LIS-TALEP
163, avenue de Luminy - Case 901
13288 MARSEILLE CEDEX 9
(+33) 4 86 09 06 72
Carlos.Ramisch [at] lis-lab.fr
https://pageperso.lis-lab.fr/carlos.ramisch
Birth date: Oct 01, 1986

Carlos Ramisch

Updated on November 22, 2025

Assistant professor (maître de conférences)
Computer science and computational linguistics

Education

2023 **Habilitation à diriger des recherches (HDR)**
Aix Marseille University
Multiword expressions in computational linguistics
Adviser: Alexis Nasr

2009–2012 **PhD in computer science**
University of Grenoble (France) & Federal University of Rio Grande do Sul (Brazil)
A generic and open framework for multiword expressions treatment: from acquisition to applications
Advisers: Aline Villavicencio & Christian Boitet

2007–2009 **Research master in informatics (M2R)**
ENSIMAG - Grenoble Institute of Technology (France)
Multi-word terminology extraction for domain-specific documents
Advisers: Christian Boitet & Aline Villavicencio

2004–2009 **Bachelor in computer science**
Federal University of Rio Grande do Sul (Brazil) "Láurea acadêmica" award

Professional experiences

2013–now **Assistant professor (maître de conférences)**, *LIS-TALEP & Aix-Marseille University*
Research: Computational linguistics, multiword expressions, machine learning, syntax and semantics
Teaching: probabilities, automata theory, compilers, natural language processing, machine learning

2019–2020 **Visiting researcher (délégation CNRS)**, *IRIT-MELODI (Toulouse)*
Research: computational linguistics, multiword expressions, lexical resources, heterogeneous data

2012–2013 **Researcher and teaching and assistant (ATER)**, *LIG-GETALP & Joseph Fourier University*
Research: computational linguistics, machine translation, lexical resources
Teaching: natural language processing, human-computer interaction

2009–2012 **PhD researcher and teaching assistant**, *LIG-GETALP & University of Grenoble*, Grenoble
Research: computational linguistics, development/evaluation of framework for multiword expressions
Teaching: algorithms in C

2008–2009 **Junior researcher (intern)**, *Xerox Research Centre Europe, CLT team*, Grenoble
Improvements and optimizations for a dynamic translation memory

2004–2007 **Junior researcher (intern)**, *Federal University of Rio Grande do Sul*, Porto Alegre
Autonomous language acquisition: multiword expressions and corpora
Artificial intelligence applied to the evaluation of petrol reservoirs

Interests

Scientific: Computational linguistics, machine learning, multiword expressions, semantics, language models
Technical: Python (torch, transformers), C, Java, Linux, PHP & web, bash, SQL.
Leisure: Hiking, cycling, work-out, dancing, sociology, learning languages

Lectures and courses (computer science)

2024–now	Data science & ML (1st, 4th year, AMU)	2013–17	Linux and LaTeX (1st year, AMU)
2021–now	Methodology (MsC–PhD, AMU)	2013	Algorithms in C (1st year, AMU)
2018–now	Probability/statistics (2nd year, AMU)	2013	Comput. linguistics (4th year, UJF)
2014–now	Comput. linguistics (3rd–5th year, AMU)	2013	Interactive design in C (1st year, ENSIMAG)
2018–20	Automata theory (2nd year, AMU)	2012	Human-computer interaction (3rd year, UJF)
2014–19	Compilers (3rd year, AMU)	2009–12	Algorithms in C (1st year, UJF)

Funded projects

2022-2023	Co-leader of working group 1, UniDive (COST) — https://unidive.lisn.upsaclay.fr/
2022-2025	Co-leader, SELEXINI (ANR) — https://selexini.lis-lab.fr/
2020-2021	Local coordinator, France-Ireland Ulysses grant with ADAPT Centre, DCU (PHC)
2016-2019	Local coordinator, PARSEME-FR (ANR) — http://parsemefr.lis-lab.fr
2013-2017	Participant, PARSEME (EU COST Action) — http://typo.uni-konstanz.de/parseme/
2014-2016	Local coordinator, AIM-WEST (CNRS-FAPERGS-FAPESP) — http://aim-west.imag.fr/
2011-2014	Co-leader, CAMELEON (CAPES-COFECUB 707-11) — http://cameleon.imag.fr

Scientific community development

- Member of the gender-equality commission at LIS — https://parite.lis-lab.fr/
- Co-organizer of the outreach event *Les Cigales* — https://www.fr-cirm-math.fr/lescigales/
- 2020-24: co-leader of TALEP research team at LIS — https://www.lis-lab.fr/talep/
- 2020-22: Section representative of the MWE Section of SIGLEX — https://multiword.org/
- Co-organizer of PARSEME shared tasks on MWE identification 2026, 2020, 2018, 2017
- Co-organizer of MWE workshops 2022, 2021, 2020, 2018, 2017, 2013, 2011, 2010
- Editorial board member of Phraseology and Multiword Expressions series, Lang. Science Press
- Tutorials/courses on MWEs at ESSLLI 2018/2024, PARSEME meetings, USP, UFRGS, etc.
- Area/PC chair of *SEM 2012, NAACL 2019, ACL 2020, ACL 2022, PROPOR 2018
- Local organisation: TALN 2012 (Grenoble), 2014 (Marseille), 2022 (Avignon), 2025 (Marseille)
- Reviewer: journals (CoLi, ARR, LRE,...), conferences (*ACL, COLING, EMNLP, CoNLL,...)
- Outreach to general public: Nuit des chercheurs, Petits débrouillards, ...

Adviser and co-adviser

2022–now	Anna Mosolova (PhD)	2018–2019	Anne-Laure Méalier (post-doc)
2022–now	Ioana Ivan (master+PhD)	2014–2017	Silvio Cordeiro (PhD)
2021–2024	Leonardo Pinto Arata (PhD)	2021	Thibault Roux (master)
2022–2023	Tithir Kumar Saha (engineer)	2021	Adrien Pupier (master)
2022	Nathan Chometton (master)	2018–2019	Cindy Aloui (master)
2022	Quentin Barrouyer (master)	2018–2019	Nicolas Zampieri (master)
2022	Baptiste Souche (master)	2016–2018	Natalie Vargas (master)
2018–2022	Léo Bouscarrat (PhD)	2013	Alexander Kobzar (master)
2016–2021	Manon Scholivet (master+PhD)	2012–2014	Sandra Castellanos (master)
2016–2020	Caroline Pasquer (PhD+postdoc)		

Undergrad interns: Vitor de Araujo (2011–2012), Alexandre Rondon (2014), Tristan Mollet (2017), Solange Lâm (2018), Lucas Pagès (2019), Charles d'Audigier (2019), Yannis Coutouly (2021), Marius Thorre (2022).

