Seventh Conference
of the
European Chapter
of the
Association for
Computational Linguistics

March 27–31, 1995

University College Dublin

Belfield, Dublin, Ireland

Preface

This volume contains the papers prepared for the Seventh Conference of the European Chapter of the Association for Computational Linguistics, held 27–31 March 1993 in Dublin, Ireland.

The conference programme features invited talks, tutorials and submitted papers in the main session and in the student session.

The Programme Committee for the main session received a large number of submissions from all over the world. The general quality of the submissions was high. Out of a total of 153 submissions for the main session, 36 were accepted. Almost without exception, each paper was reviewed by three referees. We trust that the resulting programme offers an inspiring and representative sample of current research in both the theoretical and the applied aspects of the field.

Due to the volume of submissions, the reviewing process was both time–consuming and difficult. We would like to take this opportunity to thank the members of the program committee for taking on an average of 27 reviews and, in many cases, for contributing their own travel funds to attend the meeting of the program committee held at the University of Tübingen:

Harry Bunt, *Tilburg University*
Jan van Eijck, *CWI, Amsterdam*
Eva Ejerhed, *University of Umea*
Dale Gerdemann, *University of Tübingen*
Jan Hajic, *Universita Karlova, Prague*
Julia Hirschberg, *AT&T Bell Laboratories*
Ed Hovy, *USC Information Sciences Institute*
Judith Klavans, *Columbia University*
Lori Lamel, *LIMSI CNRS*
Mark Liberman, *University of Pennsylvania*

Tsuneko Nakazawa, *NTT Laboratories*
Philip Resnik, *Sun Microsystems Labs*
William Rounds, *CWI and Univ. of Michigan*
Louisa Sadler, *University of Essex*
Harold Somers, *UMIST, Manchester*
Oliviero Stock, *IRST, Trento*
Henry Thompson, *IRCS and Edinburgh*
Harald Trost, *Univ. of Vienna*
Annie Zaenen, *Rank Xerox Research, Meylan*

We would also like to thank those colleagues who helped with the reviewing process at the request of members of the program committee.

Many people contributed to the preparations of this meeting, and we are extremely grateful to them: Patrick Hanks, Stuart Shieber and Wolfgang Wahlster for the invited talks; Martin Kay, Shalom Lappin, and Alan Smeaton for giving tutorials; John Nerbonne for organising the tutorials; all ACL and EACL officials, for replying promptly to our frequent appeals for advice; Paul Jacobs and Steven Krauwer for providing reusable resources from previous ACL-affiliated meetings; Thorsten Brants for organising the student sessions. The student session is described in more detail in a separate preface (p. iv).

Special thanks are due to Allan Ramsay and his assistants in Dublin for the preparation of these proceedings and for handling the local organisation of the meeting.

Local help in Tübingen with preparing the papers for the committee and with mailings came from Iris Haberbosch, Katerina Magdou and Stephanie Schwarz.

The meeting in Dublin is the first (E)ACL conference which offers support for students from the Don and Betty Walker International Student Fund to help defray costs for attending the meeting. We would like to take this opportunity to thank all contributors to the fund for their generous help. Your contributions are a tribute to the inspiration that the field of computational linguistics has received from the many years of dedication by Betty Walker and the late Don Walker.

We dedicate these Proceedings to Betty and Don Walker.

Steven P. Abney and Erhard W. Hinrichs

Conference Chairs

Student Session

Preface

For the second time the European Conference of the ACL includes a student session. This part of the conference differs from the main conference in its emphasis on promising work in progress and provides an opportunity for student researchers to receive valuable feedback on their work from other members of the computational linguistics community.

We received 37 submissions for this student session, eight of which were accepted; one submission was accepted as a reserve paper. We allowed electronic submissions for the student session, and most of the authors (35 of 37) submitted electronically. Except for some minor formatting problems the committee had very good experiences with electronic reviewing and electronic discussion. We would like to thank the authors for their submissions. An equally important contribution was made by the student programme committee. Every paper was read by at least three of the 20 committee members. Thanks to their efforts, we were able to make a well-informed decision about which papers to select.

Also we would like to thank the organizers of the main conference, who supported us with their advice throughout the preparations of the session. We welcome their decision to enlarge the student part from six to eight papers, which still left a strong competition between the authors.

