

Liste des publications

Google Scholar h-index : 24
2.792 citations ; 664 citations depuis 2018.

1 brevet ; auteur de 3 livres ; 6 logiciels avec leur manuel ; co-éditeur de 29 ouvrages collectifs ; auteur de 9 chapitres de livre, 25 articles et 44 papiers de conférences internationales.

1 brevet

2001 : co-auteur avec Brian O'Donovan (Manager, IBM-LOTUS) du brevet (Patent) #GB820010431: *Method and apparatus for performing decomposition of an input string into the concatenation of non-overlapping non-gapped words or word formation elements from large electronic dictionary.*

3 livres (auteur)

— Silberztein, M. (2016). *The Formalisation of Natural Languages: the NooJ approach.* Wiley Eds. : Hoboken NJ, USA (346 pages).

— Silberztein, M. (2015). *La formalisation des langues : l'approche de NooJ.* ISTE Ed. : Londres. (429 pages).

(Le livre de 2016 est une adaptation du livre de 2015 à la formalisation de la langue anglaise)

— Silberztein, M. (1993). *Dictionnaires électroniques et analyse automatique de textes : le système INTEX,* Masson Ed. : Paris. (240 pages).

6 logiciels et leur manuel (auteur)

— WebNooJ, 2023. Il s'agit d'une version simplifiée de NooJ utilisée en Travaux Dirigés. WebNooJ utilise une API écrite à partir de RA.

Disponible sur : <https://webnooj.univ-fcomte.fr>

— RA, 2022-. Un nouveau moteur pour l'analyse et la génération automatique de textes. Ecrit en Swift ; il est compilable en natif (sans machine virtuelle type Java ou .NET) sur machines LINUX, MacOS et Windows.

Disponible sur : <https://nooj.univ-fcomte.fr>

— ATISHS, 2022. ATISHS est une application d'analyse statistique de corpus qui utilise des ressources linguistiques pour identifier des unités de traitement, puis effectue une douzaine de mesures statistiques sur ces unités (distribution, évolution, collocation, analyse factorielle, score standard, etc.). Ce logiciel a été financé par le projet ATISHS de la Région de Franche-Comté.

Disponible sur : <https://atishs.univ-fcomte.fr>

— NooJ, 2002-. NooJ contient plus 150.000 lignes de code C# et 100.000 lignes de code en Java. Manuel de 210 pages en anglais.

Disponible sur <https://nooj-univ-fcomte.fr>.

INTEX représente 100.000 lignes de code C et C++.

— INTEX, 1993-2000. INTEX contient plus de 100.000 lignes de code en C et en C++. La première version d'INTEX était écrite en Objective C sur ordinateur NextStep. Manuel de 230 pages en anglais.

Disponible sur <http://intex.univ-fcomte.fr>.

Editeur ou coéditeur de 3 numéros de revues scientifiques

Silberztein Max, Ed. *Outils informatiques pour la linguistique*. Langue française n°203 (2019).

Atigui Faten, Kornysheva Elena, Métails Elisabeth, Silberztein Max Eds. *Application of Natural Language Processing to Information Systems*. Special issue of *Data @ Knowledge Engineering Journal*. Elsevier B.V., Amsterdam (2019).

Blandine Courtois, Max Silberztein Eds. *Dictionnaires électroniques du français*. Langue française, n°87, Larousse (1990).

Editeur d'un livre collectif

Silberztein M. Eds. *Linguistic Resources for Natural Language Processing*. Springer Ed. (2024), 200 pages, 10 chapitres.

Coéditeur de 25 volumes de sélection d'actes

Bartulović A., Mijić L., Silberztein M. Eds. *Formalising Natural Languages: Applications to Natural Language Processing and Digital Humanities*, 17th International Conference, NooJ 2022, Rosario, Argentina, May 31–June 2, 2023, Revised Selected Papers (2024).

González C., Rodrigo A., Silberztein M. Eds. *Formalising Natural Languages: Applications to Natural Language Processing and Digital Humanities*, 16th International Conference, NooJ 2022, Rosario, Argentina, June 14–16, 2022, Revised Selected Papers (2023).