Awards, fellowships and invitations

2017 Invited talk at MUMTTT conference
2016 Erasmus grant for visiting the LexiCon group at the University of Granada (Spain)
2014 Merit Scholarship from FRQNT for scientific visit to UQAM (Canada)
2013 Qualification French CNU 27 (informatics) and 7 (linguistics)
2009–2012 PhD fellowship (*allocation*) of the French Ministry of Higher Education and Research
2007–2008 CAPES Brafitec scholarship for exchange student (France-Brazil)
2007 Bachelor degree with honors from UFRGS (*láurea acadêmica*)

Languages

Brazilian Portuguese (mother tongue), **English** (fluent, TOEFL 117/120), **French** (fluent, DALF C1 98/100), **German** (fluent), **Spanish** (fluent), **Italian** (beginner), **Arabic** (beginner).

Language and software resources

- **PARSEME corpora**: `http://hdl.handle.net/11372/LRT-2842`
- **mwetoolkit tool**: `http://mwetoolkit.sourceforge.net/`
- **Veyn tool**: `https://github.com/zamp13/Veyn`
- **varIDE tool**: `https://gitlab.com/cpasquer/SharedTask2018_varIDE`
- **minimantics tool**: `https://github.com/ceramisch/minimantics`
- **MORPH dataset**: `http://pageperso.lis-lab.fr/~carlos.ramisch/?page=downloads/morph`
- **Compound dataset**: `http://pageperso.lis-lab.fr/~carlos.ramisch/?page=downloads/compounds`
- **Universal Dependencies contributor**: `http://universaldependencies.org`

Developed in collaboration with many colleagues, see URLs for details.

Publication stats

106 publications in total: 1 book, 7 book chapters, 11 peer-reviewed international journals, 58 peer-reviewed international conferences, 7 peer-reviewed national conferences, 11 book-proceeding edition, and 11 other. *h-index=25* (Google Scholar, 2025).

Full list of publications

2025

Ikram Belmadani, Benoit Favre, Richard Dufour, Frédéric Bechet, and Carlos **Ramisch**. Adaptation des connaissances médicales pour les grands modèles de langue : Stratégies et analyse comparative. In *Actes des 32ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN), volume 1 : articles scientifiques originaux*, pages 50–72, Marseille, France, 2025. Association pour le Traitement Automatique des Langues. `https://talnarchives.atala.org/TALN/TALN-2025/42.pdf`.

Ioana Ivan, Carlos **Ramisch**, and Alexis Nasr. Evaluating pretrained causal language models for synonymy. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 12533–12551, Vienna, Austria, 2025. ACL. **CORE2023 rank: A***. `https://aclanthology.org/2025.findings-acl.649/`.

Anna Mosolova, Marie Candito, and Carlos **Ramisch**. In the LLM era, word sense induction remains unsolved. In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 17161–17178, Vienna, Austria, 2025. Association for Computational Linguistics. **CORE2023 rank: A***. `https://aclanthology.org/2025.findings-acl.882/`.

Anna Mosolova, Marie Candito, and Carlos **Ramisch**. Raffinage des représentations des tokens dans les modèles de langue pré-entraînés avec l'apprentissage contrastif : une étude entre modèles et entre langues. In *Actes des 32ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN), volume 1 : articles scientifiques originaux*, pages 667–682, Marseille, France, 2025. ATALA. `https://talnarchives.atala.org/TALN/TALN-2025/139.pdf`.

Beatriz Sánchez-Cárdenas, Pablo Rienda, Nuria Medina-Medina, and Carlos **Ramisch**. Towards a tool for extracting specialized argument structures. In *Proceedings of the 4th International Conference on Multilingual Digital Terminology Today (MDTT 2025)*, volume 3990, Thessaloniki, Greece, 2025. CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3990/short22.pdf.

Manon Scholivet, Agata Savary, Louis Estève, Marie Candito, and Carlos **Ramisch**. SELEXINI — un grand corpus français, divers et parsé automatiquement. In *Actes des 32ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN), volume 2 : traductions d'articles publiés*, pages 57–57, Marseille, France, 2025. ATALA. https://talnarchives.atala.org/TALN/TALN-2025/35.pdf.

Manon Scholivet, Agata Savary, Louis Estève, Marie Candito, and Carlos **Ramisch**. SELEXINI – a large and diverse automatically parsed corpus of French. In *Proceedings of the 18th Workshop on Building and Using Comparable Corpora (BUCC)*, pages 83–98, Abu Dhabi, UAE, 2025. ACL. https://aclanthology.org/2025.bucc-1.10/.

2024

Leonardo Pinto Arata, Laura Ordonez Magro, Carlos **Ramisch**, Jonathan Grainger, and Arnaud Rey. The dynamics of multiword sequence extraction. *Quarterly Journal of Experimental Psychology*, 77(12):2439–2462, 2024. **Impact Factor: 1.5**. https://doi.org/10.1177/17470218241228548.

François Delon, Gabriel Bédubourg, Léo Bouscarrat, Jean-Baptiste Meynard, Aude Valois, Benjamin Queyriaux, Carlos **Ramisch**, and Marc Tanti. Infectious risk events and their novelty in event-based surveillance: new definitions and annotated corpus. *Language Resources and Evaluation*, 59(1):277–295, 2024. **Impact Factor: 1.7**. https://doi.org/10.1007/s10579-024-09728-w.

Anna Mosolova, Marie Candito, and Carlos **Ramisch**. Injecting Wiktionary to improve token-level contextual representations using contrastive learning. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 34–41, St. Julian's, Malta, 2024. ACL. **CORE2023 rank: A**. https://aclanthology.org/2024.eacl-short.5/.

Renata Ramisch, Carlos Ramisch, and Aline Villavicencio. Expressões multipalavras: fogo de palha ou osso duro de roer? In Helena Caseli and Maria das Graças Volpe Nunes, editors, *Processamento de Linguagem Natural: Conceitos, Técnicas e Aplicações em Português*, chapter 5. BPLN, 3 edition, 2024. https://brasileirapln.com/livro-pln/3a-edicao/parte-palavras/cap-mwe/cap-mwe.html.

