Βιβλιογραφία

[1]   Abney, S. (1991). Parsing by Chunks. In Berwick, Abney, and Tenny (eds), *Principle-based Parsing.* Kluwer Academic Publishers.

[2]   Abney, S. (1996). Statistical Methods and Linguistics. In Klavans, J. & Resnik, P. (eds) *The Balancing Act*, Cambridge, MA: MIT Press.

[3]   Allen, J. (1987). *Natural Language Understanding.* Benjamin Cummings.

[4]   Antonacci, F. and C. Calamani (1992). Capturing the Deep Meaning of Texts through Deduction and Inference. *IBM Journal of Research and Development*, 36(3), pp. 333-345.

[5]   Antworth, E. (1990). *PC-KIMMO: A Two-Level Processor for Morphological Analysis.* Summer Institute of Linguistics.

[6]   Baayen, H., H. Van Halteren, and F. Tweedie (1996). Outside the Cave of Shadows: Using Syntactic Annotation to Enhance Authorship Attribution. *Literary and Linguistic Computing*, 11(3), pp. 121-131.

[7]   Biber, D. (1988). *Variation Across Speech and Writing.* Cambridge University Press.

[8]   Biber, D. (1989). A Typology of English Texts. *Linguistics*, 27, pp. 3-43.

[9]   Biber, D. (1990). Methodological Issues Regarding Corpus-based Analyses of Linguistic Variations. *Literary and Linguistic Computing*, 5, pp. 257-269.

[10] Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8, pp. 1-15.

[11] Biber, D. (1995). *Dimensions of Register Variation: A Cross-linguistic Comparison.* Cambridge University Press.

[12] Bikel, D., S. Miller, R. Schwartz, and R. Weischedel (1997). Nymble: a High-Performance Learning Name-finder. In *Proc. of the Fifth Conference on Applied Natural Language Processing (ANLP'97)*, pp. 194-201.

[13] Brill, E. (1995). Transformation-Based Error-Driven Learning and Natural Language Processing: A Case Study in Part-of-Speech Tagging. *Computational Linguistics*, 21(4), pp. 543-565.

[14] Brill, E. and R. Mooney (1997). An Overview of Empirical Natural Language Processing. *AI Magazine*, 18(4), pp.13-24.

[15] Brinegar, C. (1963). Mark Twain and the Quintus Curtius Snodgrass Letters: A Statistical Test of Authorship. *Journal of the American Statistical Association*, 58, pp. 85-96.

[16] Bourigault, D. (1992). Surface Grammatical Analysis for the Extraction of Terminological Noun Phrases. In *Proc. of the Fifteenth Int. Conference on Computational Linguistics*, 3, pp. 977-981.

[17] Brunet, E. (1978). *Vocabulaire de Jean Giraudoux: Structure et Evolution.* Slatkine.

[18] Burrows, J. (1987). Word-patterns and Story-shapes: The Statistical Analysis of Narrative Style. *Literary and Linguistic Computing*, 2(2), pp. 61-70.

[19] Burrows, J. (1992). Not Unless You Ask Nicely: The Interpretative Nexus Between Analysis and Information. *Literary and Linguistic Computing*, 7(2), pp. 91-109.

[20] Cardie, C. (1997). Empirical Methods in Information Extraction. *AI Magazine*, 18(4), pp. 65-80.

[21] Charniak, E. (1997). Statistical Techniques for Natural Language Parsing. *AI Magazine*, 18(4), pp. 33-44.

[22] Chomsky, N. (1957). *Syntactic Structures.* The Hague, Mouton.

[23] Church, K. (1988). A Stochastic Parts Program and Noun Phrase Parser for Unrestricted Text. In *Proc. of Second Conference on Applied Natural Language Processing*, pp. 136-143.

[24] Church, K. and M. Liberman (1991). A Status Report on the ACL/DCI. In *Proc. of the 7th Annual Conference of the UW Centre for the New OED and Text Research: Using Corpora*, pp. 84-91.

[25] Cowie, J. and W. Lehnert (1996). Information Extraction. *Communications of the ACM*, 39(1), pp. 80-91.

[26] Croce, B. (1922). *Aesthetics.* New York, MacMillan.

[27] Damerau, F. (1975). The Use of Function Word Frequencies as Indicators of Style. *Computers and the Humenities*, 9, pp. 271-280.

[28] Darbyshire, A. (1971). *A Grammar of Style.* London, Andre Deutch.

[29] Dermatas, E. and G. Kokkinakis (1995). Automatic Stochastic Tagging of Natural Language Texts. *Computational Linguistics*, 21(2), pp. 137-164.

