

Referências

- (Afonso et al., 2001) Susana Afonso, Eckhard Bick, Renato Haber e Diana Santos. Floresta sintá(c)tica: um treebank para o português. Em Anabela Gonçalves e Clara Nunes Correia, editoras, *Actas do XVII Encontro Nacional da Associação Portuguesa de Linguística (APL 2001)*. Lisboa, Portugal. 2-4 de Outubro de 2001. p. 533–545. APL. <http://www.linguateca.pt/documentos/AfonsoetalAPL2001.pdf>. 14, 31
- (Afonso et al., 2002) Susana Afonso, Eckhard Bick, Renato Haber e Diana Santos. Floresta sintá(c)tica: a treebank for Portuguese. Em Manuel González Rodrigues e Carmen Paz Suarez Araujo, editores, *Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002)*. Las Palmas de Gran Canaria, Espanha. 29-31 de Maio de 2002. p. 1698–1703. ELRA. <http://www.linguateca.pt/documentos/AfonsoetalLREC2002.pdf>. 14
- (Agrawal e Srikant, 1994) Rakesh Agrawal e Ramakrishnan Srikant. Fast algorithms for mining association rules. Em Jorge B. Bocca, Matthias Jarke e Carlo Zaniolo, editores, *Proceedings of the 20th International Conference Very Large Data Bases (VLDB'94)*. Santiago de Chile, Chile. 12-15 de Setembro de 1994. p. 487–499. Morgan Kaufmann. 47
- (Aho et al., 1988) Alfred V. Aho, Brian W. Kernighan e Peter J. Weinberger. *The AWK Programming Language*. Addison-Wesley. 1988. 5
- (Aires e Aluísio, 2001) Rachel Aires e Sandra Aluísio. Criação de um Corpus com 1.000.000 de Palavras Etiquetado Morfossintaticamente. Relatório Técnico NILC-TR-01-8. Núcleo Interinstitucional de Linguística Computacional. 2001. 66
- (Aires, 2005) Rachel Virgínia Xavier Aires. *Uso de marcadores estilísticos para a busca na Web em português*. Tese de doutoramento. Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo. Agosto de 2005. <http://www.linguateca.pt/documentos/TeseDoutRachelAires.pdf>. 76, 88
- (Allan et al., 2003) James Allan, Jay Aslam, Nicholas Belkin, Chris Buckley, Jamie Callan, Bruce Croft, Sue Dumais, Norbert Fuhr, Donna Harman, David J. Harper, Djoerd Hiemstra, Thomas Hofmann, Eduard Hovy, John Lafferty Wessel Kraaij, Victor Lavrenko,

- David Lewis, Liz Liddy, R. Manmatha, Andrew McCallum, Jay Ponte, John Prager, Dragomir Radev, Philip Resnik, Stephen Robertson, Roni Rosenfeld, Salim Roukos, Mark Sanderson, Rich Schwartz, Amit Singhal, Alan Smeaton, Howard Turtle, Ellen Voorhees, Ralph Weischedel, Jinxi Xu e Cheng Xiang Zhai. Challenges in Information Retrieval and Language Modeling: Report of a Workshop held at the Center for Intelligent Information Retrieval, University of Massachusetts Amherst, September 2002. *SIGIR Forum*. 2: 31–47. 2003. 72
- (Almeida e Simões, 2001) José João de Almeida e Alberto Simões. Text to speech – "A rewriting system approach". *Procesamiento del Lenguaje Natural*. 27:247–255. Setembro de 2001. <http://alfarrabio.di.uminho.pt/~albie/publications/text2speech.pdf>. 100, 104
- (Aluísio et al., 2003) Sandra M. Aluísio, Gisele M. Pinheiro, Marcelo Finger, Maria das Graças Volpe Nunes e Stella E.O. Tagnin. The Lacio-Web Project: overview and issues in Brazilian Portuguese corpora creation. Em Dawn Archer, Paul Rayson, Andrew Wilson e Tony McEnery, editores, *Proceedings of the 2nd International Conference on Corpus Linguistics Corpus Linguistics (CL2003)*. Lancaster, Reino Unido. 28-31 de Março de 2003. p. 14–21. <http://www.nilc.icmc.usp.br/lacioweb/downloads/CL2003Lacio.zip>. 88
- (Arasu et al., 2001) Arvind Arasu, Junghoo Cho, Hector Garcia-molina, Andreas Paepcke e Sriram Raghavan. Searching the Web. *ACM Transactions on Internet Technology*. 1:2–43. 2001. 76
- (Armstrong et al., 2006) Stephen Armstrong, Marian Flanagan, Yvette Graham, Declan Groves, Bart Mellebeek, Sara Morrissey, Nicolas Stroppa e Andy May. MaTrEx: machine translation using examples. Em *TC-STAR OpenLab Workshop on Speech Translation*. Trento, Itália. 31 de Março-1 de Abril de 2006. 9
- (Auer et al., 2007) Sören Auer, Christian Bizer, Georgi Kobilarov, Jens Lehmann, Richard Cyganiak e Zachary Ives. DBpedia: A Nucleus for a Web of Open Data. Em Karl Aberer, Key-Sun Choi, Natasha Noy, Dean Allemang, Kyung-Il Lee, Lyndon Nixon, Jennifer Golbeck, Diana Maynard Peter Mika, Riichiro Mizoguchi, Guus Schreiber e Philippe Cudré-Mauroux, editores, *6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007, Busan, Korea, November 11-15, 2007, Proceedings*. 2007. p. 722–735. Springer. 84
- (Banerjee e Pedersen, 2003) Satanjeev Banerjee e Ted Pedersen. The Design, Implementation, and Use of the Ngram Statistics Package. Em Alexander Gelbukh, editor, *Proceedings of the 4th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing 2003)*. Cidade do México, México. 16-22 de Fevereiro de 2003. p. 370–381. Springer. 66, 67

