

**O**f Joseph Whene ye gettes Euell. Do hi be kepte in psonis  
if he had lured othete ihu  
Woye Woye

And al manged  
per sent  
Woye ny  
And oye  
Whene  
Allo  
But  
ffor  
As  
he  
as  
ye  
ho  
pe  
ye  
pi p  
An  
App  
Com  
pze  
Wip  
in a  
Enc  
of  
Aud  
At  
An  
To  
An  
y  
ffo  
aya  
Di  
he

SLJN

NEWSLETTER

no. 18 June 1998

## Contents:

1. SLIN and HEL Conferences and Seminars
2. Conference Reports: SLIN Seminar (R.Dury), Historical Corpora Workshop and ICAME 19 (R.Facchinetti)
3. Papers read at the Rome Seminar (M.Dossena, R.Facchinetti, R.Dury)
4. A SLIN web site
5. A list of subscribers' electronic mail addresses

Edited by Prof. Nicola Pantaleo. Institute of English. Faculty of Foreign Languages and Literatures. University of Bari. Via Garruba 6/b - 70122 BARI

Tel. 080/5616520 - 080/5717426 Fax 080/5717459 - 5717515

E-mail: n.pantaleo@lingue.uniba.it

Co-edited and printed by Dr. Gabriele Rinelli

All materials to be published (reports, reviews, short articles etc.) should be sent (possibly with a disquette) to: Gabriele Rinelli - Via Fratelli Cervi, 22 - 72100

BRINDISI - Tel. 0831/517387 - 583630 - e-mail: naogao@tin.it

Printed by Word & Image Processing - Tel. 080/5576003

## 1. SLIN and HEL Conferences and Seminars

§ The SLIN Biennial Seminar on "Corpus Material from Old English to Present English" seems to have kept the promises of the vigil: it was at once a kind of 'making the point' of the respective experiences and a further step in the ideal march for an apt updating of research methodologies and instruments. It was both informative and inspiringly projectual, a nice cocktail of friendly exchanges and competence sharing. My irrecoverable optimism will certainly be balanced by colleagues' more sober private evaluations and also by Richard's report on p. 25. Thanks must be given to our host David Hart and to the speakers Marina Dossena, Richard Dury, Roberta Facchinetti, Maurizio Gotti - our formidable Bergamo stronghold - Gabriella Mazzon, Maria Teresa Zagrebelski. It was decided, given the special interest of all talks, to devote most of this, and a large part of the next issue of NL to publishing a substantially faithful version of the papers read at the Worksop (see section 3 below).

Incidentally, we wish to express our congratulations to our colleagues and contributors Antonio Bertacca, Richard Dury and Gabriella Mazzon, who have lately become *professori associati* of English language after passing a severe national competition, and formulate a hope for their prompt return to their main scientific and disciplinary collocation.

§§ The 9th SLIN National Conference, as already anticipated, will be held in Naples, at Suor Orsola

Benincasa and Federico II University, on **May 13-14** and dedicated to Thomas Frank who will be commemorated in his University on the tenth sad anniversary of his death. It was also agreed during the workshop which concluded the Rome Seminar that the Conference heading should be English Historical Pragmatics, a study field in interesting progress. More details including subtopic specification, bibliographical information, host speaker designation, accommodation facilities as well as a provisional programme will be provided in the next issue of *NL*.

§§§ **10/ICEHL** and **XXXI SLE Meeting** are going to be celebrated on 21-26 and 26-30 in Manchester and Saint Andrews respectively. Reports on both Conferences are expected to find room in the *NL* provided volunteer note-takers will kindly offer.

## 2. Conference Reports.

### § **STORIA DELLA LINGUA INGLESE - THIRD BIENNIAL MEETING**

The Storia della Lingua Inglese 'club' met in Rome for the third biennial meeting on 8<sup>th</sup> and 9<sup>th</sup> May 1998. The inverted commas in the previous sentence are used because we have yet to give ourselves a name. Of course, a name is not essential to a flourishing existence: Japan was an empire without a name for many centuries, and religion permeated society in India for an equally long time though its languages had no word for 'religion'. The suggested 'club' does not perhaps sound serious enough to be an official title, but it gives a very good idea of the unstructured and friendly nature of our meetings, experienced once again at this latest

edition of the (variously-named) 'national meeting', 'Seminar', 'Biennial Workshop' or 'Colloquium'.

It is an event now established as a welcome relaxing interlude from the busy world where we're whipped and whisked and whirled. We arrive in the magnificent city of Rome, see waving palm-trees and sun reflecting on the yellow-orange façades and this (together with the journey that has separated us from our normal worlds) is enough to put us in a good frame of mind. Then, on either side of a meal together, we meet for short presentations on teaching and research methodology, for computer demonstrations, for debates on the institutional status of our discipline and for the general exchange of information. It is an excellent formula and one can only hope it survives intact when the group and this event itself take on a stable name.

Proceedings this time (in the Dipartimento di Linguistica) were opened by an old friend, Maria Teresa Zagrebelsky, who reported on one result of her sabbatical last year during which she investigated various aspects of corpus linguistics: the recent International Corpus of Learner English. This new resource should be of great help in investigating the cause of learner errors and 'foreign-soundingness' and has already been the basis of several research papers and a laurea thesis.