Agata Savary, Daniel Zeman, Verginica Barbu Mititelu, Anabela Barreiro, Olesea Caftanatov, Marie-Catherine de Marneffe, Kaja Dobrovoljc, Gülşen Eryiğit, Voula Giouli, Bruno Guillaume, Stella Markantonatou, Nurit Melnik, Joakim Nivre, Atul Kr. Ojha, Carlos **Ramisch**, Abigail Walsh, Beata Wójtowicz, and Alina Wróblewska. UniDive: A COST action on universality, diversity and idiosyncrasy in language technology. In *Proceedings of the 3rd Annual Meeting of the Special Interest Group on Underresourced Languages at LREC-COLING 2024*, pages 372–382, Torino, Italy, 2024. ELRA and ICCL. https://aclanthology.org/2024.sigul-1.45/.

2023

Mathieu Constant, Marie Candito, Yannick Parmentier, Carlos **Ramisch**, and Agata Savary. Construction, exploitation et exploration de ressources linguistiques pour le traitement automatique des expressions polylexicales en français : le projet PARSEME-FR. In Lidia Becker, Julia Kuhn, Christina Ossenkop, Claudia Polzin-Haumann, and Elton Prifti, editors, *Digitale romanistische Sprachwissenschaft: Stand und Perspektiven*, Romanistisches Kolloquium, pages 219–250. Narr Francke Attempto Verlag GmbH + Co. KG, 2023. https://hal.science/hal-04995189.

Agata Savary, Cherifa Ben Khelil, Carlos **Ramisch**, Voula Giouli, Verginica Barbu Mititelu, Najet Hadj Mohamed, Cvetana Krstev, Chaya Liebeskind, Hongzhi Xu, Sara Stymne, Tunga Güngör, Thomas Pickard, Bruno Guillaume, Eduard Bejček, Archna Bhatia, Marie Candito, Polona Gantar, Uxoa Iñurrieta, Albert Gatt, Jolanta Kovalevskaitė, Timm Lichte, Nikola Ljubešić, Johanna Monti, Carla Parra Escartín, Mehrnoush Shamsfard, Ivelina Stoyanova, Veronika Vincze, and Abigail Walsh. PARSEME corpus release 1.3. In *Proceedings of the 19th Workshop on Multiword Expressions (MWE 2023)*, pages 24–35, Dubrovnik, Croatia, 2023. ACL. https://aclanthology.org/2023.mwe-1.6.

Agata Savary, Sara Stymne, Verginica Barbu Mititelu, Nathan Schneider, Carlos **Ramisch**, and Joakim Nivre. PARSEME meets universal dependencies: Getting on the same page in representing multiword expressions. *Northern European Journal of Language Technology*, 9:14, 2023. https://nejlt.ep.liu.se/article/view/4453.

Carlos **Ramisch**. *Multiword expressions in computational linguistics: down the rabbit hole and through the looking glass*. PhD thesis, Aix Marseille University, Marseille, France, 2023. https://theses.hal.science/tel-04216223 Slides https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2023/p01.pdf Video https://amupod.univ-amu.fr/video/27969-expressions-polylexicales-en-traitement-automatique-des-langues-carlos-ramisch-soute

Carlos **Ramisch**, Abigail Walsh, Thomas Blanchard, and Shiva Taslimipoor. A survey of MWE identification experiments: The devil is in the details. In *Proceedings of the 19th Workshop on Multiword Expressions (MWE 2023)*, pages 106–120, Dubrovnik, Croatia, 2023. ACL. https://aclanthology.org/2023.mwe-1.15.

2022

Archna Bhatia, Paul Cook, Shiva Taslimipoor, Marcos Garcia, and Carlos **Ramisch**, editors. *Proceedings of The 18th Workshop on Multiword Expressions at LREC 2022*. European Language Resources Association, Marseille, France, 2022. https://aclanthology.org/2022.mwe2022-1.

Fernando Zagatti, Paulo Augusto de Lima Medeiros, Esther da Cunha Soares, Lucas Nildaimon dos Santos Silva, Carlos **Ramisch**, and Livy Real. mwetoolkit-lib: Adaptation of the mwetoolkit as a python library and an application to MWE-based document clustering. In *Proceedings of The 18th Workshop on Multiword Expressions at LREC 2022*, pages 112–117, Marseille, France, 2022. European Language Resources Association. https://aclanthology.org/2022.mwe2022-1.16.

Nicolas Zampieri, Carlos Ramisch, Irina Illina, and Dominique Fohr. Identification des expressions polylexicales dans les tweets (identification of multiword expressions in tweets). In *Actes de la 29e Conférence sur le Traitement Automatique des Langues Naturelles. Volume 1 : conférence principale*, pages 365–373, Avignon, France, 2022. ATALA. https://aclanthology.org/2022.jeptalnrecital-taln.36.

Nicolas Zampieri, Carlos **Ramisch**, Irina Illina, and Dominique Fohr. Identification of multiword expressions in tweets for hate speech detection. In *Proceedings of the Language Resources and Evaluation Conference*, pages 202–210, Marseille, France, 2022. European Language Resources Association. https://aclanthology.org/2022.lrec-1.22.

2021

Léo Bouscarrat, Antoine Bonnefoy, Cécile Capponi, and Carlos **Ramisch**. AMU-EURANOVA at CASE 2021 task 1: Assessing the stability of multilingual BERT. In *Proceedings of the 4th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE 2021)*, pages 161–170, Online, 2021. ACL. https://aclanthology.org/2021.case-1.21.

Paul Cook, Jelena Mitrović, Carla Parra Escartín, Ashwini Vaidya, Petya Osenova, Shiva Taslimipoor, and Carlos **Ramisch**, editors. *Proceedings of the 17th Workshop on Multiword Expressions (MWE 2021)*. ACL, Online, 2021. https://aclanthology.org/2021.mwe-1.0.

2020

Cindy Aloui, Carlos **Ramisch**, Alexis Nasr, and Lucie Barque. SLICE: Supersense-based lightweight interpretable contextual embeddings. In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 3357–3370, Barcelona, Spain (Online), 2020. International Committee on Computational Linguistics. **CORE2020 rank: A**. https://www.aclweb.org/anthology/2020.coling-main.298.