- (Baptista, 2001) Jorge Baptista. *Sintaxe dos Predicados Nominais construídos com o Verbo-suporte SER DE*. Tese de doutoramento. Universidade do Algarve. 2001. 15
- (Barker e Szabpakowicz, 1998) Ken Barker e Stan Szabpakowicz. Semi-Automatic Recognition of Noun Modifier Relationships. Em *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics (COLING-ACL'98)*. Montreal, Canadá. 10-14 de Agosto de 1998. p. 96-102. ACL / Morgan Kaufmann. 36
- (Barreiro, 2008) Anabela Barreiro. ParaMT: a Paraphraser for Machine Translation. Em António Teixeira, Vera Lúcia Strube de Lima, Luís Caldas de Oliveira e Paulo Quaresma, editores, *Proceedings of the 8th International Conference on Computational Processing of the Portuguese Language (PROPOR'2008)*. Aveiro, Portugal. 8-10 de Setembro de 2008. p. 202-211. Springer. 17
- (Barreiro, 2007) Anabela Barreiro. Port4NooJ: Portuguese Linguistic Module and Bilingual Resources for Machine Translation. Em *Proceedings of the 2007 International NooJ Conference*. Barcelona, Espanha. 7-9 de Junho de 2007. Cambridge Scholars Publishing. 15
- (Barreiro e Afonso, 2007) Anabela Barreiro e Susana Afonso. Construção da lista dourada para as primeiras Morfolimpíadas do português. Em Diana Santos, editora, *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*. p. 107-118. IST Press. 20 de Março de 2007. 14
- (Belkin, 2008) Nicholas J. Belkin. Some (what) Grand Challenges for Information Retrieval. Em Craig MacDonald, Iadh Ounis, Vassilis Plachouras, Ian Ruthven e Ryan W. White, editores, *Advances in Information Retrieval, 30th European Conference on IR Research, ECIR 2008, Glasgow, UK, March 30-April 3, 2008. Proceedings*. 2008. p. 1. Springer. 72
- (Bernárdez, 2005) Enrique Bernárdez. Social cognition: variation, language, and culture in a cognitive linguistic typology. Em Francisco J. Ruiz de Mendoza e Sandra Peña Cervel, editores, *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction*. p. 191-222. Mouton de Gruyter. 2005. 27
- (Biber, 1988) Douglas Biber. *Variation across speech and writing*. Cambridge University Press. Cambridge, Reino Unido. 1988. 90
- (Biber et al., 2000) Douglas Biber, Susan Conrad e Randi Reppen. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge University Press. Cambridge, Reino Unido. 2000. 95
- (Bick, 2000) Eckhard Bick. *The Parsing System "Palavras": Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Tese de doutoramento. Universidade de Aarhus, Dinamarca. Aarhus University Press. Novembro de 2000. 28, 32, 92, 112

- (Bick et al., 2007) Eckhard Bick, Diana Santos, Susana Afonso e Rachel Marchi. Floresta Sintá(c)tica: Ficção ou realidade? Em Diana Santos, editora, *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*. p. 291–300. IST Press. 20 de Março de 2007. 2
- (Boldi e Vigna, 2005) Paolo Boldi e Sebastiano Vigna. MG4J at TREC 2005. Em *Proceedings of the The 14th Text REtrieval Conference (TREC 2005)*. 2005. p. 4. 81
- (Borba, 2002) F. S. Borba, editor. *Dicionário de usos do português contemporâneo do Brasil*. Ática. 2002. 112
- (Borbinha et al.,) José Luís Borbinha, Gilberto Pedrosa, Diogo Reis, João Luzio, Bruno Martins, João Gil e Nuno Freire. DIGMAP - Discovering Our Past World with Digitised Maps. Em László Kovács, Norbert Fuhr e Carlo Meghini, editores, *Research and Advanced Technology for Digital Libraries, 11th European Conference, ECDL 2007, Budapest, Hungary September 16-21, 2007, Proceedings*. LNCS. Springer. 53
- (Braga e Resende Jr, 2007) Daniela Braga e Fernando Gil Vianna Resende Jr. Módulos de Processamento de Texto Baseados em Regras para Sistemas de Conversão Texto-Fala em Português Europeu. Em *XXII Encontro Nacional da Associação Portuguesa de Linguística*. 2007. p. 141–155. APL. 100
- (Braga et al., 2008) Daniela Braga, Pedro Silva, Manuel Ribeiro, Mário Henriques e Miguel Sales Dias. HMM-based Brazilian Portuguese TTS. Em António Teixeira, Vera Lúcia Strube de Lima, Luís Caldas de Oliveira e Paulo Quaresma, editores, *Proceedings of the 8th International Conference on Computational Processing of the Portuguese Language (PROPOR'2008)*. Aveiro, Portugal. 8-10 de Setembro de 2008. Springer. 33
- (Broder, 2002) Andrei Broder. A Taxonomy of Web Search. *SIGIR Forum*. 36(2):3–10. 2002. 76
- (Busa e Johnston, 1996) Federica Busa e Michael Johnston. Qualia Structure and the Compositional Interpretation of Compounds. Em Evelyne Viegas, editora, *Proceedings of the ACL SIGLEX Workshop on Breath and Depth of Semantic Lexicons*. 1996. p. 167–187. Kluwer. 36
- (Cabral et al., 2008) Luís Miguel Cabral, Diana Santos e Luís Fernando Costa. SUPeRB: Building bibliographic resources on the computational processing of Portuguese. Em *PROPOR 2008 Special Session: Applications of Portuguese Speech and Language Technologies*. Aveiro, Portugal. 10 de Setembro de 2008. <http://www.linguateca.pt/documentos/CabralSantosCostaMLDC08.pdf>. 30

- (Caminada, 2008) Nuno Caminada. *Identificação Automática de Expressões Cristalizadas Preposicionais em Corpora da Língua Portuguesa*. Tese de doutoramento. Instituto Militar de Engenharia, Rio de Janeiro, Brasil. 2008. 69
- (Candeias, 2007) Sara Candeias. *Vocalismo dos "clíticos", Sistema fonológico da Beira Interior e algumas considerações sintáctico-semântica*. Tese de doutoramento. Departamento de Línguas e Cultura da Universidade de Aveiro. 2007. 102
- (Cardoso, 2008) Nuno Cardoso. REMBRANDT - Reconhecimento de entidades mencionadas Baseado em relações e análise detalhada do texto. Em Cristina Mota e Diana Santos, editoras, *Desafios na avaliação conjunta do reconhecimento de entidades mencionadas: O Segundo HAREM*. Linguateca. 2008. 80
- (Cardoso e Silva, 2007) Nuno Cardoso e Mário J. Silva. Query Expansion through Geographical Feature Types. Em *4th Workshop on Geographic Information Retrieval (GIR 2007)*. Lisboa, Portugal. 9 de Novembro de 2007. ACM press. <http://www.geo.unizh.ch/~rsp/gir07/>. 80, 81
- (Cardoso et al., 2006a) Nuno Cardoso, Leonardo Andrade, Alberto Simões e Mário J. Silva. The XLDB Group participation at CLEF 2005 ad hoc task. Em Carol Peters, Frederic Gey, Julio Gonzalo, Henning Müeller, Gareth J.F. Jones, Michael Kluck, Bernardo Magnini e Maarten de Rijke, editores, *Accessing Multilingual information Repositories: 6th Workshop of the Cross-Language Evaluation Forum, CLEF 2005. Vienna, Austria, September 2005. Revised Selected papers*. p. 54–60. Volume 4022 de LNCS. Springer. 2006. 11
- (Cardoso et al., 2006b) Nuno Cardoso, Bruno Martins, Leonardo Andrade, Marcirio Silveira Chaves e Mário J. Silva. The XLDB Group at GeoCLEF 2005. Em Carol Peters, Frederic Gey, Julio Gonzalo, Henning Müeller, Gareth J.F. Jones, Michael Kluck, Bernardo Magnini e Maarten de Rijke, editores, *Accessing Multilingual information Repositories: 6th Workshop of the Cross-Language Evaluation Forum, CLEF 2005. Vienna, Austria, September 2005. Revised Selected papers*. p. 997–1006. Volume 4022 de LNCS. Springer. 2006. 52
- (Cardoso et al., 2007) Nuno Cardoso, Bruno Martins, Daniel Gomes e Mário J. Silva. WPT 03: a primeira coleção pública proveniente de uma recolha da web portuguesa. Em Diana Santos, editora, *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*. p. 279–288. IST Press. 20 de Março de 2007. 56, 76
- (Cardoso et al., 2008a) Nuno Cardoso, David Cruz, Marcirio Chaves e Mário J. Silva. Using Geographic Signatures as Query and Document Scopes in Geographic IR. Em Carol Peters, Valentin Jijkoun, Thomas Mandl, Henning Müller, Doug W. Oard, Anselmo Peñas, Vivien Petras e Diana Santos, editores, *Advances in Multilingual and Multimodal Information Retrieval: 8th Workshop of the Cross-Language Evaluation Forum, CLEF 2007, Budapest*,