New information-technology resources remained the theme with Roberta Facchinetti's report on a CD-ROM containing three eighteenth-century grammars, the ICAME Corpus collection (bringing together several important corpora, including historic corpora), and the 'W3-Corpora' Internet site that allows easy searching of corpus files.

Marina Dossena then gave us a guide through the 'Storia della Lingua Inglese' pages of the Bergamo Anglistica website and proposed that a site should be established for our 'club' (which, presumably having a name at the top of its page, will force us to definitely leave our not-unpleasant period of namelessness).

Proceedings were then adjourned until the following day and we divided briefly before meeting again around the tables of our usual trattoria (perhaps we could call ourselves a cenacle, a symposium?), where we had a pleasant meal before strolling back to our hotels through the warm Roman night.

The following morning's work opened with the present writer giving a report on project work with a 'Storia della Lingua Inglese' class: an outline of how projects substituted a part of the exam, the projects tried out, problems and recommendations for the future.

Gabriella Mazzon then presented software based on Access specially designed for helping in the work of searching corpora for items with variable spelling (in the specific research-project in hand, adverbs of negation) and presenting the results in a form that provides columns for further item-by-item analysis-notes. The item-by-item analysis is going to take some time, but hopefully at the end of it a button will be pressed and all will be light.

The last contribution to proceedings came from Maurizio Gotti who talked of an experiment in collaboration between English linguistics courses and Letteratura Inglese: the contribution from linguistics teachers to blocks of lessons in literature courses. This should strengthen links between

colleagues, make the position of the linguistics courses stronger and (by exam 'discounts') lighten the load of heavy Laurea study-programmes.

The business meeting led by Nicola Pantaleo concluded the morning and the event: among other matters there was the presentation of the next Convegno Nazionale (CNSLIN9, as I have suggested calling it for short) which Gabriella De Martino made us all look forward to by describing the place where the conference will take place and the social programme of visits to the monuments and museums of Naples. Back to the busy world we went, stimulated by our discussions, looking forward to meeting again in Naples.

**(Richard Dury)**

*§§ Historical Corpora Workshop* (Newcastle Co. Down, Northern Ireland, 19<sup>th</sup>-20<sup>th</sup> May 1998)

The two-day workshop was opened on 19<sup>th</sup> May in Belfast, at Queen's University. Suitably inebriated by the opening words of Prof. Leslie Clarkson, Senior Pro-Vice-Chancellor of the University, we all appreciated a talk on 'Drink and Drunkenness in Anglo-Saxon Writing', by the distinguished Anglo-Saxonist Hugh Magennis. His witty, learned references to the convivial and social values of drinking in the Anglo-Saxon world, followed by an equally convivial dinner in the heart of Belfast, were a perfect prelude to the happenings of the following day when the workshop was officially opened in Newcastle Co. Down, some 30 miles south-east of Belfast. The venue was the Slieve Donard Hotel, an enchanted castle-looking nineteenth century building, nestled at the foot of the spectacular, legendary Mourne Mountains, surrounded by the Donard forest and situated right on a sweeping five-mile sandy beach.

Never could a setting have been more appropriate to allow history to flood back with the breathing of the nearby sea, when Matti Rissanen (University of Helsinki) provided a detailed, comprehensive overview of the historical corpora currently available and of those in preparation (cfr. APPENDIX 1: English historical corpora in 1998: an overview). He was then followed by other representatives of 'the Helsinki team'; more specifically, Irma Taavitsainen provided a report of her work in progress on the *Corpus of Early English Medical Writing, 1375-1750*" (cfr. APPENDIX 2: the *Corpus of Early English Medical Writing*). In her talk ("Metatextual comments and the evolution of early medical writing"), she analysed metatextual passages, expressing the author's attitude towards the addressee or towards the subject matter. Such passages are used to clarify the author's intention and to direct the reaction of the addressee; consequently, they may lend a more personal tone to the texts and shed light on the evolution and change of style.

Jukka Keränen and Terttu Navalainen, both from the University of Helsinki, dealt with the *Corpus of Early English Correspondence*. Jukka ("CEEC and *ye shal find*") announced that the corpus is now complete; it covers the years 1417-1681 and in its full form it contains 2,710,306 words from 6,039 letters, written by 780 informants (women are responsible for one fifth of all letters). Although complete, the corpus is not ready for publishing, due to copyright issues. However, a sample will soon be available, possibly by the end of this year, covering 21 letter collections for a total of 500,000 words. Terttu ("The Helsinki Correspondence Corpus") illustrated the research she had developed on the corpus especially as regards the use of 'you', both singular and plural, the 3<sup>rd</sup> person singular ending "-th/-s", and multiple negation.

With the help of Joseph Schmied and Claudia Claridge - both from Technische Universität Chemnitz -, Lou Burnard ("The world is ruled and governed by opinion"), from the University of Oxford, described the background and development of the *Lampeter Corpus of Early Modern English Tracts* (<http://www.tu-chemnitz.de/~ehe/real/lamphome.htm>) covering the period 1640 to 1760 and a wide range of discourse types. As well as describing the corpus itself and its intended applications, Burnard focused on the procedures by which the corpus has been constructed and made conformant with the TEI Recommendations, and on the consequences of those procedures for re-usability and accessibility of the resulting resource.

