

Curriculum Vitae

Patrizia Paggio

Born in Rome on 2nd March 1959.
Nationality: Italian

Current position

Senior researcher at the University of Copenhagen, Centre for Language Technology (CST)

Education

- 1997: PhD in computational linguistics at the University of Copenhagen. Title of thesis: *The Treatment of Information Structure in Machine Translation*. Reviewers: Elisabet Engdahl and Eva Hajičová.
- 1988: Master's in English (major) and Computational Linguistics (subsidiary subject) at the University of Copenhagen.
- 1981: Laurea in Modern Languages and Literatures at the University of Rome (English and Russian).
- 1977: Diploma in Classical Studies at Liceo Mamiani in Rome.

Personal grants

- 2008: 3-year grant from the Danish Research Council for the project "Verbal and bodily communication".
- 2002: 3-year grant from the Carlsberg Foundation in Copenhagen for the project "Information Structure in Danish".
- 1992: EU (then CEC) research grant for a one-year study of discourse phenomena in machine translation.

Research positions

- 1991- : Employed as a researcher by Center for Sprogteknologi, senior researcher since April 1999.
- 1992: Guest researcher at the Centre for Cognitive Science at the University of Edinburgh.
- 1987-1990: Employed as a researcher by the MT project Eurotra-DK

Research projects and tasks

Selected projects and tasks at CST (1991-2009)

Current:

- Nordic project NOMCO: CST coordinator, responsible for the creation of a multimodal resource for Danish as well as investigations based on this resource.
- Project "Verbal and bodily communication": fundraiser and main researcher, responsible for the development of theories and corpora on multimodal communication.
-

Completed:

- Project “Information Structure in Danish”: fundraiser and main researcher, responsible for the development of a model of information structure for the Danish language.
- EU-funded MOSES project: main fundraiser and CST project manager, responsible for the analysis of Danish queries in the context of a methodology for ontology-based annotation and search of data on university Web sites.
- Nordic network MUMIN, on multimodal interfaces: main fundraiser and coordinator.
- Danish research project OntoQuery, on ontology based information retrieval: responsible for design and development of syntactic and semantic analysis.
- Danish research project “The Staging of Virtual 3D-Spaces”: CST project manager for CST, and responsible for the design and implementation of the natural language dialogue component of a multimedia application.
- GRASP project: project manager and responsible for the design and development of Danish grammars in the HPSG and LFG frameworks.
- EU-funded SCARRIE project, on spelling and grammar checking: responsible for Danish grammar development and the establishment of a general evaluation methodology.
- EU-funded LCC (Language Consulting Centre) project: contributions to seminars on language technology aimed at Danish industry.
- EU-funded TEMAA project (A Testbed Study of Evaluation Methodologies: Authoring Aids): CST project manager.
- PaTrans MT project: adaptation of the Eurotra English grammars for a commercial MT system.

Tasks at Eurotra-DK (1987-1990)

- Design and implementation of substantial parts of the Danish analysis and generation grammars.
- Implementation of Italian-Danish transfer component.
- Coordination of the Special Group of Interest on Preference, which designed the Eurotra preference mechanism.

Previous work experience

- 1985-1989: Part time teacher in Italian literature at Copenhagen Business School.
 1985-1987: Teacher of Italian at the Studieskolen, Copenhagen.
 1983-1986: Teacher of Italian at AOF, Copenhagen
 1982-1987: Teacher of Italian at HOF, Copenhagen
 1981-1982: Freelance translator English-Italian.
 1980-1982: Teacher of English at the Oxford Institute in Rome.

Recent administrative experience

Current tasks:

- Member of HumanT, think tank of the Faculty of Humanities
 Member of the advisory group of the Centre for IT Innovation at the University of Copenhagen
 Staff representative in the CST advisory board
 Deputy member of study committee at the Department of Danish Studies and Linguistics
 Scientific coordinator of Master’s programme in IT and Cognition

Recent tasks:

- 2008 Chair of assessment committee concerning an assistant professorship at CST.
- 2007 Chair of study committee responsible for the creation of the Master's programme in IT and Cognition at the University of Copenhagen. The programme, launched in the Autumn 2007, is an interdisciplinary initiative involving the Institute of Nordic Studies and Linguistics, the Institute of Psychology and the Institute of Computer Science.
- Member of assessment committee concerning an assistant professorship at the IT University of Copenhagen

Selected teaching experience**Current teaching:**

Courses in language technology, neurolinguistics, information search and multimodal communication for the Master's programme It og Cognition at the University of Copenhagen.