Bigey M., Richeton A., Silberztein M., Thomas I., Eds. *Formalising Natural Languages: Applications to Natural Language Processing and Digital Humanities*, 15th International Conference, NooJ 2021, Besançon, France, June 9–11, 2021, Revised Selected Papers (2022).

Bekavac, B., Kocijan, K., Silberztein, M., Šojat, K. Eds. *Formalising Natural Languages: Applications to Natural Language Processing and Digital Humanities*, 14th International Conference, NooJ 2020, Zagreb, Croatia, June 5–7, 2020, Revised Selected Papers (2020).

Héla Fehri, Slim Mesfar, Max Silberztein Eds. *Formalising Natural Languages with NooJ* (20 selected papers from the 13th International NooJ Conference, Hammamet, Tunisia, 2019), CCIS Series. Springer Verlag Eds: Heidelberg (2020).

Ignazio Mauro-Mirto, Mario Monteleone, Max Silberztein Eds. *Formalising Natural Languages with NooJ* (21 selected papers from the 12th International NooJ Conference, Palermo, Italy, 2018), CCIS Series. Springer Verlag Eds: Heidelberg (2019 to appear).

Max Silberztein, Faten Atigui, Elena Kornysheva, Elisabeth Métails, Farid Meziane Eds. *Natural Language Processing and Information Systems*. Proceedings of the 23rd

- International Conference on Applications of Natural Languages to Information Systems, NLDB 2018, Paris, France, June 13-15. Springer Verlag, LNCS Series: Heidelberg (2018).
- Anabela Barreiro, Kristina Kocijan, Peter Machonis, Max Silberztein Eds. Proceedings of the LR4NLP Workshop at COLING2018 (2018). Available at: <http://www.aclweb.org/anthology/W18-38> (2018).
- Samir Mbarki, Mohamed Mouroud and Max Silberztein Eds. *Formalizing Natural Languages with NooJ and Its Natural Language Processing Applications* (20 selected papers from the 11th International NooJ Conference, Kenitra, Morocco, 2017), CCIS Series. Springer Verlag Eds: Heidelberg (2018).
- Kristina Kocijan, Peter Machonis, Max Silberztein Eds. Proceedings of the Linguistic Resources for Automatic Natural Language Generation LiRA@NLG workshop at the International Natural Language Generation Conference: <http://aclweb.org/anthology/W/W17/#3800> (2017).
- Linda Barone, Mario Monteleone and Max Silberztein Eds. *Formalising Natural Languages with NooJ* (21 selected papers from the 10th International NooJ Conference, Ceske Budejovice, Czech Republic 2016), CCIS Series. Springer Verlag Eds: Heidelberg (2017).
- Yury Hetsevich, Tatsiana Okryt and Max Silberztein Eds. *Formalising Natural Languages with NooJ* (20 papers selected from the 9th International NooJ conference, Minsk, Belarus 2015), CCIS Series. Springer Verlag Eds: Heidelberg (2016).
- Mario Monteleone, Maria Pia Di Buono and Max Silberztein Eds. *Formalising Natural Languages with NooJ* (22 papers selected from the 8th International NooJ conference, Sassari, Italy, 2014). Cambridge Scholars Publishing: Cambridge (2015).
- Svetla Koeva, Slim Mesfar and Max Silberztein Eds. *Formalising Natural Languages with NooJ* (18 papers selected from the 7th International NooJ conference, Saarbrücken, Germany, 2013). Cambridge Scholars Publishing: Cambridge (2014).
- Anaïd Donabédian and Max Silberztein Eds. *Formalising Natural Languages with NooJ* (31 papers selected from the 6th International NooJ conference, Paris, 2012). Cambridge Scholars Publishing: Cambridge (2013).
- Kristina Vuckovic and Max Silberztein Eds. *Selected papers from the NooJ 2011 International Conference and Workshop* (23 papers selected from the 5th International NooJ conference, Dubrovnik, Croatia 2011). Cambridge Scholars Publishing: Cambridge (2012).
- Zoe Gavriilidou and Max Silberztein Eds. *Selected papers from the NooJ 2010 International Conference and Workshop* (21 papers selected from the 4th International NooJ conference, Komotini, Greece 2010). Univ. of Thrace Ed, Grèce (2011).
- Mohamed Ben Hamadou, Slim Mesfar and Max Silberztein Eds. *NooJ 2009 International Conference and Workshop* (18 papers selected from the 3rd International NooJ conference, Tozeur, Tunisia, 2009). In Finite-State Language Engineering. Centre de publication universitaire. Tunisie (2010).
- Tamas Varadi, Judit Kuti and Max Silberztein Eds. *Applications of Finite-State Language Processing* (20 papers selected from the 2nd International NooJ conference, Budapest, Hungary 2008). Newcastle: Cambridge Scholars Publishing (2010).
- Xavier Blanco and Max Silberztein Eds. *Proceedings of the 2007 International NooJ Conference* (18 papers selected from the 1st International NooJ conference, Barcelona, Spain 2007). Newcastle: Cambridge Scholars Publishing (2009).