[30] DiMarco, C. and G. Hirst (1993). A Computational Theory of Goal-Directed Style in Syntax. *Computational Linguistics*, 19(3), pp. 451-499.

[31] Dyer, M. (1983). *In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension.* MIT Press.

[32] Edwards, A. (1979). *Multiple Regression and the Analysis of Variance and Covariance.* San Francisco, CA: W. H. Freeman.

[33] Eisenbeis, R., and R. Avery (1972). *Discriminant Analysis and Classification Procedures: Theory and Applications.* Lexington, Mass.: D.C. Health and Co.

[34] Enkvist, N. (1973). *Linguistic Stylistics.* The Hague, Mouton.

[35] Fakotakis, N., A. Tsopanoglou, and G. Kokkinakis (1993). A Text-independent Speaker Recognition System Based on Vowel Spotting. *Speech Communication*, 12, pp. 57-68.

[36] Ferguson, C. (1994). Dialect Register and Genre: Working assumptions about conventionalization. In D. Biber and E. Finegan (eds) *Sociolinguistics Perspectives on Register*, pp. 15-30, Oxford University Press.

[37] Forsyth, R. and D. Holmes (1996). Feature-Finding for Text Classification. *Literary and Linguistic Computing*, 11(4), pp. 163-174.

[38] Fowler, R. (1972). Style and the Concept of Deep Structure. *Journal of Literary Semantics*, 1, pp. 5-24.

[39] Goodman, J. (1997). Global Thresholding and Multiple-Pass Parsing. In *Proc. of the Second Conference on Empirical Methods in Natural Language Processing*, pp. 11-25.

[40] Gray, B. (1969). *Style: The Problem and its Solution.* The Hague, Mouton.

[41] Guthrie, L., J. Pustejofsky, Y. Wilks, and B. Slator (1996). The Role of Lexicons in Natural Language Processing. *Communications of the ACM*, 39(1), pp. 63-72.

[42] Hobbs, J., D. Appelt, J. Bear, D. Israel, M. Kameyama, M. Stickel, and M. Tyson (1996). FASTUS: a Cascaded Finite-State Transducer for Extracting Information from Natural-Language Text. In E. Roche and Y. Schabes (eds) *Finite State Devices for Natural Language Processing.* Cambridge MA: MIT Press.

[43] Holmes, D. (1992). A Stylometric Analysis of Mormon Scripture and Related Texts. *Journal of the Royal Statistical Society,* Series A, 155(1), pp. 91-120.

[44] Holmes, D. (1994). Authorship Attribution. *Computers and the Humanities*, 28, pp. 87-106.

[45] Holmes, D. and R. Forsyth (1995). The Federalist Revisited: New Directions in Authorship Attribution. *Literary and Linguistic Computing*, 10(2), pp. 111-127.

[46] Honore, A. (1979). Some Simple Measures of Richness of Vocabulary. *Association for Literary and Linguistic Computing Bulletin*, 7(2), pp. 172-177.

[47] Hovy, E, (1987). *Generating Natural Language Under Pragmatic Constraints.* Ph.D. Dissertation, Yale University.

[48] Johnson, R., and D. Wichern (1992). *Applied Mutlivariate Statistical Analysis* (third edition). Prentice-Hall International Inc.

[49] Karlgren, J., and D. Cutting (1994). Recognizing text Genres with Simple Metrics Using Discriminant Analysis. In *Proc. of the 15th International Conference on Computational Linguistics (COLING '94)*.

[50] Karlgren, J. (1996). Stylistic Variation in an Information Retrieval Experiment. In *Proc. of the 2nd Conference on New Methods in Language Processing (NEMALP-2)*.

[51] Karlgren, J. (1999). Stylistic Experiments in Information Retrieval. In T. Strzalkowski (ed.) *Natural Language Information Retrieval*, Kluwer Academic Publishers, pp. 147-166.

[52] Karlsson, F., A. Voutilainen, J. Heikkila, and A. Anttila (1995). *A Language-Independent System for Parsing Unrestricted Text.* Mouton de Gruyter.

[53] Kavallieratou E., E. Stamatatos, N. Fakotakis, and G. Kokkinakis (2000). Handwritten Character Segmentation Using Transformation-Based Learning. In *Proc. of the 15th Int. Conf. on Pattern Recognition (ICPR2000)*.

[54] Kessler, B., G. Nunberg, and H. Schutze (1997). Automatic Detection of Text Genre. In *Proc. of 35th Annual Meeting of the Association for Computational Linguistics (ACL/EACL '97)*, pp. 32-38.