Earlier teaching:

Graduate and postgraduate courses in formal semantics at the University of Copenhagen, and machine translation at the Copenhagen Business School. Courses on multimodality at several Nordic PhD schools.

Other:

External examination reviewer (*censor*) for the University of Copenhagen, the Copenhagen Business School and the University of Southern Denmark.

Workshop organisation and demonstrations

- Workshop on Multimodal Communication at Nodalida 2009, Odense (Denmark).
- Workshop on Multimodal Corpora at LREC2008, Marrakesh (Morocco).
- Workshop on Multimodal Corpora at LREC2006, Genoa (Italy).
- MUMIN workshop on Annotation of Multimodal Corpora, KTH in Stockholm (SE), June 2004.
- Ontolex 2004 workshop at LREC2004, Lisbon (Portugal).
- Workshop on Multimodal Corpora at LREC2004, Lisbon (Portugal).
- 1st Nordic Symposium on Multimodal Communication, Marienlyst (Denmark) under the sponsorship of the MUMIN network, 2003.
- Workshop and PhD course on multimodality, MUMIN network, Tampere (Finland), 2002.

Programme committees

- 17th Nordic Conference of Computational Linguistics NODALIDA, Odense 2009.
- 11th Conference of the Italian Association for Artificial Intelligence (AI*IA '09), Reggio Emilia 2009.
- 16th Nordic Conference of Computational Linguistics NODALIDA, Tartu 2007.

Editorial work and reviewing**Conference paper reviews:**

- LREC 2008, Marrakesh 2008
- LREC 2006, Genoa 2006

- 19th International Joint Conference on Artificial Intelligence, Edinburgh 2005.
- LREC 2004, Lisbon 2004
- Annual Conference of the Association of Computational Linguistics ACL '04, Barcelona 2004.
- 5th International Conference on Multimodal Interfaces (ICMI-PUI'03), Nov. 5-7th 2003.
- FQAS 2002 (Flexible Question Answering Systems), Copenhagen, October 2002.
- FQAS 2000 (Flexible Question Answering Systems), Warsaw, October 2000.

Editorial boards:

- Nordic European Journal of Language Technology
- Special number of Journal of Linguistic Resources on Multimodal Corpora
- Post-proceedings for various workshops (cf workshop organization)

Selected journal reviews:

Acta Linguistica Hafnensia, Nordic Journal of Linguistics, Journal of Language Resources and Evaluation

Project reviews:

Several projects for the Nordic Research Council and the EU Commission

Scientific associations

Association of Computational Linguistics, Copenhagen Linguistic Circle, Italian Association of Artificial Intelligence.

Languages

Italian (mother tongue), Danish (proficient), English (proficient), Swedish (intermediate), French (beginner)

Refereed publications

Publications marked with a star are enclosed for review.

Paggio, P. (2009) The information structure of Danish grammar constructions. In *Nordic Journal of Linguistics*. 32.1, 137–164. Cambridge University Press.

Paggio, P., D. H. Haltrup and L. Offersgaard (2009) Ontology-based question-answering with feature structures. In Anders Søgaard and Petter Haugereid (eds.) *Typed Feature Structure Grammars*, pp.159–173. Peter Lang Publishing Group.

Jokinen, K., C. Navarretta & P.Paggio (2008) Distinguishing the communicative functions of gestures. In Proceedings of the 5th Joint Workshop on Machine Learning and Multimodal Interaction, 8-10 September 2008, Utrecht, The Netherlands. Springer LNCS 5237, pp. 38–49.

Allwood, J., L. Cerrato, K. Jokinen, C. Navarretta & P. Paggio (2007) The MUMIN Coding Scheme for the Annotation of Feedback, Turn Management and Sequencing. In J. C. Martin et al. (eds) *Multimodal Corpora for Modelling Human Multimodal Behaviour*. Special issue of the International Journal of Language Resources and Evaluation. Springer. Vol.41, no.3-4, pp.273–287.

