

# The Analytics of Word Sociology

Violeta Seretan

## References

- [Archer2006] Vincent Archer. 2006. Acquisition semi-automatique de collocations à partir de corpus monolingues et multilingues comparables. In *Proceedings of Rencontre des Etudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL 2006)*.
- [Baldwin and Villavicencio2002] Timothy Baldwin and Aline Villavicencio. 2002. Extracting the unextractable: A case study on verb-particles. In *Proceedings of the Sixth Conference on Computational Natural Language Learning (CoNLL 2002)*, pages 98–104, Taipei, Taiwan.
- [Bally1909] Charles Bally. 1909. *Traité de stylistique française*. Klincksieck, Paris.
- [Bannard et al.2003] Colin Bannard, Timothy Baldwin, and Alex Lascarides. 2003. A statistical approach to the semantics of verb-particles. In *Proceedings of the ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, pages 65–72, Sapporo, Japan, July. Association for Computational Linguistics.
- [Bartsch2004] Sabine Bartsch. 2004. *Structural and Functional Properties of Collocations in English. A Corpus Study of Lexical and Pragmatic Constraints on Lexical Cooccurrence*. Gunter Narr Verlag, Tübingen.
- [Basili et al.1994] Roberto Basili, Maria Teresa Pazienza, and Paola Velardi. 1994. A “not-so-shallow” parser for collocational analysis. In *Proceedings of the 15th conference on Computational linguistics*, pages 447–453, Kyoto, Japan.
- [Beauchesne2001] Jacques Beauchesne. 2001. *Dictionnaire des cooccurrences*. Guérin, Montréal.
- [Benson et al.1986] Morton Benson, Evelyn Benson, and Robert Ilson. 1986. *The BBI Dictionary of English Word Combinations*. John Benjamins, Amsterdam/Philadelphia.
- [Benson1990] Morton Benson. 1990. Collocations and general-purpose dictionaries. *International Journal of Lexicography*, 3(1):23–35.
- [Bourigault1992] Didier Bourigault. 1992. LEXTER, vers un outil linguistique d’aide à l’acquisition des connaissances. In *Actes des 3èmes Journées d’acquisition des Connaissances*, Dourdan, France.
- [Breidt1993] Elisabeth Breidt. 1993. Extraction of V-N-collocations from text corpora: A feasibility study for German. In *Proceedings of the Workshop on Very Large Corpora: Academic and Industrial Perspectives*, pages 74–83, Columbus, USA.
- [Calzolari and Bindi1990] Nicoletta Calzolari and Remo Bindi. 1990. Acquisition of lexical information from a large textual Italian corpus. In *Proceedings of the 13th International Conference on Computational Linguistics*, pages 54–59, Helsinki, Finland.
- [Charest et al.2012] Simon Charest, Jean Fontaine, and Jean Saint-Germain. 2012. *Le Grand Druide des cooccurrences*. Les ditions Druide.
- [Choueka1988] Yaacov Choueka. 1988. Looking for needles in a haystack, or locating interesting collocational expressions in large textual databases. In *Proceedings of the International Conference on User-Oriented Content-Based Text and Image Handling*, pages 609–623, Cambridge, USA.

