BONNEAU Julien
1979-10-07
French
Contact:
2 rue Pierre Fermaud 34 070 Montpellier, France
(+0033)6 07 11 83 75
bonneau_julien@yahoo.fr
http://www.unice.fr/bcl/auteur47

PHD IN LINGUISTICS

UMR 7320: BCL (BASES, CORPUS, LANGAGE) – CNRS-UNSA

Research Interests

Discourse Analysis, Textual & Lexical Semantics, Corpus Linguistics, Statistical NLP, Text Mining

Education

2012	**PhD** in **Linguistics,** Université Nice Sophia Antipolis (UNSA), Nice
2010	**Training course "Introduction of Logicial R for statistics"**, CNRS
2009	**Training course "Numerical calculation"**, A. Noullez, CNRS
2007	**MSc** (Master) in **Natural Language Processing (NLP)**, Université Toulouse Le Mirail (UTM), Toulouse
2002	**Bs** (Deug) in **Economic Science & Applied Mathematics** (MASS)**,** UTM, Toulouse
1998	**BACCALAUREAT S, Mathematics** (High School Diploma)

Academic Appointments & Teaching Experience

2010-2011	**Temporary assistant professor (ATER)**, UNSA,Nice
	Lectures and tutorials: Linguistic methodology (L1), Syntax (L1), General Linguistic (L1, L3)
2007-2010	**Teaching Assistant** (Moniteur), UNSA, Nice
	Lectures and tutorials: Terminology (L3), Informatics for NLP (L3), Linguistic methodology (L1)
2006-2007	**Teaching Assistant** in Informatics, UTM, Toulouse
2005-2006	**Tutor in Linguistics,** UTM, Toulouse

Academic Services & Organizational skills

2007-2011	**Elected student** to the board of BCL Laboratory, CNRS-UNSA, Nice
2007-2010	**Co-Founder and Speaker** of the **LyRe Project:** organization of popular science lectures for high school students - https://sites.google.com/site/projetlyre/projet-lyre
	Founder and Speaker of the linguistics conference **PhD student's day**, Nice
2010	**Co-Founder & Speaker** of the multidisciplinary cycle **Political decision day**, Nice

Computer skills

Programming Language	Fluent in Perl, Xml, Xslt; Experience with Prolog, C, Matlab, CSS
Textometrics, statisticals and NLP tools	R (GNU S), Hyperbase, Lexico, Alceste, TXM, Trameur, Sato, Spad-t, DTM, Monoconc, Collocate, Gephi, etc.

Languages skills

French: Native	**English:** Fluent	**German:** Basic	**Spanish:** Basic

Main publications

Bonneau J. (2012a) *Mathematics and descriptive modulations for interpretative ways in text corpora. Application to the Mendès France corpus (1922-1982)*, Doctoral dissertation, UNSA.

Our processing is a statistical application of a semantic model – François Rastier's semantic theories – on a corpus of Pierre Mendès France's discourses. Our statistical way determine particularity by generality : words by paragraph, paragraphs by text, texts by corpus. We describe how situation – historical or social practice – influe on textual structure. We study the best distribute high frequency words in the corpus with two inovent tools : the cotextual environnement and the asymmetric cooccurrence. The former produce "smallworlds" : description of lexical semantic networks in texts; the latter describe rythms of lexical variations. In our corpus, these rythms are associated to distinct argumentative values : persuasive vs informative.

Bonneau J. (2012b) « La cooccurrence asymétrique : propriétés quantitatives en disponibilité et énergie » In *Proceedings of 11th International Conference on the Statistical Analysis of Textual Data (JADT)*, Liège.

As for a target word of a text, we can define two relationships: the part given to this target word by every word it is related to in a co-text; the part given by this target word to every word it is related to in the co-text. So, that defines two assymetric co-occurrences. This article, focusing on the effects sizes and the qualitative evaluation into a corpora of various textual genders by the same author, Pierre Mendès France. It is an exploration of the properties of these indices on different part of speech for an automatic text comparisons method.

Bonneau J. & A. Dister (2010) « Logométrie et modélisation des interactions discursives. L'exemple des entretiens semi-directifs » In *Proceedings of 10th JADT*, Roma.

Continuing our previous work (Bonneau J., 2008), this paper focusses on interactions and mutual influences between interviewers and interviewees in semi-structured interviews. On a spoken language corpus, we study the disfluencies in the specific interviews organization according to the dialogic resonance situations.

Bonneau J. (2009) « Arguments logométriques pour la définition d'une description externe de corpus : Application au corpus Pierre Mendès France ». In *Proceedings of 6th International Conference on Corpus Linguistics (JLCL)*, Lorient.

This paper presents a method of text classification for defining partitioning corpus. The textual data processing uses endogenous statistical method and external coherence criteria. We apply our approach to a multiparameter corpus of political speech, in order to build a corpus partitioned into genre and time periods.

Bonneau J. (2008) « Outils d'aide à l'exploitation d'entretiens semi-directifs : Etude de l'interaction entre intervieweur et interviewés sur un corpus ethnoécologique » In *Proceedings of 9th JADT*, Lyon.

This paper presents a combination of textual data analysis software, to measure the interaction between interviewer and interviewees in semi-structured interviews. These interviews were conducted with the aim to make possible mediation between the various actors and the land managers on the Parc national des Cévennes. From formalization (TEI and XML) to data processing, the goal is to get a "mapping" of enunciative and thematic indices, own to describe the interactive interview situations.

Bonneau J. (2007) *Apports linguistiques pour l'analyse de corpus en sciences humaines. Exemples d'applications sur des entretiens semi-directifs*, Master's thesis, UTM.

This study assesses the contribution of corpus linguistics in humanities. More specifically, qualitative (ethnoscientific – produced in collaboration with Museum National d'Histoire Naturelle) and quantitative (Alceste methodology) analysis are conducted on the same corpus of witness semi-structured interviews. Advocating multidisciplinarity, these two methods are the starting point for critics and supplements from processings and tools of corpus linguistics.