Thorsten Brants

Organization

Student Programme Committee

António H. Branco (Lisbon) Anne-Marie Mineur (Saarbrücken)

Thorsten Brants (Saarbrücken, *Chair*) Lineke Oppentocht (Leiden)

Luca Dini (Pisa) Susanne Riehemann (Stanford, CA)

Gregor Erbach (Saarbrücken) Alexander Rosen (Praha)

Francesc Ribas i Framis (Barcelona) Elina Savino (Bari)

Anette Frank (Stuttgart) Hana Skoumalova (Praha)

Zelal Güngördü (Edinburgh) Lena Strömbäck (Linköping)

Jussi Karlgren (Stockholm, *Co-Chair*) David Tugwell (Edinburgh)

Krister Linden (Helsinki) Kees Vermeulen (Utrecht)

Gemma Lyons (Dublin) Barbara Wesenick (München)

Tutorials

Tuesday, March 28

9:00   Tutorial Registration

9:30–12:30,   Martin Kay, Xerox PARC and Stanford
2:00–5:00   **Theory and Implementation of Finite-State Phonology**

This two-part tutorial presents a set of mathematical and computational tools for manipulating and reasoning about regular languages and regular relations and argues that they provide a solid basis for computational phonology. It shows in detail how this framework applies to ordered sets of context-sensitive rewriting rules and also to grammars in Koskenniemi's two-level formalism. This analysis provides a common representation of phonological constraints that supports efficient generation and recognition by a single interpreter.

      Recommended Reading: RONALD M. KAPLAN AND MARTIN KAY, "Regular Models of Phonological Rule Systems," *Computational Linguistics* 20(3), 1994, 331–378.

9:30–12:30   Alan Smeaton, Dublin City University
                 Natural Language Processing and Information Retrieval

The tutorial is aimed at an NLP audience who want to find about a possible application area for NLP resources and techniques ... information retrieval.

      What is information retrieval ... functionality, including document retrieval, filtering, and routing ... application areas, present and perceived. Conventional approaches to IR ... it is essential to see what other techniques have been used in IR to appreciate where NLP can/cannot be of use .... indexing techniques, boolean retrieval, vector space model, probabilistic modelling, term weighting and relevance feedback, query expansion, ranking. Storage structures ... this will be *very* short, just touching on the engineering problems of managing gigabytes of text.

      Three sections cover the guts of the NLP-IR intersection and will be peppered with illustrative examples.

- Lexical resources/morphology in indexing by word senses, base forms, grammatical categories, stemming.

- Syntax in indexing and matching, phrase identification and extraction, phrase normalisation and matching

- Semantics in indexing and matching ... KR formalisms, examples of systems/prototypes which use higher-order NLP

Issues of scale ... examples and illustrations of NLP-based IR (IR using NLP tools, techniques or resources) working on large scale collections, examples from TREC.

      Research issues and trends ... this section will be a discussion, led by me, on what I perceive as the directions in which the IR-NLP intersection will head.

2:00–5:00   Shalom Lappin, School for Oriental and African Languages, University of London
                 Computational Approaches to Ellipsis

The sentences in 1 illustrate three types of incomplete structures.

  1a. John read the paper before Bill did.

    b. Max gave flowers to Lucy, and chocolates too.

    c. No student arrived, except John.

1a is an instance of VP ellipsis, 1b is a case of "bare argument ellipsis", and 1c contains an exception phrase fragment. I will consider two possible approaches to developing a unified procedure for interpreting these distinct kinds of incomplete constituents. The first involves generating the semantic representation of an appropriate property or relation for the elided

constituent fragment. The second attempts to reconstruct a syntactic representation of a VP or sentence containing the fragment. I will consider each approach in some detail, and argue that neither the semantic nor the syntactic view can handle all three types of incomplete constituent.

I will provide motivation for the claim that VP ellipsis requires syntactic reconstruction, and that it is, in fact, a species of pseudo-gapping illustrated in 2.

2. John gave flowers to Lucy before he did chocolates to Rosa.

On this view, reconstruction is a relation between an elided VP and an equivalence class of lexically anchored syntactic structures which correspond to an antecedent VP. All elements of the equivalence class exhibit the same syntactic structure, but variation among corresponding lexical anchors with respect to a restricted set of specified features is possible. The syntactic structure of a (perhaps partially) elided VP is reconstructed by identifying its elided head with the head of an antecedent VP, and then specifying a correspondence among the arguments and adjuncts of the antecedent head on one hand and those of the elided head on the other. I discuss the algorithm for VP ellipsis resolution presented in Lappin and McCord (1990) and extended in McCord et al. as an implementation of this analysis.

