UFC

CENTRE DE RECHERCHE LUCIEN TESNIÈRE

CIRST
Centre interuniversitaire de recherche sur la science et la technologie

UQÀM

with the support of

Observatoire des sciences et des technologies

Weak Links and Strong Meaning: The Complex Phenomenon of Negational Citations

Marc Bertin & Iana Atanassova

3rd International Workshop on
Bibliometric-enhanced Information Retrieval (BIR2016)
@ ECIR 38th European Conference on Information Retrieval
20-23 March 2016, Padova, Italy

ecir PADOVA 2016
38th European Conference on Information Retrieval

Historical Perspective on Negational Citations

- Thinking about negational citations is not a new topic :

  The notion of Negative credit refers to the fact of criticizing, correcting, disclaiming and disputing other works using negative references. (Garfield)

- It is not a easy topic

  "*The end result is that there are few negational references in the scientific literature. What criticism remains is disguised, redirected away from important people, or referenced perfunctorily.*" (The Negative Reference, MacRoberts)

Problems

- Where to find negative citations? How to define the context around citations?
  - Some studies show that we have to take into consideration anaphor resolution to enhance contexts.
- Identification of the expression of negation
  - The negational reference is often hidden, diffuse or disassembled.
  - There are few negational references in scientific literature.
- External constraints
  - The nature of citations makes direct criticism inappropriate.
  - Some publishers make texts less critical.

Corpus

- Our dataset consists of six peer-reviewed academic journals published in Open Access by the Public Library of Science (PLOS).
- We have processed the entire dataset of about 80,000 research articles published up to September 2013.

Corpus

Journal	Articles	Citations	Citation contexts
PLOS Biology	1,754	170,785	91,117
PLOS Computational Biology	2,560	243,488	126,870
PLOS Genetics	3,414	332,845	185,5 37
PLOS Medicine	926	72,676	34,819
PLOS Negl. Tropical Diseases	1,872	133,022	73,211
PLOS ONE	72,158	5,363,036	2,854,082
Total	*82,684*	*6,315,852*	*3,365,636*

Objectives

- Exploratory work around the location of negational citations.
- Compare negational contexts in sentences with in-text references and in sentences without in-text references
- Study the location of negational citations in the IMRaD Structure (Introduction, Methods, Results and Discussion) of articles

Protocol

- We identified the **IMRaD structure** (Introduction, Methods, Results, and Discussion) by analysing the section titles.
- More than 97% of the research articles in the corpus contain the four section types of the IMRaD structure.
- We **segmented all sections into sentences** and extracted sentences containing in-text references.

Linguistic Processing

- We perform both Part-Of-Speech tagging and lemmatization
- To identify the relevant verbs we analyse and extract all verbs from citation contexts. Then we focus on author disagreement :
  - **not agree with / disagree with**

Some Examples

'However, we and other authors (Table S2) disagree with the low percentage of anxious-type OCD in patients with GTS as observed by Shapiro and Shapiro [30].'

'We do not agree with the methodology of using a training set without MCI to select biomarkers to differentiate between AD and MCI [1].'

'Results disagree with Bequet and Przeworski [3] who reported a split time of 1.4 million years and an ancestral effective size of (C.I.).'

'Note also that I disagree with Livezey and Zusi ([64]: character 1142) regarding the condition in Pelecanus, which I interpreted as possessing the plesiomorphic state (Figure 12).'

Categorization of the Object of Negation

Based on our observations, we have constructed the following 5 categories:

- **model or theory** (e.g. Darwin's view of species)
- **actor of science** (proper names of people)
- **object as a material product of research** (work, table, result, report)
- **object as an abstract product of research** (idea, findings, concept, notion, opinion)
- **framework** (classification, norm)

Ratio of the Negational References

- We have looked into sentences containing the negational expressions with and without in-text references.
- This ratio is an important element to define the context around citations where we will have to look for the negational expressions.