- Hungary, September 19-21, 2007, Revised Selected Papers*. p. 802–810. Volume 5152 de LNCS. Springer. 2008. 52, 85
- (Cardoso et al., 2008b) Nuno Cardoso, Mário J. Silva e Diana Santos. Handling Implicit Locality Evidence for Geographic IR. Em *Proceedings of the 17th Conference on Information and Knowledge Management (CIKM 2008)*. Napa Valley, CA, EUA. 26-30 de Outubro de 2008. ACM. 83
- (Cardoso et al., 2008c) Nuno Cardoso, Patrícia Sousa e Mário J. Silva. The University of Lisbon at GeoCLEF 2008. Em *Working Notes for the CLEF 2008 Workshop*. Aarhus, Dinamarca. 17-19 de Setembro de 2008. 85
- (Chacoto, 2005) Lucília Chacoto. *O Verbo Fazer em Construções Nominiais Predicativas*. Tese de doutoramento. Universidade do Algarve. 2005. 15
- (Chaves et al., 2005a) Marcirio Chaves, Bruno Martins e Mário J. Silva. GKB - Geographic Knowledge Base. Relatório Técnico DI/FCUL TR-05-12. Departamento de Informática, Faculdade de Ciências da Universidade de Lisboa. Julho de 2005. http://www.linguateca.pt/documentos/gkb_technical_report.pdf. 50
- (Chaves e Santos, 2006) Marcirio Silveira Chaves e Diana Santos. What Kinds of Geographical Information Are There in the Portuguese Web? Em Renata Vieira, Paulo Quaresma, Maria da Graça Volpes Nunes, Nuno J. Mamede, Cláudia Oliveira e Maria Carmelita Dias, editores, *Computational Processing of the Portuguese Language: 7th International Workshop, PROPOR 2006*. Itatiaia, Brazil, May 13-17, 2006, *Proceedings*. p. 264–267. Volume 3960. Springer. 2006. http://www.linguateca.pt/documentos/Poster_ChavesSantosPROPOR2006.pdf. 56
- (Chaves et al., 2005b) Marcirio Silveira Chaves, Mário J. Silva e Bruno Martins. A Geographic Knowledge Base for Semantic Web Applications. Em C. A. Heuser, editor, *20º Simpósio Brasileiro de Banco de Dados (SBBD 2005)*. Uberlândia, MG, Brasil. 3-7 de Outubro de 2005. p. 40–54. <http://www.linguateca.pt/documentos/ChavesetalSBBD2005.pdf>. 50, 76
- (Chaves, 2008) Marcirio Chaves. Geo-ontologias para reconhecimento de relações entre locais: a participação do SEI-Geo no Segundo HAREM. Em Cristina Mota e Diana Santos, editoras, *Desafios na avaliação conjunta do reconhecimento de entidades mencionadas: O Segundo HAREM*. Linguateca. 2008. 57
- (Copestake, 2003) Ann Copestake. Compounds revisited. Em *2nd International Workshop on Generative Approaches to the Lexicon (GL'2003)*. Geneva, Suíça. 15-17 de Maio de 2003. 36

- (Correia, 2006) Ana Teresa Varajão Moutinho Pereira Correia. Colaboração na constituição do corpus paralelo Le Monde Diplomatique (FR-PT). Relatório de estágio, Universidade do Minho. Dezembro de 2006. 8
- (Coulthard, 1977) Malcolm Coulthard. *An Introduction to Discourse Analysis*. Longman. Londres, Reino Unido. 1977. 94
- (Coulthard e Johnson, 2007) Malcolm Coulthard e Alison Johnson. *An Introduction to Forensic Linguistics: Language in Evidence*. Routledge. Londres, Reino Unido e Nova Iorque, EUA. 2007. 94, 97
- (Cunha et al., 2006) João Paulo Silva Cunha, Isabel Cruz, Ilídio Oliveira, António Sousa Pereira, César Telmo Costa, Ana Margarida Oliveira e Amândio Pereira. The RTS project: Promoting secure and effective clinical telematic communication within the Aveiro region. Em *Proceedings of the eHealth 2006 High Level Conference and Exhibition*. Málaga, Espanha. 10-12 de Maio de 2006. p. 1–10. 44
- (Cunningham et al., 2002a) Hamish Cunningham, Diana Maynard, Kalina Bontcheva e Valentin Tablan. GATE: A Framework and Graphical Development Environment for Robust NLP Tools and Applications. Em *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL'02)*. Filadélfia, PA, EUA. 6-12 de Julho de 2002. p. 168–175. Association for Computational Linguistics. 45
- (Cunningham et al., 2002b) Hamish Cunningham, Diana Maynard, Kalina Bontcheva, Valentin Tablan e Cristian Ursu. The GATE User Guide. <http://gate.ac.uk/>. 2002. 45
- (DGS, 2006) Direcção geral da saúde. Programa Nacional de Vacinação. 2006. <http://www.min-saude.pt/portal/conteudos/informacoes+uteis/vacinacao/vacinas.htm>. 45
- (Dijk, 1997) Teun A. van Dijk. Discourse as Interaction in Society. *Discourse Studies: A Multidisciplinary Introduction - Discourse as Social Interaction*. 2:1–37. 1997. 94
- (Duran e Xatara, 2006) Sanches Magali Duran e Claudia Maria Xatara. A Metalexigrafia Pedagógica. *Cadernos de Tradução*. 2:41–66. 2006. http://www.cadernos.ufsc.br/online/cadernos18/magali_xatara.pdf. 110, 112
- (Efthimiadis, 1993) Efthimis N. Efthimiadis. A User-centered Evaluation of Ranking Algorithms for Interactive Query Expansion. Em Robert Korfhage e Edie M. Rasmussen e Peter Willett, editores, *Proceedings of the 16th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval (SIGIR'93)*. Pittsburgh, PA, EUA. 27 de Junho a 1 de Julho de 1993. p. 146–159. 81
- (Efthimiadis, 1996) Efthimis N. Efthimiadis. Query expansion. *Annual Review of Information Systems and Technology*. 31:121–187. 1996. 73