- [Church and Hanks1990] Kenneth Church and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics*, 16(1):22–29.
- [Church et al.1989] Kenneth Church, William Gale, Patrick Hanks, and Donald Hindle. 1989. Parsing, word associations and typical predicate-argument relations. In *Proceedings of the International Workshop on Parsing Technologies*, pages 103–112, Pittsburgh. Carnegie Mellon University.
- [Coseriu1967] Eugenio Coseriu. 1967. Lexikalische Solidaritäten. *Poetica*, (1):293–303.
- [Cowie1978] Anthony P. Cowie. 1978. The place of illustrative material and collocations in the design of a learner’s dictionary. In P. Strevens, editor, *In Honour of A.S. Hornby*, pages 127–139. Oxford University Press, Oxford.
- [Cruse1986] D. Alan Cruse. 1986. *Lexical Semantics*. Cambridge University Press, Cambridge.
- [Dagan and Church1994] Ido Dagan and Kenneth Church. 1994. *Termight*: Identifying and translating technical terminology. In *Proceedings of the 4th Conference on Applied Natural Language Processing (ANLP)*, pages 34–40, Stuttgart, Germany.
- [Daille1994] Béatrice Daille. 1994. *Approche mixte pour l’extraction automatique de terminologie : statistiques lexicales et filtres linguistiques*. Ph.D. thesis, Université Paris 7.
- [Dias2003] Gaël Dias. 2003. Multiword unit hybrid extraction. In *Proceedings of the ACL Workshop on Multiword Expressions*, pages 41–48, Sapporo, Japan.
- [Dunning1993] Ted Dunning. 1993. Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19(1):61–74.
- [Ellis2008] Nick Ellis. 2008. Phraseology: The periphery and the heart of language. In Fanny Meunier and Sylviane Granger, editors, *Phraseology in Foreign Language and Teaching*, pages 1–13. John Benjamins, Amsterdam/Philadelphia.
- [Evert and Krenn2001] Stefan Evert and Brigitte Krenn. 2001. Methods for the qualitative evaluation of lexical association measures. In *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*, pages 188–195, Toulouse, France.
- [Evert and Krenn2005] Stefan Evert and Brigitte Krenn. 2005. Using small random samples for the manual evaluation of statistical association measures. *Computer Speech & Language*, 19(4):450–466.
- [Evert2004] Stefan Evert. 2004. *The Statistics of Word Cooccurrences: Word Pairs and Collocations*. Ph.D. thesis, University of Stuttgart.
- [Fazly2007] Afsaneh Fazly. 2007. *Automatic Acquisition of Lexical Knowledge about Multiword Predicates*. Ph.D. thesis, University of Toronto.
- [Fellbaum2007] Christiane Fellbaum, editor. 2007. *Idioms and Collocations: Corpus-based Linguistic and Lexicographic Studies*. Continuum, London.
- [Fillmore et al.1988] Charles Fillmore, Paul Kay, and Catherine O’Connor. 1988. Regularity and idiomativity in grammatical constructions: The case of *let alone*. *Language*, 64(3):501–538.
- [Firth1968] John R. Firth. 1968. A synopsis of linguistic theory, 1930–55. In F.R. Palmer, editor, *Selected papers of J. R. Firth, 1952–1959*, pages 168–205. Indiana University Press, Bloomington.
- [Fontenelle1992] Thierry Fontenelle. 1992. Collocation acquisition from a corpus or from a dictionary: a comparison. *Proceedings I-II. Papers submitted to the 5th EURALEX International Congress on Lexicography in Tampere*, pages 221–228.
- [Fontenelle1997] Thierry Fontenelle. 1997. *Turning a bilingual dictionary into a lexical-semantic database*. Max Niemeyer Verlag, Tübingen.
- [Gitsaki1996] Christina Gitsaki. 1996. *The Development of ESL Collocational Knowledge*. Ph.D. thesis, University of Queensland.