Section	With In-text References	Without In-text References	Ratio
Introduction	12.1%	20.8%	1.1
Methods	2%	0%	0
Results	14.1%	22.4%	1.1
Discussion	71.8%	56.8%	0.5

Table 1: Negational references in the IMRaD structure

Examples: sentences without in-text references

Our observations in humans also contrast with the molecular responses observed during …

These results differ from previously published agglutination data …

Thus, the full model proposed by Nietzsche has remained empirically unproven..

Our results appear to contradict this well established dogma..

However, we and other authors (Table S2) disagree with the low percentage …

Note also that I disagree with Livezey and Zusi

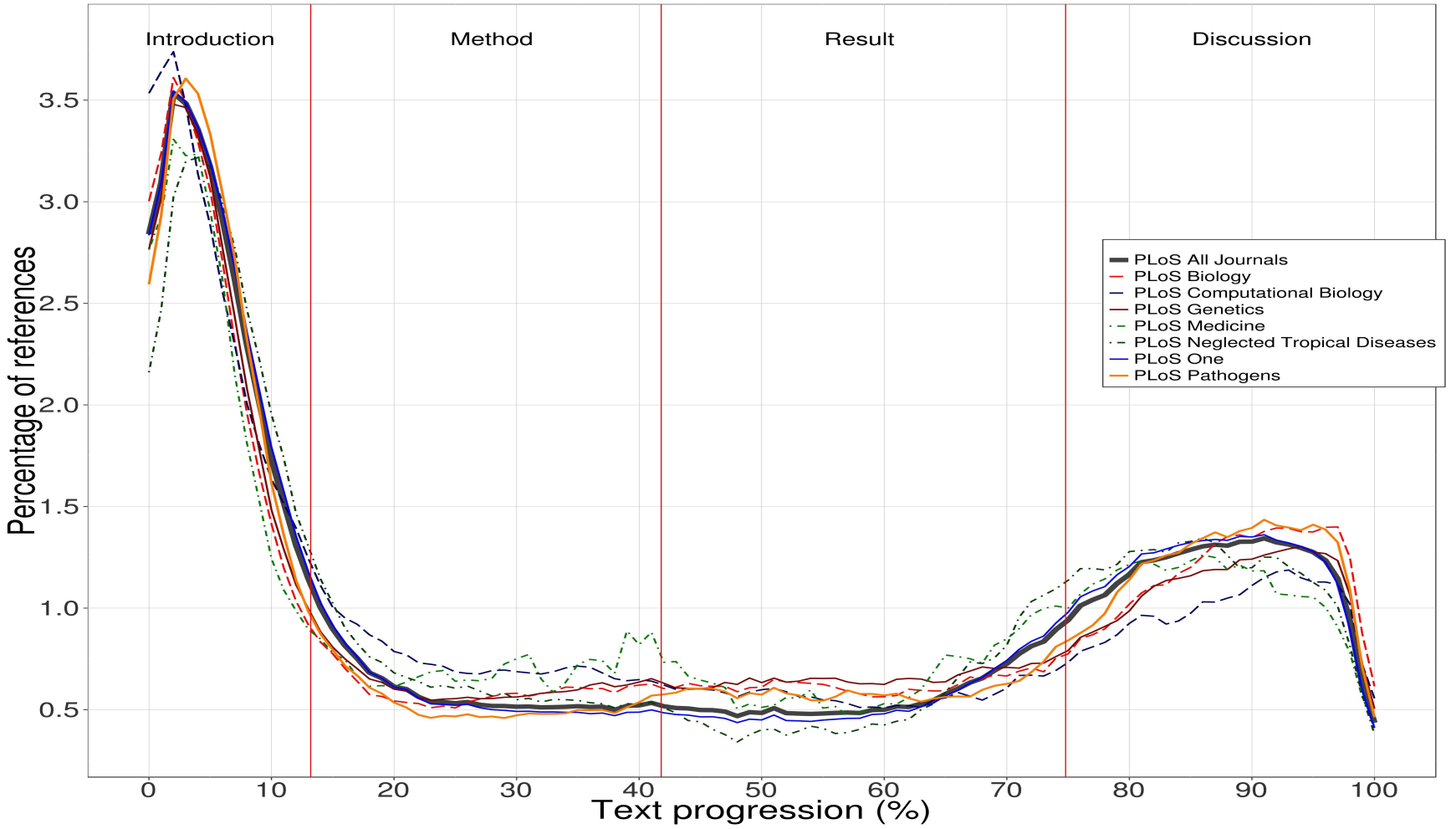
Some authors believe, these animals pose a major threat to man

Most biologists nowadays disagree with Darwin's view of species, largely because of Mayr's biological species concept.

