



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/81875/>

Version: Published Version

Article:

Atwell, ES (1987) Current European research in computational linguistics. AISBQ: Artificial Intelligence and Simulation of Behaviour Quarterly, 62. 22 - 24. ISSN: 0268-4179

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.



SUMMER 1987 No. 62

QUARTERLY NEWSLETTER OF THE SOCIETY FOR THE STUDY
OF ARTIFICIAL INTELLIGENCE & SIMULATION OF BEHAVIOUR

CURRENT EUROPEAN RESEARCH IN COMPUTATIONAL LINGUISTICS

A personal view of the Association
for Computational Linguistics
European Chapter Conference,
Copenhagen, April 1st-3rd 1987

*Eric Steven Atwell,
Department of Computer Studies,
Leeds University*

The Gods tried hard to prevent my attendance at the ACL European Chapter conference. Before leaving, my wife, my dog and I independently developed illnesses; I had to rush one to the vet just before leaving for the airport. I just made it into Copenhagen airport before atrocious weather closed it. My hotel was in the *Sailors' Entertainment* district; the receptionist suggested it would be safest to stay indoors! At least I would not be distracted from the serious business.

The Conference began with Fr. Hacker's paper, the Opening Address, and two Invited Papers from Laurence Danlos and Martin Kay. The research papers, starting in the afternoon, split into parallel sessions. My first April Fool's Day research paper was *How to Detect Grammatical Errors in a Text Without Parsing It* by yours truly in a stunning yellow outfit. This paper claimed that CLAWS, a system which computes a part-of-speech analysis without finding constituent-structure parse-trees, can be used to flag grammatical infelicities in English text.

Wolfgang Menzel's following paper discussed an intelligent tutoring system which detects and corrects student's responses to simple questions; this more 'traditional' AI system deals with a very restricted subset of English, whereas the former copes with Unrestricted English texts. Next came the first IBMer: Luis de Sopena presented the integrated environment for composing and proofreading Spanish documents developed at the IBM Madrid Scientific Centre, based on an analogous

English system from IBM Yorktown Heights. One point by is brought to mind recent debate in the AISBQ on ethical and moral issues: "Languages other than English have received little attention". Now there's a cause the egalitarian AISB should sponsor!

The next presentation, again by yours truly (co-authored with Nicos Drakos) was more theoretical, about pattern recognition techniques to extract a grammar from text. Then another IBMer, from Rome: Paolo D'Orta described their Italian speech recognition system, based on the English system built at Yorktown Heights by Fred Jelinek's R&D team. One beauty of their statistical approach is that the technique applies straightforwardly to any language. There was a sink at the front of the room; as Paolo spoke, a plumber casually walked up and started to unscrew it, but Paolo continued unwavering! Ute Ehrlich's following talk described the EVAR speech understanding system. It used 'conventional' pragmatic/semantic modelling, in contrast to the IBM system, which incorporates only word-trigram probabilities; and EVAR's 200 'concepts' contrasted with IBM's 6500-word dictionary. The final presentation of the day was due to be from Alex Martelli, another IBMer from Rome, but he didn't show; perhaps even IBM has a limited Conference budget?

The next morning, Mary McGee Wood opened with a paper on dictionary organisation. Next came a couple of papers demonstrating that Computational Linguistics thrives behind the Iron Curtain, from Petr Sgall and Eva Hajicova from Prague, Czechoslovakia. Petr expounded on dependency grammar. From Eva's title, *Fail-soft (emergency) Measures in a Production-oriented MT System*, I was half-expecting details of how the Czechs plan to avoid disaster in the event of fire in the University computing centre. In fact, "emergency measures" turned out to be things like default subroutines for coping with input words not found in the dictionary.

Next, Henrik Rue presented a Danish field grammar implemented in Borland Turbo-Prolog. Henrik ended his talk with a few 'inconclusions'; I wish more researchers could be as realistic. Thierry Guillotin's talk on Unification Categorical grammar also described a theoretical model rather than an application. The day closed with some papers peripheral to computational linguistics, by Dagmar Schmauks, Jean-Philippe Solvay, and Marie Bienkowski.

In the evening there were various demos. I thought that Mary Neff's presentation of the Wordsmith system (developed at Yorktown Heights, but mounted on IBM Denmark's miniframe for the demo to cut the phone bill) was more impressive than many of the mainstream papers. This system incorporates several published dictionaries, and Mary said they're still looking for more.

Bill Black opened next morning by discussing how we might automatically convert NL descriptions into conceptual data models. So far, his ideas are theoretical, not yet implemented; in contrast, the talk by Paola Velardi was an exposition of a large, sophisticated computer system for understanding text and converting sentences into an internal

knowledge representation formalism, developed at IBM Rome Scientific Centre. The system has been extensively tested on a corpus of Italian press releases, extracting from each article a pragmatic/semantic representation of the information encapsulated in the Italian text. We then returned to theoretical discussion: Ted Briscoe compared techniques for deterministic parsing.

During the coffee break, I was approached by a young lady from a glamorous international software house, who said her Managing Director wanted to meet me over lunch to discuss commercial exploitation of my research. Thoughts of jetting to California distracted me during Kari Valkonen's description of a parser for Finnish and Mats Wiren's discussion of how to choose an efficient rule-invocation strategy in context-free chart parsing. Mats tested out eight different strategies with a small grammar of less than 40 rules, and found the difference in parsing speed between the fastest and slowest strategies was a factor of about four; I was left wondering if it was really worth all the effort, given that hardware power and speed are increasing rapidly all the time (IBM had just announced their PS/2 range).

My lunch meeting turned out to be a Danish Pastry and coffee; I guess I had to eat at least one while in Copenhagen! I got back to hear Jean-Louis Binot's description of a part-of-speech analyser for French, which in parts sounded uncannily reminiscent of the CLAWS part-of-speech analyser for English. Then it was time to leave for the airport; with all my problems waiting for me back home, I daren't risk missing my plane!

BOOKS RECEIVED

The following books have been received and are available for review:

von der Lieth Gardner, A. "An Artificial Intelligence Approach to Legal Reasoning" MIT Press 1987

Agha, G. "Actors: a Model of Concurrent Computation in Distributed Systems" MIT Press 1987

Barton, G.E., Berwick, R.C. and Ristad, E.S. "Computational Complexity and Natural Language" MIT Press 1987

Mulmuley, K. "Full Abstraction and Semantic Equivalence" MIT Press 1987