- (Fairclough e Wodak, 1997) Norman Fairclough e Ruth Wodak. Critical Discourse Analysis. *Discourse Studies: A Multidisciplinary Introduction - Discourse as Social Interaction*. 2: 258–284. 1997. 94
- (Ferreira e Teixeira, 2008) Liliana Ferreira e António Teixeira. REMMA - Reconhecimento de entidades mencionadas do MedAlert. Em Cristina Mota e Diana Santos, editoras, *Desafios na avaliação conjunta do reconhecimento de entidades mencionadas: O Segundo HAREM*. Linguateca. 2008. <http://linguateca.dei.uc.pt/harem/encontro/remma.pdf>. 47
- (Ferreira et al., 2008) Liliana Ferreira, António Teixeira e João Paulo da Silva Neto. Ontology-driven Vaccination Information Extraction. Em *Proceedings of the 5th International Workshop on Natural Language Processing and Cognitive Science (NLPCS 2008)*. Barcelona, Espanha. 12-13 de Junho de 2008. p. 94–103. 44
- (Ferrucci e Lally, 2004) David Ferrucci e Adam Lally. UIMA: an architectural approach to unstructured information processing in the corporate research environment. *Natural Language Engineering*. 10(3-4):327–348. 2004. 45
- (Frankenberg-Garcia e Santos, 2002) Ana Frankenberg-Garcia e Diana Santos. COMPARA, um corpus paralelo de português e de inglês na Web. *Cadernos de Tradução*. IX(1):61–79. 2002. <http://www.linguateca.pt/documentos/Frankenberg-GarciaSantosCadTrad.pdf>. 36
- (Frankenberg-Garcia e Santos, 2003) Ana Frankenberg-Garcia e Diana Santos. Introducing COMPARA, the Portuguese-English parallel translation corpus. Em Federico Zanettin, Silvia Bernardini e Dominic Stewart, editores, *Corpora in Translation Education*. p. 71–87. St. Jerome Publishing. 2003. <http://www.linguateca.pt/documentos/Frankenberg-GarciaSantos2000.pdf>. 8, 15, 31
- (Geeraerts, 2005) Dirk Geeraerts. Lectal variation and empirical data in Cognitive Linguistics. Em Francisco J. Ruiz de Mendoza e Sandra Peña Cervel, editores, *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction*. p. 163–189. Mouton de Gruyter. 2005. 27
- (Geeraerts, 2006) Dirk Geeraerts. Methodology in Cognitive Linguistics. Em Gitte Kristiansen, Michel Achard, René Dirven e Francisco J. Ruiz de Mendoza, editores, *Cognitive Linguistics: Current Applications and Future Perspectives*. p. 21–49. Mouton de Gruyter. 2006. 27
- (Geeraerts et al., 1999) Dirk Geeraerts, Stefan Grondelaers e Dirk Speelman. *Convergentie en Divergentie in de Nederlandse Woordenschat. Een onderzoek naar kleding- en voetbaltermen*. Meertens Instituut. 1999. 27

- (Geiger, 2004) Paulo Geiger. *Minidicionário Caldas Aulete*. Editora Nova Fronteira. 2004. 110
- (Girju et al., 2005) Roxana Girju, Dan Moldovan, Marta Tatu e Daniel Antohe. On the semantics of noun compounds. *Computer Speech and Language*. 19:479–496. Março de 2005. 36, 38
- (Gonzalez-Marquez et al., 2007) Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson e Michael J. Spivey. *Methods in Cognitive Linguistics*. John Benjamins. 2007. 27
- (Gouveia et al., 2000) Paulo D.F. Gouveia, João P.R. Teixeira e Diamantino R.S. Freitas. Divisão Silábica Automática do Português Escrito e Falado. Em Maria das Graças Volpe Nunes, editora, *V Encontro para o processamento computacional da língua portuguesa escrita e falada (PROPOR 2000)*. São Paulo, Brasil. 19-22 de Novembro de 2000. p. 65–74. ICMC/USP. 100
- (Green, 1979) Thomas R. G. Green. The Necessity of Syntax Markers: Two Experiments With Artificial Languages. *Journal of Verbal Learning and Behaviour*. 18:481–496. 1979. 9
- (Gross, 1975) Maurice Gross. *Méthodes en Syntaxe - Régime des constructions complétives*. Hermann. 1975. 15
- (Gross, 1981) Maurice Gross. Les bases empiriques de la notion de prédicat sémantique. *Formes Syntaxiques et Prédicat Sémantiques, Langages*. 63:7–52. 1981. 15
- (Halliday, 1994) M. A. K. Halliday. *An Introduction to Functional Grammar*. Edward Arnold. Londres, Reino Unido. 1994. 95
- (Harris, 1957) Zellig Harris. Co-occurrence and transformation in linguistic structure. *Language*. 33:293–340. 1957. 15
- (Harris, 1968) Zellig Harris. *Mathematical Structures of Language*. Wiley. 1968. 15
- (Hänlein, 1998) Heike Hänlein. *Studies in Authorship Recognition - A Corpus-based Approach*. Peter Lang. Francoforte, Alemanha. 1998. 94, 95
- (Houaiss, 2001) Antonio Houaiss, editor. *Dicionário Houaiss da Língua Portuguesa*. Editora Objectiva. 2001. 112
- (Karlgrén, 2000) Jussi Karlgrén. *Stylistic Experiments for Information Retrieval*. Tese de doutoramento. Universidade de Estocolmo. 2000. 90
- (Koehn, 2005) P. Koehn. Europarl: A parallel corpus for statistical machine translation. Em *Proceedings of the Tenth Machine Translation Summit (MT-Summit X)*. Ilha de Phuket, Tailândia. 12-16 de Setembro de 2005. p. 79–86. 8