[55] Kilgarriff, A. (1997). Using Word Frequency Lists to Measure Corpus Homogeneity and Similarity between Corpora. In *Proceedings of the 5th ACL Workshop on Very Large Corpora*, pp. 231-245.

[56] King, M. (1996). Evaluating Natural Language Processing Systems. *Communications of the ACM*, 39(1), pp.73-79.

[57] Kittredge, R. and J. Lehrberger (1982). Sublanguage: Studies of Language in Restricted Semantic Domains. Berlin, De Gruyter.

[58] Knight, K. (1997). Automating Knowledge Acquisition for Machine Translation. *AI Magazine*, 18(4), pp. 81-96.

[59] Leech, G. and M. Short (1981). *Style in Fiction: A Linguistic Introduction to English Fictional Prosa.* London, Longman.

[60] Lewis, D. and K. Spärck Jones (1996). Natural Language Processing for Information Retrieval. *Communications of the ACM*, 39(1), pp. 92-101.

[61] Martin, J. (1985). Process and Text: Two Aspects of Human Semiosis. In J. Benson and W. Greaves (eds) *Systemic Perspectives on Discourse*, I, pp. 248-274, Norwood.

[62] Μίχος, Σ. (1997). *Υπολογιστική Συντακτική Επεξεργασία Κειμένων της Νέας Ελληνικής Γλώσσας Βασισμένη σε Εμπειρικές Μεθόδους.* Διδακτορική διατριβή, Πανεπιστήμιο Πατρών.

[63] Michos, S., E. Stamatatos, N. Fakotakis, and G. Kokkinakis (1996). An Empirical Text Categorizing Computational Model Based on Stylistic Aspects. In *Proc. of 8th IEEE Conference on Tools with Artificial Intelligence (ICTAI'96)*, pp. 71-77.

[64] Mikheev, A. (1997). Automatic Rule Induction for Unknown Word Guessing. *Computational Linguistics*, 23(3), pp. 405-423.

[65] Morton, A. (1965). The Authorship of Greek Prose. *Journal of the Royal Statistical Society,* Series A, 128, pp. 169-233.

[66] Mosteller, F. and D. Wallace (1984). *Applied Bayesian and Classical Inference: The Case of the Federalist Papers*. Addison-Wesley, Reading, MA.

[67] Muller, H., V. Amerl, and G. Natalis (1980). Worterkennungsverfahren als Grundlage einer Universalmethode zur automatischen Segmentierung von Texten in Satze. *Eim Verfahren zur maschinellen Satzgrenzen-bestimmung im Englischen, Sprache und Datenvarbeitung*, 1.

[68] Noussia, T. (1997). A Rule-based Hyphenator for Modern Greek. *Computational Linguistics*, 23(3), pp. 361-376.

[69] Oakman, R. (1980). *Computer Methods for Literary Research.* Columbia, University of South Carolina Press.

[70] Ohmann, R. (1964). Generative Grammars and the Problem of Literary Style. *Word*, 20, pp. 423-439.

[71] Palmer, D. and M. Hearst (1997). Adaptive Multilingual Sentence Boundary Disambiguation. *Computational Linguistics*, pp. 241-267, 23(2).

[72] Potter, R. (1991). Statistical Analysis of Literature: A Retrospective on Computers and the Humanities, 1966-1990. *Computers and the Humanities*, 25, pp. 401-429.

[73] Ramshaw, L. and Marcus M. (1995). Text Chunking Using Transformation-based Learning. In *Proc. of ACL Third Workshop on Very Large Corpora.* pp. 82-94.

[74] Raskin, R. (1986). The Quest for Style. *IEEE Transactions on Professional Communication*, 29(3), pp. 10-18.

[75] Reynar, J. and A. Ratnaparkhi (1997). A Maximum Entropy Approach to Identifying Sentence Boundaries. In *Proc. of the Fifth Applied Natural Language Processing Conference*.

[76] Riesel, E. (1971). *Stilistik der Deutschen Sprache* (second edition). Moskau, 1963.

[77] Riley, M. (1989). Some Applications of Tree-based Modeling, to Speech and Language Indexing. In *Proc. of the DARPA Speech and Natural Language Workshop*, pp. 339-352.

[78] Russel, W. (1975). Some Style Rules for a Transformational Grammar. *Linguistics*, 157, pp. 87-101.