Léo Bouscarrat, Antoine Bonnefoy, Cécile Capponi, and Carlos **Ramisch**. Multilingual enrichment of disease biomedical ontologies. In *Proceedings of the LREC 2020 Workshop on Multilingual Biomedical Text Processing (MultilingualBIO 2020)*, pages 21–28, Marseille, France, 2020. European Language Resources Association. https://www.aclweb.org/anthology/2020.multilingualbio-1.4/.

Marie Candito, Mathieu Constant, Carlos **Ramisch**, Agata Savary, Bruno Guillaume, Yannick Parmentier, and Silvio Ricardo Cordeiro. A French corpus annotated for multiword expressions and named entities. *Journal of Language Modelling*, 8(2):415–479, 2020. https://jlm.ipipan.waw.pl/index.php/JLM/article/view/265.

Stella Markantonatou, John McCrae, Jelena Mitrović, Carole Tiberius, Carlos **Ramisch**, Ashwini Vaidya, Petya Osenova, and Agata Savary, editors. *Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons*. ACL, Barcelona, Spain (Online), 2020. https://www.aclweb.org/anthology/2020.mwe-1.0.

Caroline Pasquer, Agata Savary, Jean-Yves Antoine, Carlos **Ramisch**, Nicolas Labroche, and Arnaud Giacometti. To be or not to be a verbal multiword expression: A quest for discriminating features, 2020. https://arxiv.org/abs/2007.11381.

Caroline Pasquer, Agata Savary, Carlos **Ramisch**, and Jean-Yves Antoine. Seen2Unseen at PARSEME shared task 2020: All roads do not lead to unseen verb-noun VMWEs. In *Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons*, pages 124–129, Barcelona, Spain (Online), 2020. ACL. https://www.aclweb.org/anthology/2020.mwe-1.16.

Caroline Pasquer, Agata Savary, Carlos **Ramisch**, and Jean-Yves Antoine. Verbal multiword expression identification: Do we need a sledgehammer to crack a nut? In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 3333–3345, Barcelona, Spain (Online), 2020. International Committee on Computational Linguistics. **CORE2020 rank: A**. https://www.aclweb.org/anthology/2020.coling-main.296.

Carlos **Ramisch**. Computational phraseology discovery in corpora with the MWETOOLKIT. In Gloria Corpas Pastor and Jean-Pierre Colson, editors, *Computational Phraseology*, volume 24 of *IVITRA Research in Linguistics and Literature*, pages 111–134. John Benjamins Publishing, 2020. Pre-print https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2020/p01.pdf Authenticated version https://doi.org/10.1075/ivitra.24.06ram.

Carlos **Ramisch**, Agata Savary, Bruno Guillaume, Jakub Waszczuk, Marie Candito, Ashwini Vaidya, Verginica Barbu Mititelu, Archna Bhatia, Uxoa Iñurrieta, Voula Giouli, Tunga Güngör, Menghan Jiang, Timm Lichte, Chaya Liebeskind, Johanna Monti, Renata Ramisch, Sara Stymne, Abigail Walsh, and Hongzhi Xu. Edition 1.2 of the PARSEME shared task on semi-supervised identification of verbal multiword expressions. In *Proceedings of the Joint Workshop on Multiword Expressions and Electronic Lexicons*, pages 107–118, Barcelona, Spain (Online), 2020. ACL. https://www.aclweb.org/anthology/2020.mwe-1.14.

2019

Mathieu Constant, Gülşen Eryiğit, Carlos **Ramisch**, Michael Rosner, and Gerold Schneider. Statistical MWE-aware parsing. In Yannick Parmentier and Jakub Waszczuk, editors, *Representation and Parsing of Multiword Expressions Current trends*, volume 3 of *Phraseology and Multiword Expressions*, pages 147–182. Language Science Press, Berlin, Germany, 2019. http://langsci-press.org/catalog/view/202/2026/1552-1.

Silvio Ricardo Cordeiro, Aline Villavicencio, Marco Idiart, and Carlos **Ramisch**. Unsupervised compositionality prediction of nominal compounds. *Computational Linguistics*, 45(1):1–57, 2019. **Impact Factor: 1.319**. http://www.mitpressjournals.org/doi/pdf/10.1162/coli_a_00341.

Beatriz Sánchez Cárdenas and Carlos **Ramisch**. Eliciting specialized frames from corpora using argument-structure extraction techniques. *Terminology: An International Journal of Theoretical and Applied Issues in Specialized Communication*, 25(1), 2019. **Impact Factor: 0.389**. Pre-print https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2019/p01.pdf Authenticated version https://doi.org/10.1075/term.00026.san.

Agata Savary, Silvio Ricardo Cordeiro, Timm Lichte, Carlos **Ramisch**, Uxoa Iñurrieta, and Voula Giouli. Literal occurrences of multiword expressions: Rare birds that cause a stir. *The Prague Bulletin of Mathematical Linguistics*, 112:5–54, 2019. https://ufal.mff.cuni.cz/pbml/112/art-savary-et-al.pdf.

Agata Savary, Silvio Ricardo Cordeiro, and Carlos **Ramisch**. Without lexicons, multiword expression identification will never fly: A position statement. In *Proceedings of the Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)*, pages 79–91, Florence, Italy, 2019. ACL. https://www.aclweb.org/anthology/W19-5110.

Manon Scholivet, Franck Dary, Alexis Nasr, Benoit Favre, and Carlos **Ramisch**. Typological features for multilingual delexicalised dependency parsing. In *Proceedings of the 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT 2019)*, Minneapolis, MN, USA, 2019. **CORE2018 rank: A**. https://aclweb.org/anthology/N19-1393.

Nicolas Zampieri, Carlos **Ramisch**, and Géraldine Damnati. The impact of word representations on sequential neural MWE identification. In *Proceedings of the Joint Workshop on Multiword Expressions and WordNet (MWE-WN 2019)*, pages 169–175, Florence, Italy, 2019. ACL. https://www.aclweb.org/anthology/W19-5121/.

2018

José Deulofeu, André Valli, Carlos **Ramisch**, and Alexis Nasr. Definition and representation of complex function words in dependency annotation schemes: regularity, endogeneity and parsing accuracy. In *Grammar and Corpora 2018 (GAC 2018)*, Paris, France, 2018. http://drehu.linguist.univ-paris-diderot.fr/gac-2018/abstracts/deulofeu_et_al.pdf.