Still focusing on the *Lampeter Corpus*, Claudia Claridge ("Early modern science: authorial and factual styles of writing") remarked that the emergence of the natural sciences - in a quasi-modern sense - was an important aspect of the cultural developments taking place in the 17<sup>th</sup> and 18<sup>th</sup> centuries - a fact which the *Lampeter Corpus* duly recognises by including 'science' as one of its text domains. The new 'natural philosophy', represented especially by the Royal Society, strongly advocated the empirical methods of using hypotheses, observations, and experiments in the acquisition of knowledge, thus emphasising the scientist himself and devaluating traditional 'book knowledge' to a certain degree.

Starting from this premiss, she presented the research project she is developing together with Joseph Schmied. They are using the *Lampeter Corpus* domain 'science' in order to look at the connection between this research approach and some linguistic features, or in other words linguistic representations of a more author-centred vs. a rather factual object-centred orientation. The features that have been chosen to be investigated are the different types

of conditional clauses as well as the occurrence of first and also second person pronouns.

Douglas Biber ("Dimensions of variation among 18<sup>th</sup>-century speech-based and written registers"), from Northern Arizona University, presented a multi-dimensional study of 18<sup>th</sup> century speech-based and written registers, based on an analysis of the texts of the *ARCHER* Corpus. Using factor analysis, he identified the underlying 'dimensions' of variation operative in the 18<sup>th</sup> century and used those dimensions to describe the similarities and differences among speech-based and written registers. This 18<sup>th</sup> century model of variation was compared to the multi-dimensional analysis of synchronic register variation, in order to highlight how the underlying multi-dimensional space of variation in English has evolved over the last three centuries.

Ann Curzan ("When // became all things: results of a Helsinki Corpus study on the rise of natural gender in English"), from the University of Michigan, dealt with the shift from a grammatical to a natural gender system in English and focused on the changes in anaphoric pronoun reference. Using the Old and Early Middle English parts of the Helsinki Corpus, she collected a comprehensive set of data on early English anaphoric pronouns. The results prove that early variation in the gender of anaphoric pronouns was highly patterned as this syntactic change diffused through the lexicon.

Christer Geisler ("A person to help you: postmodifying infinitives in a diachronic perspective"), from Uppsala University, first illustrated the *Upenn-Helsinki Corpus of Middle English Prose*, covering a total of 600,000 words, and highlighted how it has been tagged and parsed. Then he focused on infinitive constructions and showed that, judging from the data yielded by the corpus, the development of

such constructions may be linked to the development both of passive infinitives and of accusatives with infinitive constructions.

Gerry Knowles ("In the search of the origins of Northern English"), from the University of Lancaster, investigated the origins of the English language taking a point of view north of the Humber. He reported on a preliminary analysis of some 300 of the Middle English survey maps, paying particular attention to the role of river systems in the formation of dialects and their boundaries. This led to a straightforward explanation of prototypical Northern English as spoken north of a line from the Lune to the Humber.

Finally, Ann Curzan, from the University of Michigan, and Merja Kytö, from Uppsala University, led a session on "Tagging and parsing historical corpora", aimed at providing participants with the opportunity to share their experience in the creation and use of diachronic corpora. Starting from the *Lampeter Corpus* experience, the discussion mostly focused on ways in which complex tagging systems facilitate or hinder corpus searches for linguistic features. With regard to normalisation, the issues of abbreviation and spelling variation were also dealt with.

The whole workshop was closed by Matti Rissanen in late afternoon, to give way to the official opening of *ICAME 19-98, The 19<sup>th</sup> International Computer Archive of Modern and Medieval English Conference on English Language Research on Computerised Corpora*.

A brief note on *ICAME 19-98 conference* (Newcastle Co. Down, Northern Ireland, 20<sup>th</sup>-24<sup>th</sup> May 1998)

Although the conference was mostly synchronically geared, I believe a brief overview might be of value, in order to provide a snapshot of all the feverish work currently being

done in the field of corpus linguistics. More than a hundred distinguished academics came to Newcastle Co. Down from virtually all over the world. I will only name a few participants: Geoffrey Leech, from the University of Lancaster ("Corpus grammar and spoken English"), Geoff Barnbrook, from the University of Birmingham ("Saying what you mean: parsing Cobuild definitions"), Jim Cowie, Jeff Longwell and Charles Keller from New Mexico State University ("Using a large English corpus to refine a bilingual machine translation lexicon"), Graeme Kennedy, from Victoria University of Wellington ("The influence of Maori in the New Zealand English lexicon"), Junusaku Nakamura, from the University of Tokushima ("Semantic universe of the LOB corpus: structure of the corpus based upon the distributions of verbs, nouns and adjectives"), Yibin Ni, from the National University of Singapore ("Annotating nominal relations in English discourse"), myself, the writer of this report, from the University of Verona ("On the evolving *be going to* future in British English"), Sylviane Granger, from the University of Louvain-La-Neuve ("Optimising measures of lexical variation in EFL learner corpora"), Natalia Gvishiani, from Moscow State University ("Modals disjuncts in and through learner corpora"), Hans Lindquist, from Växjö University College ("A little more curvy at the corners: syntactic and contextual factors influencing the composition of disyllabic adjectives"), and of course the highly efficient 'Helsinki team' with a number of papers and posters. Indeed, at the conference a whole cascade of 68 papers and 12 posters were presented and discussed, each basically falling into one of the following fields:

- British National Corpus
- COBUILD corpus
- newspaper corpora

- learner corpora
  - dialectology
  - spoken corpora and discourse features
  - working on-line
  - software demonstrations
  - tagging and parsing procedures
  - International Corpus of English and ICECUP software
- Of the highly interesting aspects which emerged from the conference, I would simply like to point out Knut Hofland's illustration of the new *ICAME Collection of English Language Corpora on CD-ROM*, which is scheduled to be published in August. Among other synchronic corpora, the new CD-ROM will include the *Helsinki Corpus of Older Scots*, the *Newdigate Newsletters*, the *Lumpeter Corpus*, and, possibly, even a sample of the *Corpus of Early English Correspondence*. Details of most of the corpora included in the new CD-ROM are available at <http://www.hit.uib.no/corpora.html>.