- Dusko Vitas, Cvetana Krstev and Max Silberztein Eds. *Proceedings of the 2006 International INTEX Conference* (12 papers selected from the 9th International INTEX conference, Belgrade, Serbia 2006). Univ. of Belgrade Eds.
- Svetla Koeva, Denis Maurel and Max Silberztein Eds. *INTEX/NooJ pour la Linguistique et le Traitement Automatique des Langues* (24 papers selected from the 6th, 7th and 8th International INTEX conferences in Sofia 2003, Tours 2004, Besançon 2005). Cahiers de la MSH Ledoux. Presses Universitaires de Franche-Comté (2008).
- Eric Laporte, Christian Leclère, Mireille Piot et Max Silberztein Eds. *Hommage à la mémoire de Maurice Gross*, *Linguisticae Investigationes* (659 pages) (2004).
- Jean Royauté, Claude Muller et Max Silberztein Eds. *INTEX pour la Linguistique et le Traitement Automatique des Langues* (21 papers selected from the 4th and 5th INTEX conferences in Bordeaux 2001, Marseille 2002). Cahiers de la MSH Ledoux. Presses Universitaires de Franche-Comté (2004).
- Maurice Gross et Max Silberztein eEds. *Les premières journées INTEX* (17 papers, Paris 1996). LADL, Université Paris 7. Ed.: Paris (1996).

9 chapitres de livre

- The limitations of Corpus-Based Methodes in NLP*. In *Linguistic Resources for Natural Language Processing*, Springer Eds. (2024).
- Using Linguistic Criteria to Define Multiword Units*. *Linguistic Resources for Natural Language Processing*, Springer Eds. (2024).
- (avec Vincent Bénet). *Developing resources for Russian, Applications to Linguistics and Language Teaching*. Open Access book available at: <https://www.elibrary.ru/item.asp?id=47945360> (2022).
- (avec Danielle Leeman). *Pourquoi peut-on être au salon et non être à la chambre ? Le curieux comportement des pièces du logis. Diachronie et synchronie. Mélanges en l'honneur de Hava Bat-Zeev Shyldkrot*. Amit, Bertin, Soutet Eds. L'Harmattan, collection Sémantiques (2022).
- Reconnaissance de la négation pour l'analyse d'opinion. *Comunicazione, Linguaggi, Società, Miscellany of essays in honour of Annibale Elia*. Penguin Random House, Grupo Editorial: Bogota (2021).
- Analyse et génération transformationnelle avec NooJ. In *Livelli di analisi e fenomeni di interfaccia*. Bulzoni: Rome (2016).
- Various computational devices for various Linguistic phenomena with NooJ. In *Language resources and technologies for Bulgarian*. S. Koeva, Academic Publishing House: Sofia (2013).
- Une description formalisée des déterminants français. In *Hommage à la mémoire de Maurice Gross, Linguisticae Investigationes*, E. Laporte, C. Leclère, M. Piot, M. Silberztein, eds. (2004).
- The Lexical Analysis of Natural Languages, in *Finite State Language Processing*, E. Roche and Y. Schabes, eds. The MIT Press: Cambridge, MA (1997).

26 articles parus dans des revues avec comité de lecture

- (avec Danielle Leeman) *Pourquoi peut-on être au salon et non être à la chambre ? Le curieux comportement des pièces du logis...* In *Synchronie et diachronie : l'enjeu du sens*. Bibliothèque de grammaire et de linguistique n°67 Champion : Paris (à paraître 2022).