Pedersen, B. S., P. Paggio & C. Navarretta (2007) Sprogteknologiske ressourcer til informationsøgning. In *Dansk Biblioteksforskning—Tidsskrift for informations- og kulturformidling*. Vol.3, no 2, pp.5–18.

Paggio, P., D. H. Haltrup and L. Offersgaard (2007) QA with feature structures. In *Proceedings of the 1st International Workshop on Typed Feature Structures(TFSG'06)*. Aalborg, Denmark, June 2006, pp.146–159.

Martin, J.-C., P. Paggio, P. Kuehnlein, R. Stiefelhagen, F. Pianesi eds. (2007) *Multimodal Corpora for Modelling Human Multimodal Behaviour*. Special issue of the International Journal of Language Resources and Evaluation.

Braasch, A., Navarretta, C., Nimb, S., Olsen, S., Paggio, P. Pedersen B.S. and Wedekind, J. (2006) *Sprogteknologi i dansk perspektiv*. Copenhagen: Reitzel.

Paggio, P. and Pedersen, B. S (2006) Bedre Adgang til Viden – Informationsøgning med ontologi og sprogteknologi. In Braasch, A. et al (eds) *Sprogteknologi i dansk perspektiv*. Copenhagen: Reitzel, pp.329–353.

Martin, J.-C., Kuehnlein, P., Paggio, P., Stiefelhagen, R., Pianesi, F. eds. (2006) *Proceedings of the LREC2006 Workshop on Multimodal Corpora. From Multimodal Behaviour to Usable Models*. Genova, Italy, May 2006.

Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C. and Paggio, P. (2006) A Codings Scheme for the Annotation of Feedback, Turn Management and Sequencing Phenomena. In Martin et al (eds) *Proceedings of the LREC2006 Workshop on Multimodal Corpora. From Multimodal Behaviour to Usable Models*. Genova, Italy, May 2006, pp.38–42.

Paggio, P. (2006a) Annotating Information Structure in a Corpus of Spoken Danish. In Calzolari et. (eds) *Proceedings of the 5th International Conference on Language Resources and Evaluation LREC2006*. Genova, Italy, May 2006, 4 pages.

Paggio, P. (2006b) Information Structure and Pauses in a Corpus of Spoken Danish. In *Conference Companion of the 11th Conference of the European Chapter of the Association for Computational Linguistics*, Trento, Italy, April 2006, pp. 191–194.

Paggio, P (2006c) Representing Information Structure in a Formal Grammar of Danish. In Takashi Washio, Akito Sakurai, Katsuto Nakajima, Hideaki Takeda, Satoshi Tojo, Makoto Yokoo (Eds.): *New Frontiers in Artificial Intelligence, Joint JSAI 2005 Workshop Post-Proceedings*. Springer 2006, ISBN 3-540-35470-0, pp. 93–102.

Stellato, A., Paziienza, M. T., Henriksen, L., Paggio, P. and Zanzotto, F. M. (2005) Ontology mapping to support multilingual ontology-based question answering. In *Proceedings of 4th International Semantic Web Conference*. Galway, Ireland, November 2005, 6 pages.

Allwood, J., Cerrato, L., Jokinen, K., Navarretta, C. and Paggio, P. (2005) The MUMIN annotation scheme for feedback, turn management and sequencing. In *Proceedings of the The Second Nordic Conference on Multimodal Communication*. Gothenburg. April 2005, pp.91–110.

Paggio, P. (2005) Representing Information Structure in a Formal Grammar of Danish. *Proceedings of the 2nd International Workshop on Logic and Engineering of Natural Language Semantics (LENLS2005)*. Kitakyushu, Japan. June 13-14, pp. 37–49.

Pazienza, M. T., Stellato, A., Zanzotto, F. M., Henriksen, L. And Paggio, P. (2005) Ontology Mapping to Support Ontology-Based Question Answering. In *Proceedings of the Meaning 05 workshop*. Trento (IT), February 2005, pp. 67–73.

