available online: <http://dharchive.org/paper/DH2014/Poster-126.xml>.
- FANKHAUSER, P., KNAPPEN, J. & TEICH, E. 2014b. Exploring and Visualizing Variation in Language Resources. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC-2014)*. Stroudsburg: Association for Computational Linguistics: 4125-4128. Available online: [http://www.lrec-conf.org/proceedings/lrec2014/pdf/185\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2014/pdf/185_Paper.pdf).
- FELLBAUM, C. (ed.) 1998. *WordNet. An Electronic Lexical Database*. Cambridge: MIT Press.
- HALLIDAY, M.A.K. 1988. On the Language of Physical Science. In M. GHADESSY (ed.), *Registers of Written English: Situational Factors and Linguistic Features*. London: Pinter: 162-177.
- HALLIDAY, M.A.K. 2004. *An Introduction to Functional Grammar*. London – New York: Arnold – Oxford University Press.
- HALLIDAY, M.A.K. & HASAN, R. 1985. *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Geelong: Deakin University Press.

- HO-DAC, L.-M. et al. 2012. An Empirical Approach to the Signalling of Enumerative Structures. *Discours* 10: 1-27. Available online: <http://discours.revues.org/8611>.
- HOEY, M. 1983. *On the Surface of Discourse*. London: Allen & Unwin.
- HUNDT, M., SAND, A. & SIEMUND, R. 1999. *Manual of Information to Accompany "The Freiburg-LOB Corpus of British English (FLOB)"*. Albert-Ludwigs-Universität Freiburg. Available online: <http://clu.uni.no/icame/manuals/FLOB/INDEX.HTM>.
- HUNSTON, S. 2004. Counting the Uncountable: Problems of Identifying Evaluation in a Text and in a Corpus. In A. PARTINGTON, J. MORLEY & L. HAARMAN (eds.), *Corpora and Discourse*. Bern – New York: P. Lang: 157-188.
- HUNSTON, S. 2011. *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. New York: Routledge.
- HUNSTON, S. & SINCLAIR, J. 2001. A Local Grammar of Evaluation. In S. HUNSTON & G. THOMPSON (eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press: 74-101.
- HUNSTON, S. & THOMPSON, G. (eds.) 2000. *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.
- HYLAND, K. 2005. Stance and Engagement: A Model of Interaction in Academic Discourse. *Discourse Studies* 7 (2): 173-192.
- HYLAND, K. & TSE, P. 2004. Metadiscourse in Academic Writing: A Reappraisal. *Applied Linguistics* 25 (2): 156-177.
- JOCKERS, M.L. 2013. *Macroanalysis: Digital Methods and Literary History*. Urbana: University of Illinois Press.
- KERMES, H. 2003. *Off-line (and On-line) Text Analysis for Computational Lexicography*. PhD thesis. Universität Stuttgart.
- KERMES, H. & TEICH, E. 2012. Formulaic Expressions in Scientific Texts: Corpus Design, Extraction and Exploration. *Lexicographica* 28 (1): 99-120.
- LEVY, R. 2008. Expectation-Based Syntactic Comprehension. *Cognition* 106 (3): 1126-1177.
- LEVY, R. & JAEGER, T.F. 2006. Speakers Optimize Information Density Through Syntactic Reduction. *Advances in Neural Information Processing Systems* 19. Available online: <http://papers.nips.cc/paper/3129-speakers-optimize-information-density-through-syntactic-reduction.pdf>.
- MARTIN, J.R. & WHITE, P.R.R. 2005. *The Language of Evaluation: Appraisal in English*. Basingstoke – New York: Palgrave Macmillan.
- KUČERA, H. & FRANCIS, W.N. 1967. *Computational Analysis of Present-Day American English*. Providence: Brown University Press.
- RUPPENHOFER, J. et al. 2010. *FrameNet II: Extended Theory and Practice*. Technical report. Berkeley: International Computer Science Institute. 119 p. Available online: <https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf>.
- SANZ, R.L., MUR-DUEÑAS, P. & LAFUENTE-MILLÁN, E. (eds.) 2010. *Constructing Interpersonality: Multiple Perspectives on Written Academic Genres*. Newcastle: Cambridge Scholars.

- SCHMID, H. 1994. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of the International Conference on New Methods in Language Processing (NEMLAP)* (14-16 September 1994, UMIST, Manchester, UK). Available online: <http://www.cis.uni-muenchen.de/~schmid/tools/TreeTagger/data/tree-tagger1.pdf>.
- STOTESBURY, H. 2003. Evaluation in Research Article Abstracts in the Narrative and Hard Sciences. *Journal of English for Academic Purposes* 2 (4): 327-341.
- STOTESBURY, H. 2006. Gaps and False Conclusions: Criticism in Research Article across the Disciplines. In K. HYLAND & M. BONDI PAGANELLI (eds.), *Academic Discourse across Disciplines*. Berlin – New York: P. Lang: 123-142.
- SWALES, J.M. 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge – New York: Cambridge University Press.
- TEICH, E. & FANKHAUSER, P. 2010. Exploring a Corpus of Scientific Texts Using Data Mining. In S.T. GRIES, S. WULFF & M. DAVIES (eds.), *Corpus-Linguistic Applications: Current Studies, New Directions*. Amsterdam – New York: Rodopi: 233-247.
- THOMPSON, G. 2001. Interaction in Academic Writing: Learning to Argue with the Reader. *Applied Linguistics* 22 (1): 58-78.
- THOMPSON, G. & ZHOU, J. 2001. Evaluation and Organization in Text: The Structuring Role of Evaluative Disjuncts. In S. HUNSTON & G. THOMPSON (eds.), *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press: 121-141.
- VIS, K., SANDERS, J. & SPOOREN, W. 2012. Diachronic Changes in Subjectivity and Stance – A Corpus Linguistic Study of Dutch News Texts. *Discourse, Context and Media* 1 (2-3): 95-102.
- WINTER, E.O. 1977. A Clause-Relational Approach to English Texts: A Study of some Predictive Lexical Items in Written Discourse. *Instructional Science* 6 (1): 1-92.
- WINTER, E.O. 1982. *Towards a Contextual Grammar of English: The Clause and Its Place in the Definition of Sentence*. London: Allen & Unwin.
- WINTER, E.O. 1994. Clause Relations as Information Structure: Two Basic Text Structures in English. In M. COULTHARD (ed.), *Advances in Written Text Analysis*. London – New York: Routledge: 46-68.