Linguistique et Traitement Automatique des Langues : une coopération est nécessaire. *Langue(s) & Parole* n°5 – Revista de filologia francesa y romanica (2020).

Les dictionnaires DEM et LVF au service des humanités numériques. In *L'héritage de Jean Dubois et Françoise Dubois-Charlier*. Numéro spécial de *LINX* (2020).

(avec Danielle Leeman) Le lexique comme reflet des représentations sociales. In *L'héritage de Jean Dubois et Françoise Dubois-Charlier*. Numéro spécial de *LINX* (2020).

Les dictionnaires électroniques de NooJ. In *Les Cahiers du dictionnaire* no. 6, *Classiques Garnier*, pp. 167-188 (2014).

Corpus Linguistics and Semantic Desambiguation, in G. Maiello, R. Pellegrino. In *Database, Corpora, Insegnamenti Linguistici, Linguistica*, 63, pp. 397-410 (2012).

La formalisation du dictionnaire LVF avec NooJ et ses applications pour l'analyse automatique de corpus. In *Langages* (Larousse) n° 179-180, pp. 221-241 (décembre 2010)

Orthographe compliquée ou orthographe fidèle ? In *Le français d'aujourd'hui* (Armand Colin) n° 170, pp. 85-102 (2010).

Analyse automatique des locutions en *sous* et *sur*. In *Le français moderne*, pp. 154-170, (juillet 2010).

Les unités linguistiques et leur annotation automatique. In *Modèles linguistiques* n° 55, (2007).

(avec Agnès Tutin). NooJ, un outil TAL pour l'enseignement des langues. Application pour l'étude de la morphologie lexicale en FLE. In *Apprentissage des langues et Systèmes d'Information et de Communication*, 8, pp. 123-134 (2005).

(avec Pierre-André Buvet and Fabienne Moreau) Procédures de désambiguïsation sémantique automatique pour la détermination du domaine d'un texte. In *Revue Québécoise de Linguistique*, 32/11, pp. 177-196 (2005).

(avec Gisèle Chevalier and Sylvia Kasparian). Éléments de solution pour le traitement automatique d'un français oral régional. In *TAL*, 45/2, pp. 41-62 (2004).

Finite-State Recognition of the French determiner system. In *Journal of French Language Studies* (Cambridge University Press) 13/2, , pp. 221-246 (2003).

INTEX: a Finite State Transducer toolbox, in *Theoretical Computer Science* 231/1, pp. 33-46 (1999).

Text Indexation with INTEX, in *Computer and the Humanities* 33/3, pp. 265-280 (1999).

Les graphes INTEX. In *Analyse lexicale et syntaxique : le système INTEX, # Linguisticae Investigationes*, XXII, pp. 3-29 (1998-1999)

Les expressions figées dans INTEX. In *ibid*, pp. 425-449

(with Nathalie Friburger) Le débogueur de grammaires sous INTEX. In *ibid*, pp. 413-423.

Transducteurs pour le traitement automatique des textes, in *Travaux de Linguistique*, 37 (1998).

Levée d'ambiguïtés avec INTEX. In *L'ambiguïté*. BULAG, 21, pp. 259-271 (1996).

Analyse automatique de corpus avec INTEX. In *Hommage à Jean Dubois, LINX*, 34-35: (1996).

Groupes nominaux libres et noms composés lexicalisés, in *Linguisticae Investigationes* XVII/2, pp. 405-425 (1994).

Finite state descriptions of various levels of linguistic phenomena. In *Language Research*, 28/4 (1992).

A new approach to tagging: the use of a large-coverage electronic dictionary, in *Applied Computer Translation*, 1/4 (1991).

Le système DELA, in *Linguisticae Communicatio* (1989).