- (Kohler, 2003) Janet Kohler. *Analysing Search Engine Queries for the Use of Geographic Terms*. Tese de mestrado. Universidade de Sheffield, Reino Unido. 2003. 74
- (Kristiansen e Dirven, 2008) Gitte Kristiansen e René Dirven. *Cognitive Sociolinguistics: Language Variation, Cultural Models, Social Systems*. Mouton de Gruyter. 2008. 27
- (Lassila e Swick, 1998) Ora Lassila e Ralph Swick. *Resource Description Framework (RDF) Model and Syntax*. W3C, World Wide Web Consortium. 1998. <http://www.w3.org/TR/WD-rdf-syntax/>. 47
- (Maia e Barreiro, 2007) Belinda Maia e Anabela Barreiro. Uma experiência de recolha de exemplos classificados de tradução automática de inglês para português. Em Diana Santos, editora, *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*. p. 205–216. IST Press. 20 de Março de 2007. 14
- (Maia e Matos, 2008) Belinda Maia e Sérgio Matos. *Corpógrafo V4 - Tools for Researchers and Teachers using Comparable Corpora*. Em Pierre Zweigenbaum, Éric Gaussier e Pascale Fung, editores, *LREC 2008 Workshop on Comparable Corpora*. Marraquexe, Marrocos. 31 de Maio de 2008. p. 79–82. European Language Resources Association (ELRA). <http://www.linguateca.pt/documentos/MaiaMatosW12LREC08.pdf>. 15, 95
- (Mandl et al., 2008) Thomas Mandl, Fredric Gey, Giorgio Di Nunzio, Nicola Ferro, Ray Larson, Mark Sanderson, Diana Santos, Christa Womser-Hacker e Xing Xie. *GeoCLEF 2007: The CLEF 2007 Cross-Language Geographic Information Retrieval Track Overview*. Em Carol Peters, Valentin Jijkoun, Thomas Mandl, Henning Müller, Doug W. Oard, Anselmo Peñas, Vivien Petras e Diana Santos, editores, *Advances in Multilingual and Multimodal Information Retrieval: 8th Workshop of the Cross-Language Evaluation Forum, CLEF 2007, Budapest, Hungary, September 19-21, 2007, Revised Selected Papers*. p. 745–772. Volume 5152 de LNCS. Springer. 2008. 85
- (Marques et al., 2007) Nuno C. Marques, Sebastian Bader, Vitor Rocio e Steffen Hölldobler. *Neuro-Symbolic Word Tagging*. Em *Proceedings of 13th Portuguese Conference on Artificial Intelligence (EPIA'07)*. Guimarães, Portugal. 3-7 de Dezembro de 2007. IEEE. 5
- (Martins et al., 2007a) Bruno Martins, Nuno Cardoso, Marcirio Silveira Chaves, Leonardo Andrade e Mário J. Silva. *The University of Lisbon at GeoCLEF 2006*. Em Carol Peters, Paul Clough, Fredric C. Gey, Jussi Karlgren, Bernardo Magnini, Douglas W. Oard, Maarten de Rijke e Maximilian Stempfhuber, editores, *Evaluation of Multilingual and Multi-modal Information Retrieval - 7th Workshop of the Cross-Language Evaluation Forum, CLEF 2006, Alicante, Spain, September, 2006. Revised Selected papers*. p. 986–994. Volume 4730 de LNCS. Springer. 2007. 58

- (Martins et al., 2007b) Bruno Martins, Mário Silva e Marcirio Chaves. O sistema CaGE no HAREM - reconhecimento de entidades geográficas em textos em língua portuguesa. Em Diana Santos e Nuno Cardoso, editores, *Reconhecimento de entidades mencionadas em português: Documentação e actas do HAREM, a primeira avaliação conjunta na área*. p. 97–112. 12 de Novembro de 2007. http://www.linguateca.pt/aval_conjunta/LivroHAREM/Cap08-SantosCardoso2007-Martinsetal.pdf. 52
- (Martins Junior e Moreira, 2004) José Martins Junior e Edson dos Santos Moreira. Using Support Vector Machines to Recognize Products in E-commerce Pages. Em *Proceedings of the 22nd IASTED International Multi-Conference on Applied Informatics*. Fevereiro de 2004. p. 212–217. 90
- (McDonald, 1996) David D. McDonald. Internal and external evidence in the identification and semantic categorization of proper names. Em I. Boguraev e J. Pustejovsky, editores, *Corpus processing for lexical acquisition*. p. 21–39. 1996. 83
- (McEnery e Wilson, 1996) Tony McEnery e Andrew Wilson. *Corpus Linguistics: An Introduction*. Edinburgh University Press. Edimburgo, Reino Unido. 1996. 94, 95
- (McMenamin, 2002) Gerald R. McMenamin. *Forensic Linguistics: Advances in Forensic Stylistics*. CRC Press. Boca Raton e Nova Iorque, EUA. 2002. 94
- (Melamed, 2001) I. Dan Melamed. *Empirical methods for exploiting parallel texts*. MIT Press. 2001. 10
- (Meyers et al., 2004a) Adam Meyers, Ruth Reeves, Catherine Macleod, Rachel Szekeley, Veronique Zielinska e Brian Young. The Cross-Breeding of Dictionaries. Em Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa e Raquel Silva, editoras, *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC'2004)*. Lisboa, Portugal. 26-28 de Maio de 2004. 15
- (Meyers et al., 2004b) Adam Meyers, Ruth Reeves, Catherine Macleod, Rachel Szekely, Veronika Zielinska, Brian Young e Ralph Grishman. Annotating noun argument structure for NomBank. Em Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa e Raquel Silva, editoras, *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC'2004)*. Lisboa, Portugal. 26-28 de Maio de 2004. 15
- (Mika, 2004) Peter Mika. Social Networks and the Semantic Web. Em *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI'04)*. Pequim, China. 20-24 de Setembro de 2004. p. 285–291. 74
- (Mika, 2006) Peter Mika. Ontologies Are Us: A Unified Model of Social Networks and Semantics. Em V. Richard Benjamins e Mark A. Musen Yolanda Gil, editor, *The Semantic*

- Web : ISWC 2005, 4th International Semantic Web Conference, ISWC 2005, Galway, Ireland, November 6-10, 2005, Proceedings (ISWC'2005)*. 2006. p. 522–536. Springer. 74
- (Mota e Santos, 2008) Cristina Mota e Diana Santos, editoras. *Desafios na avaliação conjunta do reconhecimento de entidades mencionadas: Actas do Encontro do Segundo HAREM*. Linguatca. 2008. i, 47, 57
- (Nogueira, 2004) Cícero Nogueira. Algoritmo para extração de Combinações do tipo V+(det)+N. Programa feito em Java. 2004. 67
- (Och e Ney, 2004) Franz Josef Och e Hermann Ney. The alignment template approach to statistical machine translation. *Computational Linguistics*. 30:417–449. 2004. 10
- (Oksefjell e Santos, 1998) Signe Oksefjell e Diana Santos. Breve panorâmica dos recursos de português mencionados na Web. Em Vera Lúcia Strube de Lima, editora, *III Encontro para o Processamento Computacional do Português Escrito e Falado (PROPOR'98)*. Porto Alegre, RS, Brasil. 3-4 de Novembro de 1998. p. 38–47. <http://www.linguatca.pt/documentos/recursos.pdf>. 44
- (Oliveira et al., 2008) Hugo Gonçalo Oliveira, Paulo Gomes e Diana Santos. PAPEL: a dictionary-based lexical ontology for Portuguese. Em António Teixeira, Vera Lúcia Strube de Lima, Luís Caldas de Oliveira e Paulo Quaresma, editores, *Proceedings of the 8th International Conference on Computational Processing of the Portuguese Language (PROPOR'2008)*. Aveiro, Portugal. 8-10 de Setembro de 2008. p. 31–40. Springer. 47
- (Olsson, 2004) John Olsson. *Forensic Linguistics: An Introduction to Language, Crime and the Law*. Continuum. Londres, Reino Unido. 2004. 94
- (Pereira e Shieber, 1987) Fernando C. N. Pereira e Stuart M. Shieber. Prolog and Natural Language Analysis - Digital Edition. Microtome Publishing. 1987. <http://www.mtome.com/Publications/PNLA/pnla.html>. 2
- (Pinheiro e Aluísio, 2003) Gisele Montilha Pinheiro e Sandra Maria Aluísio. Córpus NILC: descrição e análise crítica com vistas ao projeto Lacio-Web. Relatório Técnico NILC–TR–03–03. Núcleo Interinstitucional de Linguística Computacional. Fevereiro de 2003. 66
- (Pustejovsky, 1995) James Pustejovsky. *The Generative Lexicon*. MIT Press. 1995. 36, 38
- (Ranchhod, 1990) Elisabete Ranchhod. *Sintaxe dos predicados nominais com Estar*. INIC. 1990. 15
- (Ranchhod, 2003) Elisabete Ranchhod. O Lugar das Expressões Fixas na Gramática do Português. 2003. 66