Paggio, P. and Jongejan, B. (2005) Multimodal Communication in Virtual Environments – Communicating with the Staging virtual farm. In O. Stock and M. Zancanaro (eds) *Multimodal Intelligent Information Presentation*, Kluwer Academic Publishers, pp.27–47. ISBN: 1-4020-3051-7.

Basili Roberto, Dorte H. Hansen, Patrizia Paggio, Maria Teresa Pazienza, Fabio Massimo Zanzotto, "Ontological resources and question answering", Workshop on Pragmatics of Question Answering, held jointly with NAACL 2004 Boston, Massachusetts, May 2004, 7 pages.

Atzeni, Paolo, Roberto Basili, Dorte H. Hansen, Paolo Missier, Patrizia Paggio, Maria Teresa Pazienza, Fabio Massimo Zanzotto (2004) Ontology-Based Question Answering in a Federation of University Sites: The MOSES Case Study. *NLDB 2004*: pp. 413–420. ISSN: 0302-9743.

Martin, J.C., Den Os, E., Kühnlein, P., Boves, L., Paggio, P. and Catizone, R. (2004) *Proceedings of the workshop on Multimodal Corpora: Models of Human Behaviour for the Specification and Evaluation of Multimodal Input and Output Interfaces*, held in conjunction with LREC04, Lisbon, May 2004.

Oltremari, A., Paggio, P., Gangemi, A., Pazienza, M.T., Calzolari, N., Pedersen, B.S. and Simov, K., eds. (2004) *Proceedings of OntoLex 2004: Ontologies and Lexical Resources in Distributed Environments*. Workshop held at LREC04 in Lisbon, May 2004.

Paggio, P., Hansen, D.H., Basili, R., Pazienza, M.T. and Zanzotto, F.M. (2004) Ontology-based question analysis in a multilingual environment: the MOSES case study. In *Proceedings of OntoLex 2004: Ontologies and Lexical Resources in Distributed Environments*, workshop held at LREC04 in Lisbon, May 2004, pp. 1–8.

Pedersen, B. S. and Paggio, P. (2004) A Danish Semantic Lexicon and its Application in Content-based Querying. *Nordic Journal of Linguistics*, Vol.27, No.1, pp.97–126.

Andreasen, T., P. A. Jensen, J. F. Nilsson, P. Paggio, B. S. Pedersen & H. E. Thomsen (2004): Content-based text querying with ontological descriptors. *Data & Knowledge Engineering* 48, pp. 199–219.

Paggio, P., Jokinen, K. and Jönsson, eds, (2003a) *Proceedings of the 1st Nordic Symposium on Multimodal Communication*, CST Working Papers No. 6, Copenhagen, Denmark. September 2003.

Paggio, P., Pedersen, B.S, and Haltrup, D. (2003b) Applying Language Technology to Ontology-based Querying – the OntoQuery Project, in *Journal for Applied Artificial Intelligence*. Special Issue on Artificial Intelligence for Cultural Heritage and Digital Libraries, 17 (8-9). Taylor & Francis, Philadelphia, pp. 817–833.

Paggio, P., Jokinen, K. and Jönsson, A. (2002) MUMIN – A Nordic Network for Multimodal Interfaces. *Proceedings of the Workshop on Multimodal Resources at LREC 2002*, pp. 81–84.

Andreasen, T., Jensen, P.A., Nilsson, J.F., Paggio, P. Pedersen, B.S. and Thomsen, H.E. (2002a) Ontological Extraction of Content for Text Querying. In B. Andersson, M. Bergholtz & P. Johannesson (eds) *Natural Language and Information Systems*. June 27-28, Stockholm, 2002. The 6th International Conference on Applications of Natural Language to Information Systems, Stockholm June 2002. Springer LNCS 2553, pp.123–136.

Andreasen, T., Jensen, P.A., Nilsson, J.F., Paggio, P. Pedersen, B.S. and Thomsen, H.E. (2002b) Ontoquery: Ontology-based Querying of Texts. In J. Karlgren, P. Kanerva, & B. Gambäck (eds.) *Proceedings of the AAAI-2002 Spring Symposium Series*. Stanford University, Palo Alto, March 2002, pp. 28–31.