Stella Markantonatou, Carlos **Ramisch**, Agata Savary, and Veronika Vincze, editors. *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*, volume 2 of *Phraseology and Multiword Expressions*. Language Science Press, Berlin, Germany, 2018. http://langsci-press.org/catalog/view/204/1339/1318-1.

Johanna Monti, Silvio Ricardo Cordeiro, Carlos **Ramisch**, Federico Sangati, Agata Savary, and Veronika Vincze. Advances in multiword expression identification for the italian language: The PARSEME shared task edition 1.1. In *Fifth Italian Conference on Computational Linguistics (CLIC-IT 2018)*, Torino, Italy, 2018. http://ceur-ws.org/Vol-2253/paper34.pdf.

Caroline Pasquer, Agata Savary, Jean-Yves Antoine, and Carlos **Ramisch**. Towards a variability measure for multiword expressions. In *Proceedings of the 16th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL 2018) - Short papers*, New Orleans, LA, USA, 2018. ACL. **CORE2018 rank: A**. https://aclweb.org/anthology/N18-2068.

Caroline Pasquer, Agata Savary, Carlos **Ramisch**, and Jean-Yves Antoine. If you've seen some, you've seen them all: Identifying variants of multiword expressions. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2582–2594. ACL, 2018. **CORE2018 rank: A**. https://aclweb.org/anthology/C18-1219.

Caroline Pasquer, Carlos **Ramisch**, Agata Savary, and Jean-Yves Antoine. VarIDE at PARSEME shared task 2018: Are variants really as alike as two peas in a pod? In *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*, pages 283–289. ACL, 2018. https://aclweb.org/anthology/W18-4932.

Agata Savary, Marie Candito, Verginica Barbu Mititelu, Eduard Bejček, Fabienne Cap, Slavomír Čéplö, Silvio Ricardo Cordeiro, Gülşen Eryiğit, Voula Giouli, Maarten van Gompel, Yaakov HaCohen-Kerner, Jolanta Kovalevskaitė, Simon Krek, Chaya Liebeskind, Johanna Monti, Carla Parra Escartín, Lonneke van der Plas, Behrang QasemiZadeh, Carlos **Ramisch**, Federico Sangati, Ivelina Stoyanova, and Veronika Vincze. PARSEME multilingual corpus of verbal multiword expressions. In Stella Markantonatou, Carlos Ramisch, Agata Savary, and Veronika Vincze, editors, *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*, volume 2 of *Phraseology and Multiword Expressions*. Language Science Press, Berlin, Germany, 2018. http://langsci-press.org/catalog/view/204/1344/1319-1.

Agata Savary, Carlos **Ramisch**, Jena D. Hwang, Nathan Schneider, Melanie Andresen, Sameer Pradhan, and Miriam R L Petruck, editors. *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*. ACL, 2018. https://aclweb.org/anthology/W18-49.

Manon Scholivet, Carlos **Ramisch**, and Silvio Ricardo Cordeiro. Sequence models and lexical resources for MWE identification in french. In Stella Markantonatou, Carlos Ramisch, Agata Savary, and Veronika Vincze, editors, *Multiword expressions at length and in depth: Extended papers from the MWE 2017 workshop*, volume 2 of *Phraseology and Multiword Expressions*. Language Science Press, Berlin, Germany, 2018. http://langsci-press.org/catalog/view/204/1651/1307-1.

Carlos **Ramisch**, Silvio Ricardo Cordeiro, Agata Savary, Veronika Vincze, Verginica Barbu Mititelu, Archna Bhatia, Maja Buljan, Marie Candito, Polona Gantar, Voula Giouli, Tunga Güngör, Abdelati Hawwari, Uxoa Iñurrieta, Jolanta Kovalevskaitė, Simon Krek, Timm Lichte, Chaya Liebeskind, Johanna Monti, Carla Parra Escartín, Behrang QasemiZadeh, Renata Ramisch, Nathan Schneider, Ivelina Stoyanova, Ashwini Vaidya, and Abigail Walsh. Edition 1.1 of the PARSEME shared task on automatic identification of verbal multiword expressions. In *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*, pages 222–240. ACL, 2018. https://aclweb.org/anthology/W18-4925.

Carlos **Ramisch**, Renata Ramisch, Leonardo Zilio, Aline Villavicencio, and Silvio Ricardo Cordeiro. A corpus study of verbal multiword expressions in Brazilian Portuguese. In *Computational Processing of*

the Portuguese Language 13th International Conference, PROPOR 2018, Canela, Brazil, September 24–26, 2018, Proceedings, Lecture Notes in Artificial Intelligence, Cham, Switzerland, 2018. Springer International Publishing. Pre-print https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2018/p01.pdf Authenticated version https://doi.org/10.1007/978-3-319-99722-3.

Carlos **Ramisch** and Aline Villavicencio. Computational treatment of multiword expressions. In Ruslan Mitkov, editor, *The Oxford Handbook of Computational Linguistics*. Oxford University Press, 2nd edition, 2018. https://doi.org/10.1093/oxfordhb/9780199573691.013.56.

Aline Villavicencio, Viviane Moreira, Alberto Abad, Helena Caseli, Pablo Gamallo, Carlos **Ramisch**, Hugo Gonçalo Oliveira, and Gustavo Henrique Paetzold, editors. *Computational Processing of the Portuguese Language 13th International Conference, PROPOR 2018, Canela, Brazil, September 24–26, 2018, Proceedings*. Lecture Notes in Artificial Intelligence. Springer International Publishing, Cham, Switzerland, 2018. https://www.springer.com/us/book/9783319997216.

Nicolas Zampieri, Manon Scholivet, Carlos **Ramisch**, and Benoit Favre. Veyn at PARSEME shared task 2018: Recurrent neural networks for VMWE identification. In *Proceedings of the Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018)*, pages 290–296. ACL, 2018. https://aclweb.org/anthology/W18-4933.

2017

Marie Candito, Mathieu Constant, Carlos **Ramisch**, Agata Savary, Yannick Parmentier, Caroline Pasquer, and Jean-Yves Antoine. Annotation d'expressions polylexicales verbales en français. In *Actes de la 24e conférence sur le Traitement Automatique des Langues Naturelles (TALN 2017) : articles courts*, pages 1–9, Orléans, France, 2017. ATALA. Pre-print https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2017/p01.pdf Authenticated version http://taln2017.cnrs.fr/wp-content/uploads/2017/06/actes_TALN_2017-vol2.pdf.