- (Rocchio Jr, 1971) J. J. Rocchio Jr. Relevance Feedback in Information Retrieval. Em Gerald Salton, editor, *The SMART Retrieval System: Experiments in Automatic Document Processing*. 1971. p. 313–323. 81
- (Rocha e Santos, 2000) Paulo Alexandre Rocha e Diana Santos. CETEMPúblico: Um corpus de grandes dimensões de linguagem jornalística portuguesa. Em Maria das Graças Volpe Nunes, editora, *V Encontro para o processamento computacional da língua portuguesa escrita e falada (PROPOR 2000)*. São Paulo, Brasil. 19-22 de Novembro de 2000. p. 131–140. ICMC/USP. <http://www.linguateca.pt/documentos/RochaSantosPROPOR2000.pdf>. 2, 31, 110
- (Rodrigues e Quaresma, 1999) Irene Rodrigues e Paulo Quaresma, editores. Évora, Portugal. 20-21 de Setembro de 1999. 62
- (Rundell, 1999) Michael Rundell. Recent trends in publishing monolingual learners' dictionaries. Em R. R. K. Hartmann, editor, *Thematic Network Projects, Sub-project 9: Dictionaries - Dictionaries in Language Learning, Final Report Year Three*. p. 83–98. 1999. 110
- (Ryder, 1994) Mary Ellen Ryder. *Ordered Chaos: The Interpretation of English Noun-Noun Compounds*. University of California Press. 1994. 37
- (SAMPA,) Speech Assessment Methods Phonetic Alphabet. <http://www.phon.ucl.ac.uk/home/sampa/portug.htm>. 102
- (Sampson, 1995) Geoffrey Sampson. *English for the Computer: The SUSANNE Corpus and Analytic Scheme*. Clarendon Press. 1995. 2
- (Santos, 2007) Diana Santos, editora. *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*. IST Press. 20 de Março de 2007. <http://www.istpress.ist.utl.pt/lavaliacaoconjunta.html>. 47
- (Santos e Barreiro, 2004) Diana Santos e Anabela Barreiro. On the problems of creating a consensual golden standard of inflected forms in. Em Maria Teresa Lino, Maria Francisca Xavier, Fátima Ferreira, Rute Costa e Raquel Silva, editoras, *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC'2004)*. Lisboa, Portugal. 26-28 de Maio de 2004. p. 483–486. <http://www.linguateca.pt/Diana/download/SantosBarreiroLREC2004.pdf>. 14
- (Santos e Cardoso, 2007) Diana Santos e Nuno Cardoso, editores. *Reconhecimento de entidades mencionadas em português: Documentação e actas do HAREM, a primeira avaliação conjunta na área*. Linguateca. 12 de Novembro de 2007. http://www.linguateca.pt/aval_conjunta/LivroHAREM/Livro-SantosCardoso2007.pdf. 30

- (Santos e Cardoso, 2008) Diana Santos e Nuno Cardoso. GikiP: Evaluating geographical answers from Wikipedia. Em *5th Workshop on Geographic Information Retrieval (GIR'2008)*. Napa Valley, CA, EUA. 30 de Outubro de 2008. ACM. 85
- (Santos e Chaves, 2006) Diana Santos e Marcirio Chaves. The place of place in geographical IR. Em *3rd Workshop on Geographic Information Retrieval (GIR'2006)*. Seattle, WA, EUA. 10 de Agosto de 2006. p. 5–8. <http://www.linguateca.pt/Diana/download/acetSantosChavesGIR2006.pdf>. 56
- (Santos e Costa, 2003) Diana Santos e Luís Costa. Morfolimpiadas - Apresentação detalhada da metodologia e dos problemas identificados. Em *Encontro de Avaliação Conjunta de Sistemas de Processamento Computacional do Português (AvalON'2003)*. Universidade do Algarve, Faro. 28 de Junho de 2003. <http://www.linguateca.pt/documentos/SantosMorfolimpiadasAvalon2003.pdf>. 30
- (Santos e Inácio, 2006) Diana Santos e Susana Inácio. Annotating COMPARA, a grammar-aware parallel corpus. Em Nicoletta Calzolari, Khalid Choukri, Aldo Gangemi, Bente Maegaard, Joseph Mariani, Jan Odjik e Daniel Tapias, editores, *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'2006)*. Génova, Itália. 22-28 de Maio de 2006. p. 1216–1221. <http://www.linguateca.pt/Diana/download/SantosInacioLREC2006.pdf>. 15
- (Santos e Rocha, 2005) Diana Santos e Paulo Rocha. The key to the first CLEF in Portuguese: Topics, questions and answers in CHAVE. Em Carol Peters, Paul Clough, Julio Gonzalo, Gareth J. F. Jones, Michael Kluck e Bernardo Magnini, editores, *Multilingual Information Access for Text, Speech and Images, 5th Workshop of the Cross-Language Evaluation Forum, CLEF 2004, Bath, UK, September 15-17, 2004, Revised Selected Papers*. p. 821–832. Volume 3491 de LNCS. Springer. 2005. <http://www.linguateca.pt/documentos/SantosRochaCLEF2004Springer2005.pdf>. 58
- (Santos e Sarmiento, 2002) Diana Santos e Luís Sarmiento. O projecto AC/DC: acesso a corpora/disponibilização de corpora. Em Amália Mendes e Tiago Freitas, editores, *Actas do XVIII Encontro Nacional da Associação Portuguesa de Linguística (APL 2002)*. Porto, Portugal. 2-4 de Outubro de 2002. p. 705–717. APL. <http://www.linguateca.pt/documentos/SantosSarmientoAPL2002.pdf>. 22, 26, 62, 66, 110
- (Santos e Simões, 2008) Diana Santos e Alberto Simões. Portuguese-English word alignment: some experiments. Em *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*. Marraquexe, Marrocos. 28-30 de Maio de 2008. European Language Resources Association (ELRA). <http://www.linguateca.pt/documentos/SantosSimoiesLREC2008.pdf>. 8