Pedersen, B.S. and Paggio, P. (2002) Semantic Lexical Resources Applied to Content-based Querying – the OntoQuery Project. *Proceedings of LREC 2002*, Vol. V, pp. 1753 – 1758.

Paggio, P. and Jongejan, B. (2001) Multimodal Communication in the Virtual Farm of the Staging Project. *Proceedings of the workshop on Information Presentation and Natural Multimodal Dialogue 2001*, Verona, December 2001, pp. 41–45.

Paggio, P., Pedersen, B.S, and Haltrup, D. (2001) Applying Language Technology to Content-based Querying, in Bordoni, L. and Semeraro, G. (eds.) *Proceedings on the Workshop on Artificial Intelligence for the Cultural Heritage and Digital Libraries*, 25/9/2001, Bari (IT), pp.75–79.

Neville, A. and Paggio. P (2001) Developing a Danish grammar in the GRASP project: A construction-based approach to topology and extraction in Danish. *Proceedings of the Conference on Formal Grammars/Mathematics of Language (FGMOL) '01*. Helsinki, August 2001. Electronic Notes in Theoretical Computer Science, Volume 53.

Paggio, P. , Jongejan, B. and Madsen, C. B. (2000) Unification-based Multimodal Analysis in a 3D Virtual World: the Staging project, in *Proceedings of the CELE-Twente Workshop on Language Technology: Interacting Agents*, pp. 71–82.

Paggio, P. and Jongejan, B. (2000) Representing multimodal input in a unification-based system: the Staging project, in *Proceedings of the Workshop on Integrating Information from Different Channels in Multi-Media Contexts at ESSLLI*, pp. 24–31.

- Paggio, P. (2000a) Spelling and Grammar Correction for Danish in SCARRIE, *ANLP Proceedings*, Seattle (WA) USA, pp.256–261.
- Paggio, P. (2000b) Syntactic Analysis and Error Correction for Danish in the SCARRIE Project, *Proceedings from the 12th “Nordiske datalingvistikdager”*, Trondheim (Norway), pp. 152–161.
- Paggio, P. (2000c) Introduction to Section 3, Verbal and Non-Verbal Interaction with Virtual Worlds and Agents. In Qvortrup (ed.) *Virtual Interaction: Interaction in Virtual Inhabited 3D Worlds*. Springer-Verlag, pp. 217–220.
- Paggio, P. and Music, B. (2000) Linguistic Interaction in Staging - A Language Engineering View. In Qvortrup (ed.) *Virtual Interaction: Interaction in Virtual Inhabited 3D Worlds*. Springer-Verlag, pp. 235–249.
- Paggio, P. (2000) Information Structure and Machine Translation: Generating Danish Existential Sentences, in N. Nicolov and R. Mitkov (eds) *Recent Advances in Natural Language Processing II*, John Benjamins Publishers, Amsterdam, pp. 271–282.
- Paggio, P. and Music, B. (1998) Evaluation in the SCARRIE Project. *Proceedings of the First International Conference on Language Resources & Evaluation*. Granada, Spain, pp. 277–282.
- Paggio, P. and Underwood, N. (1998) Validating the TEMAA LE evaluation methodology: a case study on Danish spelling checkers. *Journal of Natural Language Engineering*, Cambridge: Cambridge University Press, 4(3) pp.211–228.
- Paggio, P. (1997a) Information Structure and MT: Generating Danish Existential Sentences. *Proceedings of Recent Advances in Natural Language Processing*, pp.66–73.
- Paggio, P. (1997b) *The Treatment of Information Structure in Machine Translation*. PhD thesis. CST, Copenhagen.
- Paggio, P. (1995) Theme and Focus in an MT Application, in Bosch, P. & R. van der Sandt (eds.) *Focus and Natural Language Processing*. IBM Working papers. Heidelberg, pp. 553–562.
- Bennett, P. and Paggio, P. eds. (1993) Preference in Eurotra, in *Studies in Machine Translation and Natural Language Processing*, CEC Luxembourg.
- Paggio, P. and Ørsnes, B. (1993) Automatic Translation of Nominal Compounds. A case Study of Danish and Italian, in *Rivista di Linguistica*, Torino: Rosenberg & Sellier, pp. 129–156.
- Dologlou, Y., Malnati, G. and Paggio, P. (1991) A Preference Mechanism Based on Multiple Criteria Resolution, in *Proceedings for the 5th Conference of EACL*, Berlin, pp. 281–285.
- Malnati, G. and Paggio, P. (1991) The Eurotra User Language, in *Studies in Machine Translation and Natural Language Processing Vol.2*, CEC Luxembourg, pp.41–111.