§§§ *ICAME 19-98* conference was intensive and highly productive, briefly interrupted only in the afternoon of May 23<sup>rd</sup> by the joyful toast to the results of the much awaited referendum. An apt, general conclusion of both *Historical Corpora Workshop* and *ICAME 19-98* conference was the final trip held on Sunday 24<sup>th</sup> May to the Giant's Causeway, thousands of strangely symmetrical basalt columns which jut out to sea between Port Ganny and Port Noffer. Legend has it that the whole Causeway was built by the Irish giant Finn MacCool, so as to travel dry-shod across the sea to Scotland. Be it legend or not, when the 'digital people' of the conference walked along the Causeway Coast Path and gaped at that spectacular scenery, it looked as if two historically far apart worlds had joined naturally along the boundless bridge of mind.

**Appendix 1: English Historical corpora in 1998; an overview (Matti Rissanen, University of Helsinki)**

1. *The Dictionary of Old English Corpus in Electronic Form*

<http://www.press.umich/webhome/healey/siteform.html>

Annual subscription for sites. U.S.\$200

2. *Corpus of Middle English Prose and Verse* (Humanities Texts Initiative at the University of Michigan)

<http://www.hti.umich.edu/english/mideng/>

3. *The Middle English Collection* (Electronic Text Centre, University of Virginia)

<http://ctext.virginia.edu/mideng/browse.html>

4. *Public Domain Modern English Collection* (Michigan, see 2 above)

<http://www.hti.umich.edu/english/pd-modeng>

Some early items picked up from the Collection:

Gammer Gurton's *Needle*

Bacon, Francis, *New Atlantis*

Behn, Aphra, *The City Heiress*

Bourchier, John (Lord Berners), *Froissart's Chronicles*

Elyot, Thomas, *The Governor*

Jonson, Ben, *Various plays*

Marvell, Andrew, *Miscellaneous Poems*

Milton, John, *Paradise Lost*

Spenser, Edmund, *Amoretti and Epithalamion*

Wollstonecraft, Mary, *Maria, or The Wrongs of Woman*

5. *Early Modern English Dictionaries Database Search Utility* (12 monolingual and bilingual dictionaries)

<http://chass.utoronto.ca/english/emedd.html>

6. *ICAME CD-ROM* (Norwegian Computing Centre for the Humanities)

<http://www.hd.uib.no/>

7. *Wife of Bath's Prologue on CD-ROM* (The Canterbury Tales Project)

<http://www.shaf.ac.uk/uni/projects/ctp/Main/purchase.html>

(*The General Prologue* should be published by the end of this year)

8. *Electronic Beowulf*

<http://www.uky.edu/~kiernan/BL/kportico.html>

(to be completed by this summer)

9. *Corpus of Early Middle English Tagged Texts and Maps*

Margaret Laing (M.Laing@ed.ac.uk)

in preparation

10. *Edinburgh Corpus of Older Scots*

Keith Williamson (I.K.Williamson@ed.ac.uk)

in preparation

11. *Zurich English Newspapers Corpus (ZEN)*

Udo Fries (ufries@es.unizh.ch)

12. *Corpus of Irish English*

Raymond Hickey (r.hickey@uni-essen.de)

(The Middle English sample is completed)

13. *Brooklyn-Geneva-Amsterdam-Helsinki Parsed Corpus of Old English*

Susan Pintzuk (SP20@york.ac.uk)

14. *Penn-Helsinki Parsed Corpus of Middle English*

<http://www.ling.upenn.edu/mideng/>

registration: kroch@ling.cis.upenn.edu

**Appendix 2: The Corpus of Early English Medical Writing (Irma Taavitsainen and Päivi Pahta, University of Helsinki)**

• *The aim*

The purpose of the corpus is to provide a computer-readable database for the research project “Scientific thought-styles: Early English medical writing.” The aim of the project is to describe and discuss the causes of stylistic change in medical English in a socio-historical framework. The aim is to see how the underlying ideology, i.e. scientific way of thinking, is reflected on the evolution of scientific writing, how the methodology is reflected on the language use, and how the style of writing is determined by the authors’ education and various levels and scope of audience.

• *Method of analysis*

The analysis combines a quantitative approach to discourse analysis in the variationist frame with methods of qualitative studies. The intention is to construct synchronic descriptions of genre and subgenre styles within medical writing and combine them for a diachronic line of evolution, taking into account the socio-historical context.

• *Materials and period division*

The corpus consists of medical treatises between 1375 and 1750. The first part covers the period from the appearance of vernacular medical writing in the late Middle Ages to the dawn of the new science (1375-1550), and it is tentatively divided into two subperiods according to the introduction of printing (1375-1475 and 1475-1550). In the second main period (1550-1750) medieval conventions were still present, but started to be replaced by new patterns of thought and new empirical methods. The foundation of the Royal Society in 1660 marks a dividing line in this period.