- (Santos et al., 2004) Diana Santos, Belinda Maia e Luís Sarmento. Gathering empirical data to evaluate MT from English to Portuguese. Em Lambros Kranias, Nicoletta Calzolari, Gregor Thurmair, Yorick Wilks, Eduard Hovy, Guðrún Magnúsdóttir, Anna Samiotou e Khalid Choukri, editores, *Proceedings of LREC 2004 Workshop on the Amazing Utility of Parallel and Comparable Corpora*. Lisboa, Portugal. 25 de Maio de 2004. p. 14–17. <http://www.linguateca.pt/documentos/SantosMaiaSarmentoAmazing2004.pdf>. 14
- (Santos et al., 2006) Diana Santos, Nuno Seco, Nuno Cardoso e Rui Vilela. HAREM: An Advanced NER Evaluation Contest for Portuguese. Em Nicoletta Calzolari, Khalid Choukri, Aldo Gangemi, Bente Maegaard, Joseph Mariani, Jan Odjik e Daniel Tapias, editores, *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'2006)*. Génova, Itália. 22-28 de Maio de 2006. p. 1986–1991. <http://www.linguateca.pt/Diana/download/SantosSecoCardosoVilelaLREC2006.pdf>. 56, 82
- (Santos et al., 2008a) Diana Santos, Nuno Cardoso, Paula Carvalho, Iustin Dornescu, Sven Hartrumpf, Johannes Leveling e Yvonne Skalban. Getting geographical answers from Wikipedia: the GikiP pilot at CLEF. Em Francesca Borri, Alessandro Nardi e Carol Peters, editores, *Cross Language Evaluation Forum: Working Notes for the CLEF 2008 Workshop*. Aarhus, Dinamarca. 17-19 de Setembro de 2008. p. s/pp. <http://www.linguateca.pt/Diana/download/SantosetalWNCLEF2008.pdf>. 85
- (Santos et al., 2008b) Diana Santos, Cláudia Freitas, Hugo Gonçalo Oliveira e Paula Carvalho. Second HAREM: new challenges and old wisdom. Em António Teixeira, Vera Lúcia Strube de Lima, Luís Caldas de Oliveira e Paulo Quaresma, editores, *Proceedings of the 8th International Conference on Computational Processing of the Portuguese Language (PROPOR'2008)*. Aveiro, Portugal. 8-10 de Setembro de 2008. p. 212–215. Springer. 82
- (Sarmento, 2006) Luís Sarmento. SIEMÊS - A Named Entity Recognizer for Portuguese Relying on Similarity Rules. Em Renata Vieira, Paulo Quaresma, Maria da Graça Volpes Nunes, Nuno J. Mamede, Cláudia Oliveira e Maria Carmelita Dias, editores, *Computational Processing of the Portuguese Language: 7th International Workshop, PROPOR 2006. Itatiaia, Brazil, May 13-17, 2006, Proceedings*. p. 90–99. Volume 3960. Springer. 2006. <http://www.linguateca.pt/documentos/SarmentoSiemes2006.pdf>. 56
- (Sarmento, 2007) Luís Sarmento. Ferramentas para experimentação, recolha e avaliação de exemplos de tradução automática. Em Diana Santos, editora, *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*. p. 193–203. IST Press. 20 de Março de 2007. 14
- (Sarmento et al., 2004) Luís Sarmento, Belinda Maia e Diana Santos. The Corpógrafo - a Web-based environment for corpora research. Em Maria Teresa Lino, Maria Fran-

- cisca Xavier, Fátima Ferreira, Rute Costa e Raquel Silva, editoras, *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC'2004)*. Lisboa, Portugal. 26-28 de Maio de 2004. p. 449–452. <http://www.linguateca.pt/Diana/download/SarmientoMaiaSantosLREC2004.pdf>. 15, 95
- (Sarmiento et al., 2006) Luís Sarmiento, Ana Sofia Pinto e Luís Cabral. REPENTINO - A Wide-Scope Gazetteer for Entity Recognition in Portuguese. Em Renata Vieira, Paulo Quaresma, Maria da Graça Volpes Nunes, Nuno J. Mamede, Cláudia Oliveira e Maria Carmelita Dias, editores, *Computational Processing of the Portuguese Language: 7th International Workshop, PROPOR 2006. Itatiaia, Brazil, May 13-17, 2006, Proceedings*. p. 31–40. Volume 3960. Springer. 2006. <http://www.linguateca.pt/documentos/SarmientoPintoCabral2006PROPORSpringer.pdf>. 62
- (Sarmiento et al., 2007) Luís Sarmiento, Anabela Barreiro, Belinda Maia e Diana Santos. Avaliação de Tradução Automática: alguns conceitos e reflexões. Em Diana Santos, editora, *Avaliação conjunta: um novo paradigma no processamento computacional da língua portuguesa*. p. 181–190. IST Press. 20 de Março de 2007. 14
- (Seco e Cardoso, 2006) Nuno Seco e Nuno Cardoso. Detecting user sessions in the tumba! web log. Relatório técnico. Março de 2006. <http://eden.dei.uc.pt/~nseco/tumba.pdf>. 79
- (Shuy, 2006) Roger W. Shuy. *Linguistics in the Courtroom: A Practical Guide*. Oxford University Press. 2006. 94
- (Silberztein, 2004) Max Silberztein. NooJ: A Cooperative, Object-Oriented Architecture for NLP. *INTEX pour la Linguistique et le traitement automatique des langues*. 2004. Cahiers de la MSH Ledoux. 15
- (Silva, 2005) Augusto Soares da Silva. Convergência e divergência no léxico do Português Europeu e do Português Brasileiro: resultados do estudo sobre termos de futebol e de moda. Em Joaquim Barbosa e Fátima Oliveira, editores, *Textos seleccionados do XXI Encontro da Associação Portuguesa de Linguística*. 28–30 de Setembro de 2005. p. 633–646. Colibri. 26
- (Silva, 2006a) Augusto Soares da Silva. *O Mundo dos Sentidos em Português: Polissemia, Semântica e Cognição*. Almedina. 2006. 27
- (Silva, 2006b) Augusto Soares Silva. Sociolinguística cognitiva e o estudo da convergência/divergência entre o Português Europeu e o Português Brasileiro. *Veredas: Revista de Estudos Lingüísticos*. 10. 2006. <http://www.revistaveredas.ufjf.br>. 27