I argue that, contrary to VP ellipsis, bare argument ellipsis must be resolved by means of a semantic procedure for predicate generation. I consider the higher-order unification analysis proposed in Dalrymple et al. (1991) as a possible account of this procedure.

Finally, I present arguments for treating an exception phrase fragment not as an instance of ellipsis, but as a displaced NP modifier. NP storage (Cooper (1993), Pereira (1990), and Pereira and Pollock (1991)) provides a suitable device for expressing the connection between a displaced exception phrase and the NP which it modifies.

The study of these three types of incomplete constituents indicates that the interpretation of ellipsis and constituent fragments is not a unified process. Each fragment type requires a different reconstruction procedure which operates at a distinct level of representation.

Background Reading: DALRYMPLE, M., S. SHIEBER, AND F. PEREIRA (1991), "Ellipsis and Higher-Order Unification", *Linguistics and Philosophy* 14, pp. 399-452. FIENGO, R. AND R. MAY (1994), *Indices and Identity,* MIT Press, Cambridge, MA. HAIK, I. (1987), "Bound Variables that Need to Be", *Linguistics and Philosophy* 11, pp. 503-530. HARDT, D. (1993), *Verb Phrase Ellipsis: Form, Meaning, and Processing,* unpublished Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA. LAPPIN, S. (1993A), "The Syntactic Basis of Ellipsis Resolution" in S. Berman and A. Hestvik (eds.), *Proceedings of the Stuttgart Ellipsis Workshop,* Arbeitspapiere des Sonderforschungsbereichs 340, Bericht Nr. 29-1992, University of Stuttgart, Stuttgart. LAPPIN, S. (1993B), "Ellipsis Resolution at S-Structure" in Amy Schafer (ed.), *Proceedings of NELS 23,* University of Massachusetts, Amherst, MA., pp. 255-269. LAPPIN, S. AND M. McCORD (1990), "Anaphora Resolution in Slot Grammar", *Computational Linguistics* 16, pp. 197-212. PEREIRA, F. (1990), "Categorial Semantics and Scoping", *Computational Linguistics* 16, pp. 1-10. REINHART, T. (1991), "Elliptic Conjunctions- Non-Quantificational QR" in A. Kasher (ed.), *The Chomskyan Turn,* Blackwell, Oxford, 360-384. SAG, I. (1976), *Deletion and Logical Form,* unpublished Ph.D. dissertation, MIT, Cambridge, MA. WEBBER, B. (1979), *A Formal Approach to Discourse Anaphora,* Garland Publishing Co., New York. WILLIAMS, E. (1977), "Discourse and Logical Form", *Linguistic Inquiry* 8, pp. 101-139.

Program

Wednesday, March 29

9:00–9:30	Registration
9:45–10:00	Opening remarks, welcome
10:00–11:00	**Stuart Shieber, invited speaker**
11:00–11:30	Coffee

REGULAR PAPERS | REGULAR PAPERS

11:30   Uwe Reyle
On Reasoning with Ambiguities

Andy Lauriston
Criteria for Measuring Term Recognition

12:00   Anette Frank, Uwe Reyle
Principle Based Semantics for HPSG

F. Wolinski, F. Vichot, B. Dillet
Automatic Processing of Proper Names in Texts

12:30–2:00   Lunch

REGULAR PAPERS | STUDENT SESSION

2:00   Mary Dalrymple, Andrew Kehler, John Lamping, Vijay Saraswat
The Semantics of Resource Sharing in Lexical-Functional Grammar

Kuang-hua Chen
Topic Identification in Discourse

2:30   Patrick Blackburn, Claire Gardent
A Specification Language for Lexical Functional Grammars

David Tugwell
A State-Transition Grammar for Data-Oriented Parsing

3:00   Jürgen Wedekind
Some Remarks on the Decidability of the Generation Problem in LFG- and PATR-Style Unification Grammars

Christopher C. Huckle
Grouping Words Using Statistical Context

3:30–4:00   Tea

REGULAR PAPERS | REGULAR PAPERS

4:00   Moon J. Kim, Young S. Han, Key-Sun Choi
Collocation Map for Overcoming Data Sparseness