Mathieu Constant, Gülşen Eryiğit, Johanna Monti, Lonneke van der Plas, Carlos **Ramisch**, Michael Rosner, and Amalia Todirascu. Multiword expression processing: A survey. *Computational Linguistics*, 43(4):837–892, 2017. **Impact Factor: 1.319**. https://www.mitpressjournals.org/doi/pdf/10.1162/COLI_a_00302.

Stella Markantonatou, Carlos **Ramisch**, Agata Savary, and Veronika Vincze. *Proceedings of the 13th Workshop on MWEs*. ACL, Valencia, Spain, 2017. https://aclweb.org/anthology/W17-17.

Agata Savary, Carlos **Ramisch**, Silvio Ricardo Cordeiro, Federico Sangati, Veronika Vincze, Behrang QasemiZadeh, Marie Candito, Fabienne Cap, Voula Giouli, Ivelina Stoyanova, and Antoine Doucet. The PARSEME shared task on automatic identification of verbal multiword expressions. In *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*, pages 31–47, Valencia, Spain, 2017. ACL. https://aclweb.org/anthology/W17-1704.

Manon Scholivet and Carlos **Ramisch**. Identification of ambiguous multiword expressions using sequence models and lexical resources. In *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*, pages 167–175, Valencia, Spain, 2017. ACL. https://aclweb.org/anthology/W17-1723.

Carlos **Ramisch**. Putting the horses before the cart: Identifying multiword expressions before translation. In Ruslan Mitkov, editor, *Computational and Corpus-Based Phraseology - Second International Conference, Europhras 2017, London, UK, November 13-14, 2017, Proceedings*, pages 69–84, Cham, Switzerland, 2017. Springer International Publishing. Pre-print https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2017/p02.pdf Authenticated version https://doi.org/10.1007/978-3-319-69805-2_6.

Natalie Vargas, Carlos **Ramisch**, and Helena Caseli. Discovering light verb constructions and their translations from parallel corpora without word alignment. In *Proceedings of the 13th Workshop on Multiword Expressions (MWE 2017)*, pages 91–96, Valencia, Spain, 2017. ACL. https://aclweb.org/anthology/W17-1711.

Rodrigo Wilkens, Leonardo Zilio, Silvio Ricardo Cordeiro, Felipe Paula, Carlos **Ramisch**, Marco Idiart, and Aline Villavicencio. LexSubNC: A dataset of lexical substitution for nominal compounds. In *Proceedings of the 12th International Conference on Computational Semantics (IWCS 2017)*, Montpellier, France, 2017. https://aclweb.org/anthology/W17-6941.

2016

Silvio Ricardo Cordeiro, Carlos **Ramisch**, Marco Idiart, and Aline Villavicencio. MWE-aware corpus processing with the mwetoolkit and word embeddings. In *Proceedings of the PROPOR 2016 Workshop on Corpora and Tools for Processing Corpora*, Tomar, Portugal, 2016. http://ec.europa.eu/translation/portuguese/magazine/documents/folha50_propor_papers_en.pdf.

Silvio Ricardo Cordeiro, Carlos **Ramisch**, Marco Idiart, and Aline Villavicencio. Predicting the compositionality of nominal compounds: Giving word embeddings a hard time. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1986–1997, Berlin, Germany, 2016. ACL. **CORE2018 rank: A***. https://aclweb.org/anthology/P16-1187.

Silvio Ricardo Cordeiro, Carlos **Ramisch**, and Aline Villavicencio. mwetoolkit+sem: Integrating word embeddings in the mwetoolkit for semantic MWE processing. In *Proceedings of LREC 2016*, Portoroz, Slovenia, 2016. ELRA. http://www.lrec-conf.org/proceedings/lrec2016/pdf/347_Paper.pdf.

Silvio Ricardo Cordeiro, Carlos **Ramisch**, and Aline Villavicencio. Nominal compound compositionality: A multilingual lexicon and predictive model. In *Proceedings of the 7th PARSEME General Meeting*, Dubrovnik, Croatia, 2016. https://typo.uni-konstanz.de/parseme/images/Meeting/2016-09-26-Dubrovnik-meeting/PARSEME_7_paper_21.pdf.

Silvio Ricardo Cordeiro, Carlos **Ramisch**, and Aline Villavicencio. UFRGS&LIF at SemEval-2016 task 10: Rule-based MWE identification and predominant-supersense tagging. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, pages 910–917, San Diego, CA, USA, 2016. ACL. https://aclweb.org/anthology/S16-1140.

Carlos **Ramisch**, Silvio Ricardo Cordeiro, and Aline Villavicencio. Filtering and measuring the intrinsic quality of human compositionality judgments. In *Proceedings of the 12th Workshop on MWEs*, pages 32–37, Berlin, Germany, 2016. ACL. https://aclweb.org/anthology/W16-1804.

Carlos **Ramisch**, Silvio Ricardo Cordeiro, Leonardo Zilio, Marco Idiart, Aline Villavicencio, and Rodrigo Wilkens. How naked is the naked truth? A multilingual lexicon of nominal compound compositionality. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 156–161, Berlin, Germany, 2016. ACL. **CORE2018 rank: A***. https://aclweb.org/anthology/P16-2026.

Carlos **Ramisch**, Alexis Nasr, André Valli, and José Deulofeu. DeQue: A lexicon of complex prepositions and conjunctions in French. In *Proceedings of LREC 2016*, Portoroz, Slovenia, 2016. ELRA. http://www.lrec-conf.org/proceedings/lrec2016/pdf/347_Paper.pdf.

2015

Silvio Ricardo Cordeiro, Carlos **Ramisch**, and Aline Villavicencio. Token-based MWE identification strategies in the mwetoolkit. In *Proceedings of the 4th PARSEME General Meeting*, Valetta, Malta, 2015. https://typo.uni-konstanz.de/parseme/images/Meeting/2015-03-19-Malta-meeting/WG2-WG3-CORDEIRO-et-al-abstract.pdf.