Selected project reports and other publications

Allwood, J., Cerrato, L., Dybkjaer, L., Jokinen, K., Navarretta, C. And Paggio, P. (2005) The MUMIN multimodal coding scheme, in Henrik Holmboe (ed) *Nordisk Sprogteknologi 2005*. Museum Tusulanums Forlag. Copenhagen, pp.129–156.

Jokinen, K., Räihä, K.J., Paggio, P. and Jönsson, A. (2003) Multimodal Interaction – The MUMIN workshop and PhD course, in Henrik Holmboe (ed) *Nordisk Sprogteknologi 2003*. Museum Tusulanums Forlag. Copenhagen.

Petersen, B.S. and Paggio, P. (2002) A Danish Semantic Lexicon and its Application in Content-based Querying. *CST Working Papers*. No. 4, 25 p.

Paggio, P. (2001) Parsing in OntoQuery - Experiments with LKB. In *Proceedings of 1st International OntoQuery Workshop*. Kolding 2000, pp. 90–101.

Paggio, P. (2000) Relative Clauses. In Underwood, N. L. (ed.) *The LINDA Manual*. CST Working Papers No. 1, pp. 107–124.

Paggio, P. (1999) NP and Multimodality in the Staging project. In *Proceedings of the 1999 DALF meeting*. CST Working Papers No. 3. Copenhagen, pp. 63–70.

Paggio, P. (1998) Treatment of Grammatical Errors and Evaluation in SCARRIE. In *Proceedings of the 1998 DALF meeting*. Copenhagen, pp. 65–78.

Paggio, P. (1996) “Relative Clauses”, in *Linguistic Specifications, for Danish*. LINDA project report. CST, Copenhagen, pp. 103–126.

Maegaard, B. and Paggio, P. eds. (1996) Final Report. TEMAA project report. CST, Copenhagen.

Ørsnes, B. and Paggio, P. (1994) Maskinoversættelse af substantivkomposita, in Baron, I. (ed) *NORDLEX-Projektet. Sammensatte substantiver i dansk*, Institut for Datalingvistik, HHK, pp.135–157.

Non-published talks, workshops and demonstrations

Demonstration of MOSES project for ”Danmarks ErhvervsforskningsAkademi”.

“Informationsstruktur og pauser i dansk”, invited talk for the research group ProGram, University of Copenhagen. November 2005.

“The MUMIN annotation scheme for feedback, turn management and sequencing”, talk at NodaLida, Joensuu, Finland, April 2005.

“The information structure of Danish constructions”, poster at the Third International Conference on Construction Grammar (ICCG3), Marseilles, July 2004.

“Informationsstruktur i det danske sprog”, invited talk at the Copenhagen Linguistic Circle, 2004.

“Sprogteknologi og multimodale systemer”, invited talk at the conference Sprogteknologisk Forum, Copenhagen, March 13, 2002.

“Developing a Danish grammar in the GRASP project: A construction-based approach to topology and extraction in Danish”, talk given at *the 1st International Conference on Construction Grammar* together with Anne Neville, Berkeley, April 6–8, 2001.

“The extraposition of nominal clauses in Danish. Representational Issues”, HPSG workshop at the Copenhagen Business School, 1999.

Demonstration of the SCARRIE correction program at the conference *Sproget i informations-samfundet*, Copenhagen 1999.

Demonstrations of the Eurotra-DK system and talks on MT at various international meetings and conferences, i.a. *Coling '90* in Helsinki, the *Nordic Summer School in Machine Translation* in Göteborg, and the meetings of the Copenhagen Linguistic Circle, 1990–1997.