After 1986, some authors persisted in the confusion.

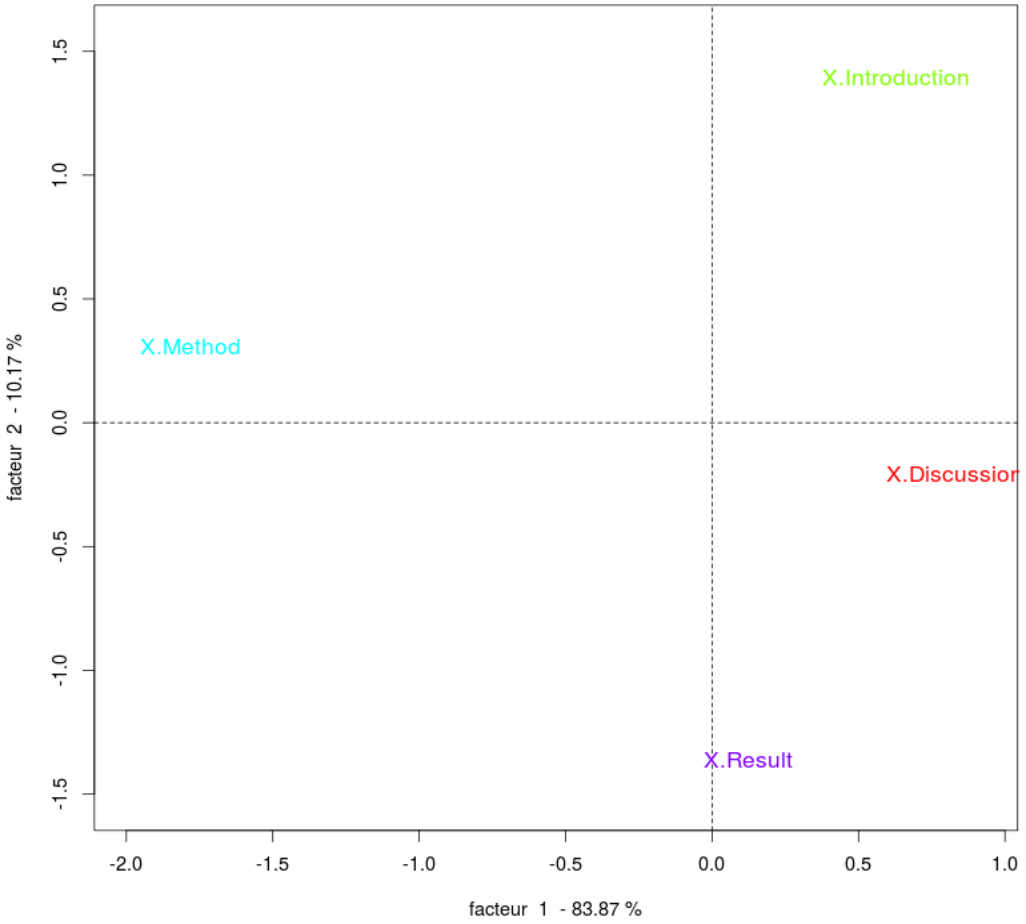
This result is similar to the clustering algorithm's results and we disagree with Chang's classification.