[79] Samuel, K. (1998). Dialogue Act Tagging with Transformation-Based Learning. In *Proc. of the 17th International Conference on Computational Linguistics (COLING-ACL '98)*

[80] Schwartz, R, L. Nguyen, and J. Makhoul (1996). Multiple-Pass Search Strategies. In C. Lee, F. Soong, and K. Paliwal (eds) *Automatic Speech and Speaker Recognition: Advanced Topics*, Kluwer Academic Publishers, pp. 429-456.

[81] Σγάρμπας, Κ. (1997). *Τεχνικές Μορφοσυντακτικής Ανάλυσης της Νέας Ελληνικής Γλώσσας*. Διδακτορική διατριβή, Πανεπιστήμιο Πατρών.

[82] Sichel, H. (1975). On a Distribution Law for Word Frequencies. *Journal of the American Statistical Association*, 70, pp. 542-547.

[83] Sichel, H. (1986). Word Frequency Distributions and Type-Token Characteristics. *Mathematical Scientist*, 11, pp. 45-72.

[84] Simpson, E. (1949). Measurement of Diversity. *Nature*, 163, pp. 168.

[85] Skut, W. and Brants T. (1998). Chunk Tagger: Statistical Recognition of Noun Phrases. In *ESSLLI-98 Workshop on Automated Acquisition of Syntax and Parsing.*

[86] Smith, M. (1987). The Revenger's Tragedy: The Derivation and Interpretation of Statistical Results for Resolving Disputed Authorship. *Computers and the Humanities*, 21, pp. 21-55.

[87] Smith, M. (1987). The Authorship of Pericles: New Evidence for Wilkins. *Journal of the Association for Literary and Linguistic Computing*, 2(4), pp. 221-230.

[88] Smith, M. (1991). The Authorship of The Revenger's Tragedy. *Notes and Queries*, 38(4), pp. 508-513.

[89] Srinivas, B. (1997). Performance Evaluation of Supertagging for Partial Parsing. In *Proc. of the Fifth International Workshop on Parsing Technologies.*

[90] Stamatatos, E., S. Michos, N. Fakotakis, and G. Kokkinakis (1997). A User-Assisted Business Letter Generetor Dealing with Text's Stylistic Variations. In *Proc. of the Ninth IEEE Conference on Tools with Artificial Intelligence (ICTAI'97)*, pp. 182-189.

[91] Strzalkowski, T. (1994). Robust Text Processing in Automated Information Retrieval. In *Proc. of the 4th Conf. On Applied Natural Language Processing,* pp. 168-173.

[92] Tabachnick, B, and L. Fidell (1996). *Using Multivariate Statistics* (third edition). HarperCollins College Publishers.

[93] Tufte, V. (1971). *Grammar as Style.* New York, Rinehart and Winston, Inc.

[94] Tweedie, F., S. Singh, and D. Holmes (1996). Neural Network Applications in Stylometry: The Federalist Papers. *Computers and the Humanities*, 30(1), pp. 1-10.

[95] Tweedie, F. and Baayen, R. (1998). How Variable may a Constant be? Measures of Lexical Richness in Perspective. *Computers and the Humanities*, 32(5), pp.323-352.

[96] Ullmann, S. (1972). *Sprache und Stil.* Tubingen, Max Niemeyer Verlag.

[97] Vachek, J. (1975). Some Remarks on Functional Dialects of Standard Languages. In H. Ringbom (ed.) *Style and Text*, Stockholm, Scriptor and Abo Akademi

[98] Voutilainen, A. (1993). NPtool, a Detector of English Noun Phrases. In *Proc. of the Workshop on Very Large Corpora: Academic and Industrial Perspectives,* Ohio State University, pp. 48-57.

[99] Waltz, D. (1982). The State of the Art in Natural-Language Understanding. In W. Lehnert and M. Ringle (eds) *Strategies for Natural Language Processing*, pp. 3-32, Lawrence Erlbaum Associates, Publishers, New Jersey.

[100] Χαραλαμπάκης, Χ. (1992). *Νεοελληνικός Λόγος: Μελέτες για την Γλώσσα, την Λογοτεχνία και το Ύφος.* Νεφέλη, Αθήνα.

[101] Yang, Y. (1999). An Evaluation of statistical approaches to text categorization. *Information Retrieval Journal*, 1(1), pp. 69-90.

[102] Yang, Y. and X. Liu (1999). A re-examination of text categorization methods. In *Proc. of ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'99)*.

[103] Yule, G. (1944). *The Statistical Study of Literary Vocabulary.*Cambridge University Press, Cambridge.