0 20th century	D Dialectology	PM Postmodernism
1 7th-11th century	DA Discourse analysis	PO Poetry
2 11th-14th century	FSP English for specific purposes	PR Pragmatics
3 14th-16th century	FA Fantastic	PRO Prose
6 16th century	FS Film studies	PS Psychoanalysis
7 17th century	FT Folk/fairy tales	PSL Psycholinguistics
8 18th century	GE Gender	R Romanticism
9 19th century	GR Grammar	RII Rhetoric
AB Autobiography	H Honour	RT Reading theory
AE Acquisition of English	HE History of English	SAF South African
AF African	HI History of ideas	SC Scottish
AIA Associazione Italiana di Anglistica	IN Indian	SE Semantics
AL Applied linguistics	IR Irish	SEM Semiotics
AM American	IS Image studies	SF Science fiction
AU Australian	LE Language education	SH Shakespeare
B Bibliographic technique	LG Linguistics	SL Sociolinguistics
	LGH History of linguistics	SLF Sociology of literature

<b>BH</b> History of the book	<b>LP</b> Literature & politics	<b>SP</b> Espionage
<b>BI</b> Biography	<b>LSC</b> Literature & science	<b>SS</b> Short story
<b>CA</b> Caribbean	<b>LX</b> Lexicology	<b>ST</b> Stylistics
<b>CAL</b> Computer-aided language learning	<b>M</b> Modernism	<b>SV</b> Syntax
<b>CE</b> Contemporary English	<b>ME</b> Media	<b>TE</b> Teacher education
<b>CH</b> Critical history	<b>MT</b> Metaphor	<b>TH</b> Theatre
<b>CL</b> Children's literature	<b>MU</b> Music	<b>TL</b> Travel literature
<b>CM</b> Computational linguistics	<b>NO</b> Novel	<b>TM</b> Teaching methods
<b>CN</b> Canadian	<b>NT</b> Narrative theory	<b>TR</b> Translation
<b>CO</b> Comparative	<b>NZ</b> New Zealand	<b>TV</b> Television
<b>CR</b> Crime	<b>OR</b> Orality	<b>U</b> Utopias
<b>CS</b> Cultural studies	<b>P</b> Philosophy	<b>VA</b> Visual arts
<b>CT</b> Critical theory	<b>PA</b> Paleography	<b>WE</b> Welsh
<b>CV</b> Civilisation	<b>PC</b> Popular culture	<b>WS</b> Women's studies
<b>CW</b> Commonwealth	<b>PH</b> Phonetics	
	<b>PL</b> Philology	

- *The present size of the corpus*

The present size of the corpus is over 800,000 words. Shorter texts are included *in toto*, and longer treatises are represented by extracts of at least 10,000 words.

(Roberta Facchinetti)

### 3. Papers read at the Rome Seminar

#### § Marina Dossena: WHAT'S IN A SITE?

Readers of this Newsletter were already aware of the existence of the new *Storia della Lingua Inglese* and *Lingua Inglese* sites at the University of Bergamo, because their launch had been announced in a previous issue (no. 16, November 1997); many of the aforementioned readers had also accessed the sites on an individual basis and thus had gained an idea of what was available; but what was still missing was an overview that allowed the webmasters to discuss their pages with actual and potential users, thus including those colleagues who still hadn't accessed the new site. We thought the biennial Rome Colloquium would be an excellent opportunity for this sort of exchange, so thanks are due to David Hart for allotting time in the programme to this project, and for securing the co-operation of a technician who made sure that everything would work smoothly - when there's any hardware or software around, the presence of a technician in the background is as valuable a Linus' blanket as any webmaster can wish for! Luckily, though, there was no need for emergency treatment of the connections, and the homepage of the Bergamo sites appeared as expected.

The first screen provides links to very detailed Erasmus pages (written by Richard Dury and greatly appreciated by students both in Bergamo and in the English

universities with which exchanges are organized) and links to the twin sites of *Storia della Lingua Inglese* (edited by Richard Dury) and for *Lingua Inglese* (edited by the present writer). The joint decision of having a common structure was aimed at securing that the contents were homogeneous, and in this respect the overall supervision of the official professor of *Storia della Lingua Inglese* at the University of Bergamo, prof. Maurizio Gotti, should also be acknowledged.

From a more technical point of view, another joint decision was to keep the use of graphics and other technological paraphernalia to a minimum, in order to guarantee speed of access and ease of information tracking; for instance, we decided not to use frames, which split the screen and constantly present the site's table of contents as these are scrolled, because frames are notoriously difficult to bookmark and the way in which they force browsers to reload often redundant information inevitably makes the whole process much slower. Our tables of contents, instead, are presented at the top of the screen, with quick links to specific paragraphs in the same file.

As is always the case when information is to be presented and distributed, the key issue was related to what should be selected and included; we thought that our sites would have three main groups of visitors: students, *tesisti*, and ourselves as researchers, so the main criterion that was adopted focused on usefulness, practicality and feasibility; hence the decision to include links to sites related to our current research interests, to British libraries and to libraries within travelling distance in Italy, and to conferences in Europe. For the sake of comprehensibility, the three levels of site use are presented in sequence.