Alexis Nasr, José Deulofeu, André Valli, and Carlos **Ramisch**. Complex function words in french: Representation and processing. In *Proceedings of the 4th PARSEME General Meeting*, Valetta, Malta, 2015. https://typo.uni-konstanz.de/parseme/images/Meeting/2015-03-19-Malta-meeting/WG3-NASR-DEULOFEU-VALLI-RAMISCH-abstract.pdf.

Alexis Nasr, Carlos **Ramisch**, José Deulofeu, and André Valli. Joint dependency parsing and multiword expression tokenization. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 1116–1126, Beijing, China, 2015. ACL. **CORE2018 rank: A***. https://aclweb.org/anthology/P15-1108.

Alexandre Rondon, Helena Caseli, and Carlos **Ramisch**. Never-ending multiword expressions learning. In *Proceedings of the 11th Workshop on MWEs*, pages 45–53, Denver, CO, USA, 2015. ACL. https://aclweb.org/anthology/W15-0908.

Carlos **Ramisch**. *Multiword Expressions Acquisition: A Generic and Open Framework*, volume XIV of *Theory and Applications of Natural Language Processing*. Springer, 2015. https://doi.org/10.1007/978-3-319-09207-2.

2014

Bruno Laranjeira, Viviane Moreira, Aline Villavicencio, Carlos **Ramisch**, and Maria José Finatto. Comparing the quality of focused crawlers and of the linguistic resources obtained from them. In *Proceedings of*

LREC 2014, Reykjavik, Iceland, 2014. ELRA. http://www.lrec-conf.org/proceedings/lrec2014/pdf/1095_Paper.pdf.

Muntsa Padró, Marco Idiart, Aline Villavicencio, and Carlos **Ramisch**. Comparing similarity measures for distributional thesauri. In *Proceedings of LREC 2014*, Reykjavik, Iceland, 2014. ELRA. http://www.lrec-conf.org/proceedings/lrec2014/pdf/619_Paper.pdf.

Muntsa Padró, Marco Idiart, Aline Villavicencio, and Carlos **Ramisch**. Nothing like good old frequency: Studying context filters for distributional thesauri. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP 2014) - short papers*, Doha, Qatar, 2014. **CORE2018 rank: A**. https://aclweb.org/anthology/D14-1047.

2013

Magali Sanches Duran, Carolina Evaristo Scarton, Sandra Maria Aluísio, and Carlos **Ramisch**. Identifying pronominal verbs: Towards automatic disambiguation of the clitic 'se' in Portuguese. In *Proceedings of the 9th Workshop on MWEs*, pages 93–100, Atlanta, Georgia, USA, 2013. ACL. https://aclweb.org/anthology/W13-1014.

Valia Kordoni, Carlos **Ramisch**, and Aline Villavicencio, editors. *Proceedings of the of the 9th Workshop on Multiword Expressions (MWE 2013)*. ACL, Atlanta, GA, USA, 2013. https://aclweb.org/anthology/W13-10.

Carlos **Ramisch**, Laurent Besacier, and Alexander Kobzar. How hard is it to automatically translate phrasal verbs from English to French? In *MT Summit 2013 Workshop on Multi-word Units in Machine Translation and Translation Technology*, Nice, France, 2013. https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2013/p01.pdf.

Carlos **Ramisch**, Aline Villavicencio, and Valia Kordoni. Introduction to the special issue on multiword expressions: From theory to practice and use. *ACM Transactions on Speech and Language Processing*, 10(2):3:1–3:10, 2013. **Impact factor: 2.491**. Pre-print https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2013/p02.pdf Authenticated version https://doi.org/10.1145/2483691.2483692.

Leonardo Zilio, Carlos **Ramisch**, and Maria José Finatto. Desenvolvimento de um recurso léxico com papéis semânticos para o português. *Linguamática*, 5(2):23–41, 2013. http://www.linguamatica.com/index.php/linguamatica/article/view/167.

2012

Roger Granada, Lucelene Lopes, Carlos **Ramisch**, Cassia Trojahn Dos Santos, Renata Vieira, and Aline Villavicencio. A comparable corpus based on aligned multilingual ontologies. In *Proceedings of the ACL 2012 First Workshop on Multilingual Modeling (MM 2012)*, Jeju, Republic of Korea, 2012. ACL. https://aclweb.org/anthology/W12-3904.

Mathieu Mangeot and Carlos **Ramisch**. A serious lexical game for building a portuguese lexical-semantic network. In *Proceedings of the ACL 2012 3rd Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources and their Applications to NLP*, Jeju, Republic of Korea, 2012. ACL. https://aclweb.org/anthology/W12-4002.

Carlos **Ramisch**. *A generic and open framework for multiword expressions treatment: from acquisition to applications*. PhD thesis, University of Grenoble (France) and Federal University of Rio Grande do Sul (Brazil), Grenoble, France, 2012. http://theses.fr/2012GRENM059.

Carlos **Ramisch**. A generic framework for multiword expressions treatment: from acquisition to applications. In *Proceedings of the ACL 2012 Student Research Workshop*, Jeju, Republic of Korea, 2012. ACL. https://aclweb.org/anthology/W12-3311.

Carlos **Ramisch**. Une plate-forme générique et ouverte pour le traitement des expressions polylexicales. In Jorge Mauricio Molina Mejia and Didier Schwab, editors, *Actes de 14e Rencontres des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL 2012)*, Grenoble, France, 2012. https://aclweb.org/anthology/F12-3011.

Carlos **Ramisch**, Vitor De Araujo, and Aline Villavicencio. A broad evaluation of techniques for automatic acquisition of multiword expressions. In *Proceedings of the ACL 2012 Student Research Workshop*, Jeju, Republic of Korea, 2012. ACL. https://aclweb.org/anthology/W12-3301.

Aline Villavicencio, Marco Idiart, Carlos **Ramisch**, Vitor Araujo, Beracah Yankama, and Robert Berwick. Get out but don't fall down: verb-particle constructions in child language. In *Proceedings of the Workshop*

on Computational Models of Language Acquisition and Loss, pages 43–50, Avignon, France, 2012. ACL. https://aclweb.org/anthology/W12-0911.

2011

Vitor De Araujo, Carlos **Ramisch**, and Aline Villavicencio. Fast and flexible MWE candidate generation with the mwetoolkit. In *Proceedings of the Workshop on MWEs: from Parsing and Generation to the Real World*, pages 134–136, Portland, OR, USA, 2011. ACL. https://aclweb.org/anthology/W11-0822.