- (Silva, 2008) Augusto Soares da Silva. *Sociolinguística cognitiva, lexicologia quantitativa e variação do Português*. Tese de doutoramento. Universidade Católica de Braga. 18 de Junho de 2008. 27
- (Silva, 2003) Mário J. Silva. The Case for a Portuguese Web Search Engine. Em Pedro Isaías, editor, *Proceedings of the IADIS International Conference WWW/Internet 2003 (ICWI 2003)*. Faro, Portugal. 5-8 de Novembro de 2003. p. 411–418. IADIS. 77
- (Silva et al., 2006) Mário J. Silva, Bruno Martins, Marcirio Silveira Chaves, Ana Paula Afonso e Nuno Cardoso. Adding Geographic Scopes to Web Resources. *CEUS - Computers Environment and Urban Systems*. 30(4):378–399. 2006. 52
- (Simões, 2004) Alberto Simões. Parallel Corpora word alignment and applications. Tese de mestrado. Faculdade de Engenharia da Universidade do Minho. 2004. <http://alfarrabio.di.uminho.pt/~albie/publications/msc.pdf>. 8
- (Simões, 2007a) Alberto Simões. Makefile::Parallel - Uma ferramenta para paralelização de processos. Em *V Simpósio Doutoral da Linguatca 2007*. Faculdade de Ciências, Universidade de Lisboa. 4 de Outubro de 2007. <http://www.linguatca.pt/documentos/MakefileParallel-AlbertoSimoes.pdf>. 11
- (Simões, 2007b) Alberto Simões. Segmentação bilingue com base na marker hypothesis. Em César Analide, Paulo Novais e Pedro Henriques, editores, *Simpósio Doutoral em Inteligência Artificial (SDIA 2007)*. Guimarães, Portugal. 3–7 de Dezembro de 2007. p. 135–144. <http://ambs.perl-hackers.net/publications/sdia07.pdf>. 9
- (Simões, 2008) Alberto Simões. *Extracção de Recursos de Tradução com base em Dicionários Probabilísticos de Tradução*. Tese de doutoramento. Faculdade de Engenharia da Universidade do Minho. Março de 2008. <http://www.linguatca.pt/documentos/SimoesPHD.pdf>. 8, 10
- (Simões e Almeida, 2003) Alberto Simões e José João Almeida. NATools – A Statistical Word Aligner Workbench. *Procesamiento del Lenguaje Natural*. 31:217–226. Setembro de 2003. <http://alfarrabio.di.uminho.pt/~albie/publications/sepln2003.pdf>. 8
- (Simões e Almeida, 2006a) Alberto Simões e José João Almeida. Combinatory Examples Extraction for Machine Translation. Em *EAMT 11th Annual Conference*. Oslo, Noruega. 19-20 de Junho de 2006. p. 27–32. <http://alfarrabio.di.uminho.pt/~albie/publications/eamt06.pdf>. 10
- (Simões e Almeida, 2006b) Alberto Simões e José João Almeida. NatServer: A Client-Server Architecture for building Parallel Corpora applications. *Procesamiento del Lenguaje Natural*. 37:91–98. Setembro de 2006. <http://alfarrabio.di.uminho.pt/~albie/publications/sepln06.pdf>. 11

- (Simões e Almeida, 2007) Alberto Simões e José João Almeida. Parallel Corpora based Translation Resources Extraction. *Procesamiento del Lenguaje Natural*. 39:265–272. Setembro de 2007. <http://www.sepln.org/revistaSEPLN/revista/39/32.pdf>. 11
- (Simões e Almeida, 2008) Alberto Simões e José João Almeida. Bilingual Terminology Extraction based on Translation Patterns. *Procesamiento del Lenguaje Natural*. 41:281–288. Setembro de 2008. <http://alfarrabio.di.uminho.pt/~albie/publications/sepln08.pdf>. 10
- (Sinclair, 1991) John M. Sinclair. *Corpus, Concordance, Collocation*. Oxford University Press. 1991. 94
- (Singhal, 2008) Amit Singhal. Web Search: Challenges and Directions. Em Craig MacDonald, Iadh Ounis, Vassilis Plachouras, Ian Ruthven e Ryen W. White, editores, *Advances in Information Retrieval, 30th European Conference on IR Research, ECIR 2008, Glasgow, UK, March 30-April 3, 2008. Proceedings*. 2008. p. 2. Springer. 72
- (Sánchez-Martínez e Forcada, 2007) Felipe Sánchez-Martínez e Mikel L. Forcada. Automatic induction of shallow-transfer rules for open-source machine translation. Em *Proceedings of the The 11th Conference on Theoretical and Methodological Issues in Machine Translation (TMI'2007)*. Skövde, Suécia. 2007. p. 181–190. 10
- (Sousa Silva, 2006) Rui Sousa Silva. Performance and Individual Act Out: The Semantics of (Re)Building and (De)Constructing in Contemporary Artistic Discourse. Tese de mestrado. Faculdade de Letras da Universidade do Porto. 2006. 95
- (SPEECHDAT, 1998) Portuguese SpeechDat (II) FDB-4000, European Language Resources Association. 1998. <http://www.elda.org/catalogue/en/speech/S0092.html>. 100, 101
- (Stamatatos et al., 2000) Efstathios Stamatatos, George Kokkinakis e Nikos Fakotakis. Automatic text categorization in terms of genre and author. *Computational Linguistics*. 26: 471–495. 2000. 92
- (Steinberger et al., 2006) Ralf Steinberger, Bruno Pouliquen, Anna Widiger, Camelia Ignat, Tomaz Erjavec, Dan Tufis e Dániel Varga. The JRC-Acquis: A multilingual aligned parallel corpus with 20+ languages. Em Nicoletta Calzolari, Khalid Choukri, Aldo Gangemi, Bente Maegaard, Joseph Mariani, Jan Odjik e Daniel Tapias, editores, *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC'2006)*. Génova, Itália. 22-28 de Maio de 2006. p. 2142–2147. 8
- (Teixeira et al., 2006) António Teixeira, Catarina Oliveira e Lurdes Moutinho. On the Use of Machine Learning and Syllable Information in European Portuguese Grapheme-Phone Conversion. p. 212–215. Volume 3960. Springer. 2006. 100

- (TIGER, 2007) TIGER PROJECT Linguistic Interpretation of a German Corpus. 2007. <http://www.ims.uni-stuttgart.de/projekte/TIGER/>. 5
- (Uzeda Garrão, 2006) Milena de Uzeda Garrão. *O Córpus não mente jamais: Sobre a identificação e uso de combinações multivocabulares do tipo "verbo + sintagma nominal"*. Tese de doutoramento. Pontifícia Universidade Católica do Rio de Janeiro. 2006. 66
- (Uzeda Garrão e Dias, 2006) Milena de Uzeda Garrão e Carmelita Dias. The corpus never lies: a statistical approach for the identification of verbal collocations. Em *Proceedings of Collocations and Idioms 1: Papers from the First Nordic Conference on Syntactic Freezes*. Joensuu, Finlândia. 19-20 de Maio de 2006. 66
- (Viana e Andrade, 1985) Maria do Céu Viana e Ernesto de Andrade. CORSO I: um conversor de texto ortográfico em código fonético para o português. Relatório Técnico n.º 6. Grupo de Fonética e Fonologia, Centro de Linguística da Universidade de Lisboa. 1985. 100
- (Wielemaker, 2008) Jan Wielemaker. SWI-Prolog 5.6 Reference Manual. 2008. <http://gollem.science.uva.nl/cgi-bin/nph-download/SWI-Prolog/refman/refman.pdf>. 2