First and foremost, information is provided on our courses, syllabi, exam dates, office hours, etc., intended as

an in-house service for our students. Then there is information concerning theses, which presents in a fairly systematic way a collection of points that is otherwise scattered in various sessions of chapter discussion with each *tesista*. The main elements here refer to recommendations on structure and layout, including a sample page, so that students may learn to present their work in a clear, legible way, with appropriate footnotes and a complete bibliography; suggestions are also provided on how to begin the research work and how to make the most of the library. The sample page had to be scanned in as an image, because the software we used does not cater for justified margins, so we could not insert it as a text file: if we had, the justified margins in the original would have been lost, whereas we wanted students to see what an actual page ought to look like.

Another key feature is the remarkable list of bibliographical resources available in our local library compiled by Richard Dury. Links to library catalogues are also a highlight of both sites, and reference to corpora, discussion lists and sites of related interest (different for *Storia della Lingua Inglese* and *Lingua Inglese*, given their different approach to the subject: diachronic in the former case and synchronic in the latter) may also be exploited for research purposes at a higher level. Among these we may refer to the following.

- diachronic linguistics: Oxford Text Archive, Lampeter Corpus of Early Modern English Tracts, Robert Cawdrey's *A Table Alphabetical* (1604), etc.;
- synchronic linguistics: British National Corpus - online searching, ICAME, Linguistic Data Consortium, COBUILD, etc.;
- journals: *International Journal of Corpus Linguistics*, *International Journal of Lexicography*, *Forum for*

*Modern Language Studies, Literary and Linguistic Computing, etc.*

Finally, we presented our links to forthcoming events in the Faculty, in Italy and abroad: for instance, we currently have links to TALC 98 (Oxford, Keble College, 24 - 27 July 1998), the 10th International Conference on English Historical Linguistics (Manchester, 21 - 26 August 1998) and the XI International SELIM Conference: Medieval English Language and Literature (Vigo, Spain, 24 - 26 September 1998). Suggestions for additions are always very welcome.

Throughout the presentation during the Rome Colloquium reference was made to ways in which users can refer from one page to another, go back to a previously visited site, and make a note of what is worth visiting again through the use of 'bookmarks' (electronic markers that allow users to list the sites that they find most useful or most interesting and access them without rewriting their often long and complex URL, i.e. electronic address). A brief explanation was also given on how to save information found in web sites through the use of the 'Save as...' option in the File menu of the browser: this allows users to download contents into a floppy disk or a directory in the hard disk and re-use them or print them at leisure; especially as regards printing, this option proves particularly convenient, because off-line printing is certainly much faster than on-line, but it may also be used to import information found in the Net into a word-processing file.

Unluckily there was no time to demonstrate how our web pages were actually prepared, i.e. to give any detailed information on the editing software that was used to write the contents and set up the links; however, this is quite self-explanatory, since icons, commands and shortcuts are very similar indeed to the ones commonly found in Word. The

only snag that potential site editors ought to bear in mind is the fact that it is very easy to forget the addition of links to back pages at the end of each page, and this might leave users not knowing what to do - unless they are already experienced enough to use the 'Back' icon of their browser, which in any case would force them to go back one page at a time, instead of going back straight to the homepage: something rather annoying and time-consuming. Another point to remember is that files should all be named in the same way (we found that the avoidance of capital letters was the safest choice) and filenames must be reused exactly in the original form when creating links: an impromptu capital letter, for instance, would cause the link to be unrecognizable. Apart from this, no other difficulties arose from a strictly technical point of view; even transferring the files to the University server by FTP (i.e., File Transfer Protocol) proved quite straightforward: this means that we can now update contents whenever it is necessary, without asking the technicians to do it for us, and we can actually use the web pages like an electronic notice-board. This may be a point worth emphasizing: the use of this medium ought to be as user-friendly as possible, both for users and for us as webmasters, and in order to achieve this we were very willing to sacrifice any hypertechnological frills that might have made the pages look more upmarket, but which did not add anything to its contents. It is also important to remember that a website is never 'finished' - it is bound to be constantly updated as new pieces of information are added and obsolete ones are deleted.

In the case of the Bergamo pages, this is particularly true because, as we said, these are the first sites on English diachronic and synchronic linguistics in Italy, so what was presented cannot be taken to be a final product, but as a stage in the process; besides, as readers will have seen

elsewhere in this Newsletter, a new site is about to be launched that will include information on the activities of English Linguistics scholars in Italy. Various links now belonging to the Bergamo site will therefore be moved into the new one, and those relating to research will obviously feature most prominently in a page that it is hoped will provide new resources for our scientific community.

In the meantime, comments and suggestions are most welcome on what has been developed so far; the address for both sites is: <<http://www.unibg.it/anglistica/anglist.htm>>

(Marina Dossena)

### §§ Roberta Facchinetti: Historical Corpora: Present and Future

In my talk I first illustrated the database *Landmarks in English Grammar: the Eighteenth Century*, published early this year by *The Survey of English Usage* of University College London. It is a collection of five grammars of English from the eighteenth century, selected by the compilers for their importance in the history of English grammar, for their contemporary influence, and for their influence on later writers:

- Charles Gildon and John Brightland's *A Grammar of the English Tongue* (1711);
- Joseph Priestley's *Rudiments of English Grammar* (1761);
- Robert Lowth's *Short Introduction to English Grammar* (1762);
- John Ash's *Grammatical Institutes* (the text was originally published in 1760, but in the 1763 fourth edition, reproduced in the CD-ROM, the author linked his work directly to Lowth's, under a new subtitle, *An Easy Introduction to Dr Lowth's English Grammar*);
- Lindley Murrays' *English Grammar* (1795).