44 papiers présentés à des colloques, puis sélectionnés pour publication par le comité scientifique

- Linguistic Resources for Corpus Processing: the ATISHS project (invited speaker). In Proceedings of the JADT 2022 Conference, Naples, (2022) <https://hal.archives-ouvertes.fr/hal-03854939/document>.
- Unary transformations for French transitive sentences. In *Formalising Natural Languages with NooJ* (12th International NooJ Conference, Palermo, Italy, 2018), CCIS Series. Springer Verlag: Heidelberg (2019).
- Using linguistic resources to evaluate the quality of annotated corpora. In *Proceedings of the LR4NLP Workshop at COLING2018*. Available at: <http://www.aclweb.org/anthology/W18-38> (2018)
- A new linguistic engine for NooJ: parsing context-sensitive grammars with finite-state machines. In *Formalizing Natural Languages with NooJ and Its Natural Language Processing Applications* (11th International NooJ Conference, Kenitra, Morocco, 2017), CCIS Series. Springer Verlag: Heidelberg (2018).
- From FOAF to English: Linguistic Contribution to Web Semantics. In *Proceedings of the Linguistic Resources for Automatic Natural Language Generation* (LiRA@NLG workshop at INLG2017). Available at: <http://aclweb.org/anthology/W/W17/#3800> (2017).
- Joe loves Lea: Transformational Analysis of Transitive Sentences. in *Formalising Natural Languages with NooJ* (9th International NooJ conference, Minsk, Belarus 2015), CCIS Series. Springer Verlag: Heidelberg (2016).
- The electronic dictionary DEM in NooJ. In *Formalising Natural Languages with NooJ* (8th International NooJ conference, Sassari, Italy, 2014). Cambridge Scholars Publishing: Newcastle (2015).
- (With Denis Le Pesant, Paul Sabatier and Marie-Hélène Stéfanini). Présentation d'un thésaurus des mots d'affects : théorie, méthodes et applications. In *Emotions in discourse*. Blumenthal, Peter/Novakova, Iva/Siepmann, Dirk ,eds. Peter Lang : Francfort s. M. (2014).
- Le dictionnaire DEM dans NooJ. In *Proceedings of TALN2014*, Marseille (2014).
- NooJ v4. In Natural Languages with NooJ. In *Formalising Natural Languages with NooJ* (7th International NooJ conference, Saarbrücken, Germany, 2013). Cambridge Scholars Publishing: Newcastle (2014).
- Analyse et génération transformationnelle avec NooJ. In *Proceedings of the 48th annual meeting of the Italian Linguistic Society* (invited speaker) (2014).
- NooJ Computational Devices. In *Formalising Natural Languages with NooJ* (6th International NooJ conference, Paris, 2012). Cambridge Scholars Publishing: Newcastle (2013).
- Variable Unification with NooJ v3. In *Selected papers from the NooJ 2011 International Conference and Workshop* (5th International NooJ conference, Dubrovnik, Croatia 2011). Cambridge Scholars Publishing: Newcastle (2012).
- Open Source Multiplatform NooJ for NLP: CESAR Project. In *Proceedings of COLING 2012. Demonstration Papers* (2012).
- NooJ. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation*. LREC Workshops and Tutorials, Istanbul. Available at www.lrec-conf.org/lrec2012 (2012).

(with Bea Ehmann). Language Technology for Multilingual Automated Content Analysis in Group Research. In *ICLASP13 Conference Book*. 13th International Conference on Language and Social Psychology (2012).

Corpus Linguistics and Semantic Disambiguation. In *Proceedings of the International Conference of Database and Corpora*, Insegnamenti Linguistici (Invited speaker) (2011).

Automatic Transformational Analysis and Generation. In *Selected papers from the NooJ 2010 International Conference and Workshop* (4th International NooJ conference, 2010). Univ. of Thrace: Komotini (2011).

Syntactic Parsing with NooJ. In *Finite-State Language Engineering* (NooJ 2009 International Conference and Workshop (3rd International NooJ conference). Centre de Publication Universitaire, Tunisia (2010).

Disambiguation Tools for NooJ. In *Applications of Finite-State Language Processing*. Cambridge Scholars Publishing: Newcastle (2010).

Complex Annotations with NooJ. In *Proceedings of the 2007 International NooJ Conference*. Cambridge Scholars Publishing: Newcastle (2009).

(with Slim Mesfar) Transducer minimization and information compression for NooJ dictionaries. In *Proceedings of the FSMNLP 2008 conference*. Frontiers in Artificial Intelligence and Applications. IOS Press: Amsterdam (2008).