- [Goldman et al.2001] Jean-Philippe Goldman, Luka Nerima, and Eric Wehrli. 2001. Collocation extraction using a syntactic parser. In *Proceedings of the ACL Workshop on Collocation: Computational Extraction, Analysis and Exploitation*, pages 61–66, Toulouse, France.
- [Granger and Meunier2008] Sylviane Granger and Fanny Meunier, editors. 2008. *Phraseology: An interdisciplinary perspective*. John Benjamins, Amsterdam/Philadelphia.
- [Grossmann and Tutin2003] Francis Grossmann and Agnès Tutin, editors. 2003. *Les Collocations. Analyse et traitement*. Éditions De Werelt, Amsterdam.
- [Halliday and Hasan1976] Michael Alexander Kirkwood Halliday and Ruqaiya Hasan. 1976. *Cohesion in English*. Longman, London.
- [Hausmann1985] Franz Josef Hausmann. 1985. Kollokationen im deutschen Wörterbuch. Ein Beitrag zur Theorie des lexikographischen Beispiels. In Henning Bergenholtz and Joachim Mugdan, editors, *Lexikographie und Grammatik. Akten des Essener Kolloquiums zur Grammatik im Wörterbuch*, Lexicographica. Series Major 3, pages 118–129.
- [Hausmann1989] Franz Josef Hausmann. 1989. Le dictionnaire de collocations. In F.J. Hausmann, O. Reichmann, H.E. Wiegand, and L. Zgusta, editors, *Wörterbücher: Ein internationales Handbuch zur Lexicographie. Dictionaries, Dictionnaires*, pages 1010–1019. de Gruyter, Berlin.
- [Heid and Raab1989] Ulrich Heid and Sybille Raab. 1989. Collocations in multilingual generation. In *Proceeding of the Fourth Conference of the European Chapter of the Association for Computational Linguistics (EACL'89)*, pages 130–136, Manchester, England.
- [Heid1994] Ulrich Heid. 1994. On ways words work together – research topics in lexical combinatorics. In *Proceedings of the 6th Euralex International Congress on Lexicography (EURALEX '94)*, pages 226–257, Amsterdam, The Netherlands.
- [Heylen et al.1994] Dirk Heylen, Kerry G. Maxwell, and Marc Verhagen. 1994. Lexical functions and machine translation. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING 1994)*, pages 1240–1244, Kyoto, Japan.
- [Hoey1991] Michael Hoey. 1991. *Patterns of Lexis in Text*. Oxford University Press, Oxford.
- [Ikehara et al.1995] Satoru Ikehara, Satoshi Shirai, and Tsukasa Kawaoka. 1995. Automatic extraction of uninterrupted collocations by n-gram statistics. In *Proceedings of first Annual Meeting of the Association for Natural Language Processing*, pages 313–316, Tokyo, Japan.
- [Jacquemin et al.1997] Christian Jacquemin, Judith L. Klavans, and Evelyne Tzoukermann. 1997. Expansion of multi-word terms for indexing and retrieval using morphology and syntax. In *Proceedings of the 35th annual meeting on Association for Computational Linguistics*, pages 24–31, Morristown, NJ, USA.
- [Jacquemin2001] Christian Jacquemin. 2001. *Spotting and Discovering Terms through Natural Language Processing*. MIT Press, Cambridge MA.
- [Justeson and Katz1995] John S. Justeson and Slava M. Katz. 1995. Technical terminology: Some linguistic properties and an algorithm for identification in text. *Natural Language Engineering*, 1(1):9–27.
- [Kilgarriff et al.2004] Adam Kilgarriff, Pavel Rychly, Pavel Smrz, and David Tugwell. 2004. The Sketch Engine. In *Proceedings of the Eleventh EURALEX International Congress*, pages 105–116, Lorient, France.
- [Kilgarriff1996] Adam Kilgarriff. 1996. Which words are particularly characteristic of a text? A survey of statistical approaches. In *Proceedings of AISB Workshop on Language Engineering for Document Analysis and Recognition*, pages 33–40, Sussex, UK.

- [Kim et al.1999] Seonho Kim, Zooil Yang, Mansuk Song, and Jung-Ho Ahn. 1999. Retrieving collocations from Korean text. In *Proceedings of the 1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora*, pages 71–81, Maryland, USA.
- [Kjellmer1987] Göran Kjellmer. 1987. Aspects of English collocations. In Willem Meijs, editor, *Corpus Linguistics and Beyond*, pages 133–140. Rodopi, Amsterdam.
- [Kjellmer1994] Göran Kjellmer. 1994. *A Dictionary of English Collocations*. Clarendon Press, Oxford.
- [Krenn and Evert2001] Brigitte Krenn and Stefan Evert. 2001. Can we do better than frequency? A case study on extracting PP-verb collocations. In *Proceedings of the ACL Workshop on Collocation: Computational Extraction, Analysis and Exploitation*, pages 39–46, Toulouse, France.
- [Krenn2000] Brigitte Krenn. 2000. *The Usual Suspects: Data-Oriented Models for Identification and Representation of Lexical Collocations*, volume 7. German Research Center for Artificial Intelligence and Saarland University Dissertations in Computational Linguistics and Language Technology, Saarbrücken, Germany.
- [Lafon1984] Pierre Lafon. 1984. *Dépouillements et statistiques en lexicométrie*. Slatkine – Champion, Genève/Paris.
- [Lea and Runcie2002] Diana Lea and Moira Runcie, editors. 2002. *Oxford Collocations Dictionary for Students of English*. Oxford University Press, Oxford.
- [Lewis2000] Michael Lewis, editor. 2000. *Teaching Collocations. Further Developments in the Lexical Approach*. Language Teaching Publications, Hove.
- [Lin1998] Dekang Lin. 1998. Extracting collocations from text corpora. In *First Workshop on Computational Terminology*, pages 57–63, Montreal, Canada.
- [Louw1993] Bill Louw. 1993. Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In Mona Baker, Gill Francis, and Elena Tognini-Bonelli, editors, *Text and Technology: In Honour of John Sinclair*, pages 157–176. John Benjamins, Amsterdam.
- [Makkai1972] Adam Makkai. 1972. *Idiom Structure in English*. Mouton, The Hague.
- [Manning and Schütze1999] Christopher D. Manning and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing*. MIT Press, Cambridge, Mass.
- [McKeown and Radev2000] Kathleen R. McKeown and Dragomir R. Radev. 2000. Collocations. In Robert Dale, Hermann Moisl, and Harold Somers, editors, *A Handbook of Natural Language Processing*, pages 507–523. Marcel Dekker, New York, USA.
- [Mel’čuk1998] Igor Mel’čuk. 1998. Collocations and lexical functions. In Anthony P. Cowie, editor, *Phraseology. Theory, Analysis, and Applications*, pages 23–53. Clarendon Press, Oxford.
- [Michiels1998] Archibald Michiels. 1998. The DEFI matcher. In T. Fontenelle, P. Hilgsmann, A. Michiels, A. Moulin, and S. Theissen, editors, *Proceedings of the Eighth Euralex International Congress (Euralex 1998)*, pages 203–211, Liège. University of Liège.
- [Moon1998] Rosamund Moon. 1998. *Fixed expressions and idioms in English: A corpus-based approach*. Clarendon Press Oxford, Oxford.
- [Orliac and Dillinger2003] Brigitte Orliac and Mike Dillinger. 2003. Collocation extraction for machine translation. In *Proceedings of Machine Translation Summit IX*, pages 292–298, New Orleans, Louisiana, USA.
- [Pearce2001] Darren Pearce. 2001. Synonymy in collocation extraction. In *Proceedings of the NAACL Workshop on WordNet and Other Lexical Resources: Applications, Extensions and Customizations*, pages 41–46, Pittsburgh, USA.
- [Pecina2008] Pavel Pecina. 2008. *Lexical Association Measures: Collocation Extraction*. Ph.D. thesis, Charles University in Prague.