Each text has been electronically scanned from facsimile editions produced by R.C. Alston in the series *English Linguistics 1500-1800*, published by the Scolar Press. The compilers have added an index to each grammar book, including grammatical terms as well as references to cited authors - Shakespeare, Milton, Pope, Dryden, Swift, and Jonson among others - and containing hypertext links to the relevant sections in the grammars. A contents list has also been compiled for each text, using the original section and chapter headings of the books. Finally, a number of topics have been indexed, such as 'language and thought', 'French language', and 'English Academy'.

All texts can be analysed by means of *Adobe Acrobat Reader* v. 3.01 for Windows. Thanks to this search software, one can either simply read the text, skip from one page to the other, or find a word which has been indexed by the compilers of the database. Researchers can perform full-text searches across the five grammars, expand or limit their searches by selecting a specified index and defining a search query:

- *wordstemming*: allows one to find words with similar stems (e.g. 'building' and 'builder')
- *sounds-like*: finds words of similar sounds (e.g. 'Smith' and 'Smyth')
- *thesaurus*: finds words of similar meaning ('nouns' and 'substantives').

Researchers can increase the screen magnification in order to view even the smallest details of the optically scanned pages, which can also be copied to the clipboard and printed. Finally, a select bibliography, which is highly valuable for researchers and undergraduates writing their theses, has also been included.

While showing this database, mention has been given to *Project Gutenberg* (<http://promo.net/pg/>), started in 1971,

whose eventual goal is to complete - by the end of the year 2001, the 10,000 book Electronic Public Library, providing Etext editions of books a short time after they enter Public Domain. Currently, more than one thousand texts, mostly English, are included, the majority of which date back to 18<sup>th</sup> and 19<sup>th</sup> centuries. Here is a small selection of the writers whose texts have been included so far:

- |                       |                              |
|-----------------------|------------------------------|
| • Benjamin Franklin   | • Niccolò Machiavelli        |
| • Charles Darwin      | • Plato                      |
| • Chrétien De Troyes  | • François Rabelais          |
| • Christopher Marlowe | • William Shakespeare        |
| • Winston Churchill   | • Benedict Spinoza           |
| • Daniel Defoe        | • Thomas Carlyle             |
| • Charles Dickens     | • Thomas Malory              |
| • Edgar Allan Poe     | • Virgil (English and Latin) |
|                       | • Walter Raleigh             |

Of interest might be the following texts as well:

- *95 Theses*, by Martin Luther (in English and Latin)
- *An Open Letter on Translating*, by Martin Luther
- *Gettysburg Address*, by Abram Lincoln
- *La Divina Commedia* (in English and Italian)
- *The Anglo-Saxon Chronicle*, transl. By James Ingram
- *The United States Presidents' Inaugural Speeches*
- *The United States' Bill of Rights*
- *The United States' Declaration of Independence*

All the over one thousand texts can be freely downloaded and are even searchable on the net by means of the *World Wide Web Access to Corpora*, provided by the Department of Language and Linguistics of the University of Essex. The Internet search facility is not a permanent service yet, since it was first launched two years ago as a project and is still under evaluation. I have personally

evaluated it with my students at Verona (degrees in Scienze dell'Educazione and Lettere) and have found it of great value. The service aims at providing free access to existing linguistic corpora via the World Wide Web to students and researchers in Linguistics and related disciplines; no specific ability is required, the user only needs access to the WWW and can perform corpus searches using a web browsing interface (such as Netscape, Lynux, Internet Explorer, etc.). The user-friendly search software allows one to search and view in collocation a specific word or expression (*exact match*), or a pattern in the middle, at the end or at the beginning of a word. More specific details can be obtained when visiting the project site at <http://clwww.essex.ac.uk/w3c/>.

In the final part of my talk I anticipated the historical corpora due for completion this year and have specifically dealt with the new edition of ICAME CD-ROM. This and other corpora under way were subsequently presented at the *Historical Corpora Workshop* (19<sup>th</sup>-20<sup>th</sup> May 1998) and at *ICAME 19-98, the 19<sup>th</sup> International Computer Archive of Modern and Medieval English Conference on English Language Research on Computerised Corpora* (20<sup>th</sup>-24<sup>th</sup> May), organised by the Queen's University of Belfast (see report above).

(Roberta Facchinetti)

### §§§ Richard Dury: EXPERIMENTS WITH SLIN PROJECT WORK

#### *Bari and Bergamo experiments*

The group work (from 1994/95) on which Nicola Pantaleo reported at the Spring Seminar of 1996 inspired me to try out a single project at the end of the 96/7 academic year, continued this year with projects substituting the *corsa istituzionale* part of the exam.

The Bari and Bergamo experiments differed in the following ways: (i) numbers (four groups of 8-10 in a class of 25-30 in Bari vs. 4 pairs in a class of 8 regular attenders in Bergamo); (ii) time devoted to projects (c. 40% of the course covered by group projects in Bari, about 12% of the course (25% of the *istituzionale*) in Bergamo); (iii) the type of projects (Bari projects were linguistic analyses of passages from the same text, while Bergamo projects were a variety of different projects using OED on CD-ROM and Library resources; (iv) the final product (oral reporting and written report vs. just a written report in Bergamo); (v) the relationship to the exam (reports are the basis for the oral examination in Bari, while in Bergamo they substitute the *istituzionale* part of the oral examination).