An Alternative Approach to Tagging. In *Proceedings of NLDB 2007* (Invited Speaker). LNCS series, Springer-Verlag, pp. 1-11 (2008)

NooJ's Linguistic Annotation Engine. In *INTEX/NooJ pour la Linguistique et le Traitement Automatique des Langues* (6th, 7th and 8th International INTEX conferences in Sofia 2003, Tours 2004, Besançon 2005). *Cahiers de la MSH Ledoux*, Presses Universitaires de Franche-Comté (2008).

(With Huei-Chi Lin). Ressources lexicales chinoises pour le TALN. In *Proceedings of TALN2007* (2007).

NooJ: A Linguistic Annotation System for Corpus Processing. In *Proceedings of the HLT/EMNLP2005 Conference* (2005).

NooJ's Dictionaries. In *Proceedings of the 2nd Language and Technology Conference*, (2005).

Named Entity Recognition with NooJ. In *Actes du colloque Terminologie et Intelligence Artificielle TIA2005*. LORIA/INSA de Rouen (2005).

(With Agnès Tutin) NooJ: un outil TAL de corpus pour l'Enseignement des Langues et de la linguistique. In *Actes de la journée ATALA : TAL et apprentissage des langues* (2004).

NooJ: an Object-Oriented Approach. In *INTEX pour la Linguistique et le Traitement Automatique des Langues* (Papers selected from the 4th and 5th INTEX conferences in Bordeaux 2001, Marseille 2002). *Cahiers de la MSH Ledoux*. Presses Universitaires de Franche-Comté (2004).

(With Pierre-André Buvet) Analyse sémantique de textes journalistiques. In *Proceedings of TALN2002*, Corpus et Web. Maison des Sciences de l'Homme de Saint-Denis (2002).

INTEX at IBM. in *Actes des Troisièmes Journées INTEX, Revue informatique et Statistique dans les sciences humaines*: Liège (2000).

INTEX: a walkthrough. In *Proceedings of the Third International Workshop on Implementing Automata* (WIA'98). Springer Verlag: Berlin-New York: (1999).

INTEX: a Finite State Transducer toolbox, in *Proceedings of the Second International Workshop on Implementing Automata* (WIA'97). Springer Verlag: Berlin-New York (1998).

- Normalisation des textes, in *Actes des 4èmes Journées internationales d'Analyse statistique des Données Textuelles* (JADT'98). Nice: Université de Nice - Sophia Antipolis (1998).
- INTEX 4.1 for Windows, in *Proceedings of COLING'98*, Tutorial Notes. Montréal: COLING, (1998).
- Expérience d'étiquetage avec levée d'ambiguïtés, in *Actes du colloque INTEX*. Special Issue of *Linguisticae Investigationes*. Amsterdam (1996).
- Dictionnaires électroniques et comptage des mots, in *Actes des 3èmes Journées internationales d'Analyse statistique des Données Textuelles* (JADT'95). Rome: CNR, (1995).
- Describing Natural Languages with INTEX, in *Proceedings of the conference on Object Oriented Computing in Natural Sciences* (OOCNS'95). Grenoble: IMAG (1995).
- INTEX: a corpus processing system, in *Proceedings of COLING'94*. Kyoto (1994).
- Reconnaissance automatique des mots d'un texte, in *Actes du Colloque Lexiques-Grammaires Comparés: la variation en français*. Jacques Labelle, Christian Leclère, eds. Université du Québec A Montréal, 1992. *Linguisticae Investigationes Supplementa* 17, (1994).
- (With Blandine Courtois) Les dictionnaires DELAS et DELAC, in *Actes du Colloque International RELAI : La description des langues naturelles en vue d'applications linguistiques*. Québec : Université Laval (1989).
- Analyse lexicale, in *Actes de la quatrième rencontre mondiale des départements d'Etudes françaises AUPELF*. New Delhi: Jawaharlal Nehru University (1988).
- The Lexical Analysis of French, in *Proceedings of the Conference on Electronic Dictionaries and Automata in Computational linguistics*. Lecture notes in Computer Sciences, Berlin-New York: Springer-Verlag, (1987).