- [Sag et al.2002] Ivan A. Sag, Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger. 2002. Multiword expressions: A pain in the neck for NLP. In *Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics (CICLING 2002)*, pages 1–15, Mexico City.
- [Schulte im Walde2003] Sabine Schulte im Walde. 2003. A collocation database for German verbs and nouns. In *Proceedings of 7th Conference on Computational Lexicography and Corpus Research*, Budapest, Hungary.
- [Seretan2011] Violeta Seretan. 2011. A collocation-driven approach to text summarization. In *Actes de la 18e conférence sur le Traitement Automatique des Langues Naturelles (TALN 2011)*, pages 9–14, Montpellier, France.
- [Sinclair1987] John Sinclair, editor. 1987. *Collins COBUILD English Language Dictionary*. Collins.
- [Sinclair1991] John Sinclair. 1991. *Corpus, Concordance, Collocation*. Oxford University Press, Oxford.
- [Sinclair1995] John Sinclair. 1995. *Collins Cobuild English Dictionary*. Harper Collins, London.
- [Smadja1993] Frank Smadja. 1993. Retrieving collocations from text: Xtract. *Computational Linguistics*, 19(1):143–177.
- [Stubbs2002] Michael Stubbs. 2002. *Words and Phrases: Corpus Studies of Lexical Semantics*. Blackwell, Oxford.
- [Todiraşcu et al.2008] Amalia Todiraşcu, Dan Tufiş, Ulrich Heid, Christopher Gledhill, Dan Ştefănescu, Marion Weller, and François Rousselot. 2008. A hybrid approach to extracting and classifying verb+noun constructions. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco.
- [Tutin2004] Agnès Tutin. 2004. Pour une modélisation dynamique des collocations dans les textes. In *Proceedings of the Eleventh EURALEX International Congress*, pages 207–219, Lorient, France.
- [van der Wouden1997] Ton van der Wouden. 1997. *Negative Contexts. Collocation, polarity, and multiple negation*. Routledge, London and New York.
- [Villada Moirón2005] María Begoña Villada Moirón. 2005. *Data-driven identification of fixed expressions and their modifiability*. Ph.D. thesis, University of Groningen.
- [Wermter and Hahn2004] Joachim Wermter and Udo Hahn. 2004. Collocation extraction based on modifiability statistics. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004)*, pages 980–986, Geneva, Switzerland.
- [Wu and Zhou2003] Hau Wu and Ming Zhou. 2003. Synonymous collocation extraction using translation information. In *Proceeding of the Annual Meeting of the Association for Computational Linguistics (ACL 2003)*, pages 120–127, Sapporo, Japan.
- [Yarowsky1993] David Yarowsky. 1993. One sense per collocation. In *Proceedings of ARPA Human Language Technology Workshop*, pages 266–271, Princeton.
- [Zinsmeister and Heid2003] Heike Zinsmeister and Ulrich Heid. 2003. Significant triples: Adjective+Noun+Verb combinations. In *Proceedings of the 7th Conference on Computational Lexicography and Text Research (Complex 2003)*, Budapest, Hungary.