[104] _, (1991). *Proceedigs of the 29th Annual Meeting of the Association for Computational Linguistics (ACL'91).* University of California.

[105] _, (1996). *Proceedings of the 34th Annual Meeting of the ACL (ACL'96).* Santa Cruz, CA.

[106] _, (1995). *Proceedings of the 6th Message Understanding Conference (MUC-6).* Columbia, Advanced Research Projects Agency, Information Technology Office, Maryland.

[107] _, (1992). *Συντακτικό της Νεάς Ελληνικής.* Οργανισμός Εκδοσεων Διδακτικών Βιβλίων, Αθήνα.

Δημοσιεύσεις

Περιοδικά

[1] **Stamatatos, E.**, N. Fakotakis, and G. Kokkinakis (to appear). Computer-Based Authorship Attribution without Lexical Measures. *Computers and the Humanities.*

[2] Michos S., **E. Stamatatos**, and N. Fakotakis (1999). Supporting Multilinguality in Library Automation Systems Using AI Tools. *Applied Artificial Intelligence*, 13(7), pp. 679-704.

[3] **Stamatatos, E.**, N. Fakotakis, and G. Kokkinakis (under review). On Detecting Sentence and Chunk Boundaries in Unrestricted Text Based on Minimal Resources. *Journal of Natural Language Engineering.*

[4] **Stamatatos, E.**, N. Fakotakis, and G. Kokkinakis (under review). Automatic Text Categorization in Terms of Genre and Author. *Computational Linguistics.*

Βιβλία

[5] Michos S., **E. Stamatatos**, N. Fakotakis, and G. Kokkinakis (1996). Categorising Texts by Using a Three-Level Functional Style Description. In A. Ramsay (ed.) *Artificial Intelligence: Methodology, Systems, Applications*, Frontiers in Artificial Intelligence and Applications, 35, IOS press, pp 191-198.

Συνέδρια με πλήρη κρίση

[6]   **Stamatatos, E.**, N. Fakotakis, and G. Kokkinakis (2000). Text Genre Detection Using Common Word Frequencies. In *Proc. of the 18th Int. Conf. on Computational Linguistics (COLING2000)*.

[7]   Kavallieratou E., **E. Stamatatos**, N. Fakotakis, and G. Kokkinakis (2000). Handwritten Character Segmentation Using Transformation-Based Learning. In *Proc. of the 15th Int. Conf. on Pattern Recognition (ICPR2000)*.

[8]   **Stamatatos, E.**, N. Fakotakis, and G. Kokkinakis (1999). Automatic Authorship Attribution. In *Proc. of the 9th Conf. of the European Chapter of the Association for Computational Linguistics (EACL'99),* pp. 158-164.

[9]   **Stamatatos, E.**, N. Fakotakis, and G. Kokkinakis (1999). Automatic Extraction of Rules for Sentence Boundary Disambiguation. In *Proc. of the Workshop in Machine Learning in Human Language Technology, Advance Course on Artificial Intelligence (ACAI'99),* pp. 88-92.

[10] **Stamatatos, E.**, S. Michos, N. Fakotakis, and G. Kokkinakis (1997). A User-Assisted Business Letter Generetor Dealing with Text's Stylistic Variations. In *Proc. of the 9th IEEE Conference on Tools with Artificial Intelligence (ICTAI'97)*, pp. 182-189.

[11] Michos S., **E. Stamatatos**, N. Fakotakis, and G. Kokkinakis (1996). An Empirical Text Categorizing Computational Model Based on Stylistic Aspects. In *Proc. of the 8th IEEE Conference on Tools with Artificial Intelligence (ICTAI'96)*, pp. 71-77.

Συνέδρια με κρίση περίληψης

[12] **Stamatatos, E.**, N. Fakotakis, and G. Kokkinakis (2000). A Practical Chunker for Unrestricted Text. In *Proc. of the 2nd Int. Conf. of Natural Language Processing (NLP2000)*, pp. 139-150.

[13] **Stamatatos, E.**, S. Michos, C. Patelodimou, and N. Fakotakis (1997). TRANSLIB: An Advanced Tool for Supporting Multilingual Access to Library Catalogues. In *Proc. of the 2nd Workshop on Multilinguality in Software Industry (MULSAIC'97)*, pp. 33-40.

[14] Michos S., **E. Stamatatos**, N. Fakotakis, and G. Kokkinakis (1996). Identification of Functional Style Based on a Three-Level Stylistic Description. In *Proc. of Workshop on Language Engineering for Document Analysis and Recognition (LEDAR'96)*.