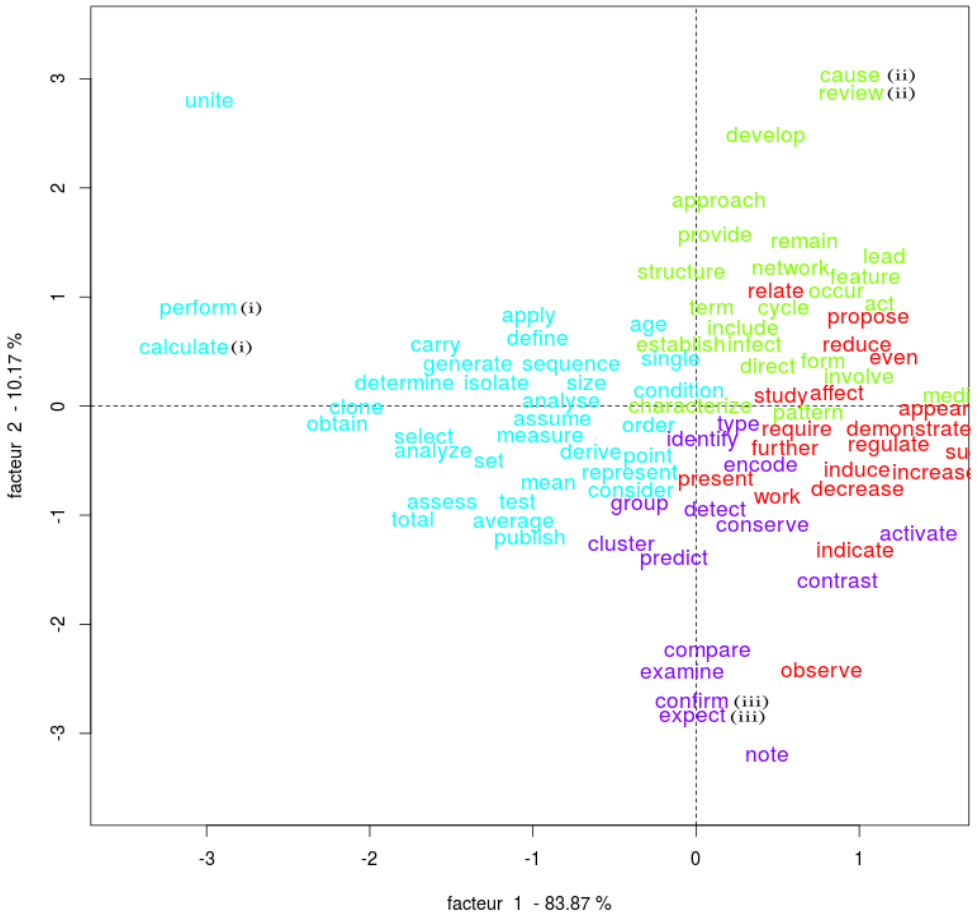
Building Metrics



The Invariant Distribution of References in Scientific Articles (2016)
Marc Bertin, Iana Atanassova, Vincent Larivière and Yves Gingras, JASIST

Factorial Correspondance Analysis of Verbs Around Citations





Metrics

In-text Reference distributions

Part-Of-Speech tagging and lemmatization

Verb distributions and AFC

Semantic level

Negational citations ... and others

IMRaD

Perspectives

- Reproducible research
  - corpus available soon

- Thinking about ethics and not only about applications

Thank you!

Marc Bertin bertin.marc@gmail.com
CIRST, Université du Québec à Montréal, Canada

Iana Atanassova iana.atanassova@univ-fcomte.fr
Centre Tesnière, Université de Franche-Comté, France

Bertin, M.; Atanassova, I.; Larivière, V. & Gingras, Y. *''The Invariant Distribution of References in Scientific Articles''* Journal of the Association for Information Science and Technology (JASIST), 2015

Bertin, M. & Atanassova, I. Factorial Correspondence Analysis Applied to Citation Contexts European Conference on Information Retrieval, Proceedings of the First Workshop on Bibliometric-enhanced Information Retrieval co-located with 37th European Conference on Information Retrieval (ECIR 2015), 2015

Bertin, M. & Atanassova, I. Presutti, V.; Stankovic, M.; Cambria, E.; Cantador, I.; Iorio, A. D.; Noia, T. D.; Lange, C.; Recupero, D. R. & Tordai, A. (Eds.) Semantic Web Evaluation Challenge: SemWebEval 2014 at ESWC 2014 "*Extraction and Characterization of Citations in Scientific Papers*" Communications in Computer and Information Science, Springer, 2014, 120-128