Bill Keller, David Weir
A tractable extension of linear indexed grammar

4:30   Brett Kessler
Computational dialectology in Irish Gaelic

Chris Brew
Stochastic HPSG

5:00   Mark Davis, Ted Dunning, Bill Ogden
Text Alignment in the Real World: Improving Alignments of Noisy Translations Using Common Lexical Features, String Matching Strategies and N-Gram Comparisons

Annius V. Groenink
Literal Movement Grammars

Thursday, March 30

9:30–10:30   **Patrick Hanks, invited speaker**

10:30–11:00   Coffee

REGULAR PAPERS   REGULAR PAPERS

11:00	Gregory Grefenstette, Simone Teufel *Corpus-based Method for Automatic Identification of Support Verbs for Nominalizations*	David Milward *Incremental Interpretation of Categorial Grammar*
11:30	Rens Bod *The Problem of Computing the Most Probable Tree in Data-oriented Parsing and Stochastic Tree Grammars*	Mark Hepple *Mixing Modes of Linguistic Description in Categorial Grammar*
12:00	Francesc Ribas Framis *On Learning more Appropriate Selectional Restrictions*	Glyn Morrill *Higher-order Linear Logic Programming of Categorial Deduction*

12:30–2:00   Lunch

REGULAR PAPERS   REGULAR PAPERS

2:00	Hinrich Schuetze *Distributional Part-of-Speech Tagging*	Suresh Manandhar *Deterministic Consistency Checking of LP Constraints*
2:30	Jean-Pierre Chanod, Pasi Tapanainen *Tagging French – comparing a statistical and a constraint-based method*	Guido Minnen, Dale Gerdemann, Thilo Götz *Off-line optimization for Earley-style HPSG processing*
3:00	Atro Voutilainen *A syntax-based part-of-speech analyser*	Gregor Erbach *ProFIT: Prolog with Features, Inheritance and Templates*

3:30–4:00   Tea

4:00–5:00   EACL Business Meeting

Evening: Conference banquet

Friday, March 31

9:30–10:30   **Wolfgang Wahlster, invited speaker**

10:30–11:00   Coffee

REGULAR PAPERS	STUDENT SESSION
11:00 Jan Alexandersson, Elisabeth Maier, Norbert Reithinger *A Robust and Efficient Three-Layered Dialog Component for a Speech-to-Speech Translation System*	Pierre Sablayrolles *The Semantics of Motion*
11:30 Andrei Mikheev, Steven Finch *A Workbench for Acquisition of Ontological Knowledge from Natural Language*	Saliha Azzam *An Algorithm to Coordinate Anaphora Resolution and PPs*
12:00 David Carter *Rapid Development of Morphological Descriptions for Full Language Processing Systems*	Hercules Dalianis *Aggregation in the NL-generator of the Visual and Natural Language Specification Tool*

12:30–2:00   Lunch

REGULAR PAPERS	REGULAR PAPERS
2:00 Atro Voutilainen, Timo Jarvinen *Specifying a shallow grammatical representation for parsing purposes*	Richard Crouch *Ellipsis and Quantification: A Substitutional Approach*
2:30 Caroline Lyon, Bob Dickerson *A fast partial parse of natural language sentences using a connectionist method*	Michael Strube, Udo Hahn *ParseTalk about Sentence- and Text-level Anaphora*
3:00 Kong Joo Lee, Cheol Jung Kwon, Jungyun Seo, Gil Chang Kim *A Robust Parser Based on Syntactic Information*	Beryl Hoffman *Integrating "Free" Word Order Syntax and Information Structure*

3:30–4:00   Tea

REGULAR PAPERS	STUDENT SESSION
4:00 Janet Hitzeman, Marc Moens, Claire Grover *Algorithms for Analysing the Temporal Structure of Discourse*	Patrick Sturt *Incorporating "Unconscious Reanalysis" into an Incremental, Monotonic Parser*
4:30 Rani Nelken, Nissim Francez *Splitting the reference time: temporal anaphora and quantification in DR*	Tanya Bowden *Cooperative Error Handling and Shallow Processing*

Reserve Student Paper: Frank Keller
Towards an Account of Extraposition in HPSG

Table of Contents

Index of Authors