Magali Sanches Duran and Carlos **Ramisch**. How do you feel? investigating lexical-syntactic patterns in sentiment expression. In *Proceedings of Corpus Linguistics 2011: Discourse and Corpus Linguistics Conference*, Birmingham, UK, 2011. https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2011/p07.pdf.

Magali Sanches Duran, Carlos **Ramisch**, Sandra Maria Aluísio, and Aline Villavicencio. Identifying and analyzing Brazilian Portuguese complex predicates. In *Proceedings of the Workshop on MWEs: from Parsing and Generation to the Real World*, pages 74–82, Portland, OR, USA, 2011. ACL. https://aclweb.org/anthology/W11-0812.

Valia Kordoni, Carlos **Ramisch**, and Aline Villavicencio, editors. *Proceedings of the ACL Workshop on Multiword Expressions: from Parsing and Generation to the Real World (MWE 2011)*. ACL, Portland, OR, USA, 2011. https://aclweb.org/anthology/W11-08.

2010

Helena Caseli, Carlos **Ramisch**, Maria das Graças Volpe Nunes, and Aline Villavicencio. Alignment-based extraction of multiword expressions. *Language Resources and Evaluation Special Issue on Multiword expression: hard going or plain sailing*, 44(1-2):59–77, 2010. **Impact Factor: 0.656**. Pre-print https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2010/p07.pdf Authenticated version https://doi.org/10.1007/s10579-009-9097-9.

Éric Laporte, Preslav Nakov, Carlos **Ramisch**, and Aline Villavicencio, editors. *Proceedings of the COLING Workshop on Multiword Expressions: from Theory to Applications (MWE 2010)*. The Coling 2010 Organizing Committee, Beijing, China, 2010. https://aclweb.org/anthology/W10-37.

Evita Linardaki, Carlos **Ramisch**, Aline Villavicencio, and Aggeliki Fotopoulou. Towards the construction of language resources for Greek multiword expressions: Extraction and evaluation. In Stelios Piperidis, Milena Slavcheva, and Cristina Vertan, editors, *Proceedings of the LREC Workshop on Exploitation of multilingual resources and tools for Central and (South) Eastern European Languages*, pages 31–40, Valetta, Malta, 2010. ELRA. http://www.lrec-conf.org/proceedings/lrec2010/workshops/W22.pdf.

Carlos **Ramisch**, Helena Caseli, Aline Villavicencio, André Machado, and Maria José Finatto. A hybrid approach for multiword expression identification. In Thiago A. S. Pardo, António Branco, Aldebaro Klautau, Renata Vieira, and Vera Lúcia Strube de Lima, editors, *Proceedings of the 9th International Conference on Computational Processing of Portuguese Language (PROPOR 2010)*, volume 6001 of *Lecture Notes in Computer Science (Lecture Notes in Artificail Intelligence)*, pages 65–74, Porto Alegre, RS, Brazil, 2010. Springer. Pre-print https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2010/p11.pdf Authenticated version https://doi.org/10.1007/978-3-642-12320-7_9.

Carlos **Ramisch**, Aline Villavicencio, and Christian Boitet. Multiword expressions in the wild? the mwetoolkit comes in handy. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010) — Demonstrations*, pages 57–60, Beijing, China, 2010. The Coling 2010 Organizing Committee. **CORE2018 rank: A**. https://aclweb.org/anthology/C10-3015.

Carlos **Ramisch**, Aline Villavicencio, and Christian Boitet. mwetoolkit: a framework for multiword expression identification. In Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odjik, Stelios Piperidis, Mike Rosner, and Daniel Tapias, editors, *Proceedings of LREC 2010*, Valetta, Malta, 2010. ELRA. http://www.lrec-conf.org/proceedings/lrec2010/pdf/803_Paper.pdf.

Carlos **Ramisch**, Aline Villavicencio, and Christian Boitet. Web-based and combined language models: a case study on noun compound identification. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING 2010)*, pages 1041–1049, Beijing, China, 2010. The Coling 2010 Organizing Committee. **CORE2018 rank: A**. https://aclweb.org/anthology/C10-2120.

Aline Villavicencio, Carlos **Ramisch**, André Machado, Helena Caseli, and Maria José Finatto. Identificação de expressões multipalavra em domínios específicos. *Linguamática*, 2(1):15–33, 2010. http://linguamatica.com/index.php/linguamatica/article/download/43/57.

2009

Carlos **Ramisch**. Multi-word terminology extraction for domain-specific documents. Master's thesis, École Nationale Supérieure d'Informatique et de Mathématiques Appliquées, Grenoble, France, 2009. https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2009/p01.pdf.

2008

Christian Boitet, Valérie Bellynck, Mathieu Mangeot, and Carlos **Ramisch**. Towards higher quality internal and outside multilingualization of web sites. In Pushpak Bhattacharyya, editor, *Summer Workshop on Ontology, NLP, Personalization and IE/IR (ONII-08)*, Mumbai, India, 2008. Centre for Indian Language Technology (CFILT), IITB (Indian Institute of Technology Bombay). https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2008/p03.pdf.

Carlos **Ramisch**, Paulo Schreiner, Marco Idiart, and Aline Villavicencio. An evaluation of methods for the extraction of multiword expressions. In *Proceedings of the LREC Workshop Towards a Shared Task for MWEs (MWE 2008)*, pages 50–53, Marrakech, Morocco, 2008. https://pageperso.lis-lab.fr/carlos.ramisch/download_files/publications/2008/p01.pdf.

Carlos **Ramisch**, Aline Villavicencio, Leonardo Moura, and Marco Idiart. Picking them up and figuring them out: Verb-particle constructions, noise and idiomaticity. In Alex Clark and Kristina Toutanova, editors, *Proceedings of the Twelfth Conference on Natural Language Learning (CoNLL 2008)*, pages 49–56, Manchester, UK, 2008. ACL. **CORE2018 rank: A**. https://aclweb.org/anthology/W08-2107.

2007

A. Villavicencio, V. Kordoni, Y. Zhang, M. Idiart, and C. **Ramisch**. Validation and evaluation of automatically acquired multiword expressions for grammar engineering. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, pages 1034–1043, Prague, Czech Republic, 2007. ACL. **CORE2018 rank: A**. https://aclweb.org/anthology/D07-1110.