### Projects

The homework sheets were inspired by exercises in John Algeo's *Problems in the Origin and Development of the English Language* (New York, Harcourt Brace Jonovitch, 3<sup>rd</sup> ed. 1982) - grouping together cognate forms in different European languages; supplying English and German cognates and deducing from these the rule of the 2<sup>nd</sup> consonant shift, and replies to a series of short questions.

The projects mainly involved the use of the OED on CD-ROM. I also gave students a bibliography of books in the Faculty Library that would give the most basic information for the introductory paragraphs. Here are the OED projects:

- (i) Words from Italian 1500-1599
- (ii) Adoption of Latin past participle in *-atus* > English *-ate*
- (iii) The suffixes *-dom*, *-hood*, *-th*
- (iv) Evolution of the meaning of words in the semantic area 'child'
- (v) Words borrowed from Scots into 'standard English'
- (vi) The lowering of the short vowel in words like *up*

However, in the face of some difficulty with the software, I also supplied some Library-based projects:

- (vii) The development of monolingual dictionaries in Britain and France
- (viii) Language academies
- (ix) Development of periphrastic 'do'

### Problems

In the 1997/8 class only one student had any familiarity with the computer so that basic skills down to the use of the mouse had to be learnt. They also found difficulty with the OED command language for the OED software, which is not simple. The result was that I had to invest more time than expected in basic instruction and help in making OED searches. In both classes I found that students are unfamiliar with the layout of a report, the need for uniformity of subtitles, the need to give each table a number and an explanatory title, etc.

It was not possible to teach students fully how to use the computer, how to use the software, how to do research in a library, how to present data and how to write a project report - as well as teaching History of the English Language! These are skills that should be learnt as a matter of course by students for all subjects, but it is unrealistic to expect them to be picked up in a short space of time.

Some of the projects themselves raised unsuspected problems: the loanword projects ran into all the limitations of the information in the OED, the Scottish loans project had to face the problems of the concept of 'standard English'.

### Conclusions and recommendations

(i) Instruction in the use of software takes time. Perhaps the solution would be to supply tables of data already obtained from corpora for the students' analysis and comment. This

material could be collected in a dispensa and re-used from year to year.

(ii) As well as corpora work students also need instruction and exercises in the skills of finding information in a Library.

(iii) Guidance in report-writing and data-presentation is necessary. Perhaps the most effective way of doing this is in a master:apprentice type of relationship, i.e. students produce a piece of work and then the teacher comments, corrects, suggests and asks for a revised version.

(iv) Projects should therefore also be presented to the class by those who have written them. This will also encourage the development of other useful skills.

Asked whether the project work was more of an investment of time than preparing for the exam, one group said it was about the same, others said more (but not vastly more). In compensation, the work involved the acquisition of other skills that were considered valuable.

(Richard Dury)

#### 4. A SLIN web site

I am very glad to submit the following project which was launched during the Rome Workshop and immediately (and very generously) "appropriated" by the Bergamo colleagues. So, I leave it to them to explain it and enclose the relative form which I eagerly recommend all Italian colleagues to fill and send in soon.

##### THE PLAN OF A NEW SLIN SITE

Soon after the Conference in Genoa we began to outline ways in which our two subjects, Storia della Lingua Inglese

and Lingua Inglese respectively, could appear in the Net, so that both students and researchers could find information quickly and easily.

The results of much brainstorming and several hours at the keyboard are on your screens and now, a full year later, we find ourselves thinking of a new SLIN site with a somewhat different perspective in mind. First of all, we have certainly learnt from our pioneer work (ours were the first English Linguistics sites in Italy), but, most importantly, this time our effort will be aimed at providing something that we, as a group, can share and the focus will be largely on research.

We owe this to the kind way in which our initiative was discussed at the latest Colloquium in Rome, and we certainly accept gladly our role as webmasters of the forthcoming site, because we are confident that all our colleagues will contribute with their ideas and suggestions, so here are a few notes to start from...

Even in the northbound train we began to think of what the new site might look like, and we thought that it should not be "truly one, but truly two" - in the sense that it should both present the SLIN group and its research and allow us to gather information as if it were a specialized search engine; in other words, we thought that it could be useful if the site included:

- a. information about the group, its conferences, colloquia, research interests, publications, etc.; this would mean a brief introduction (possibly by the editor of SLIN Newsletter?) and a directory of research interests and publications for each of us;
- b. conference announcements, links to similar groups, discussion lists, online corpora, libraries, etc.

In this sense the new site would both look to the world and mirror the group into the world, presenting its academic interests and allowing other researchers to contact us if they share the same interests (see form enclosed).

(Richard Dury & Marina Dossena)

## 5. A list of subscribers' electronic mail addresses

Here follows a list which I will be very glad to complete, amend and update with new entries as soon as information comes:

Bertacca Antonio	Lingue, Urbino	E-mail: lingue @ bib.uniurb.it
Brownlee Nicholas	L.inglese, Firenze	" doclin @ cesitl.unifi.it
Del Lungo Gabriella	Economia, Firenze	" dellungo @posta.cce.unifi.it
Di Martino Gabriella	Scienze pol., Napoli	" simonell @ cds.unina.it
Denton John	Lettere, Firenze	" doclin @ cesitl.unifi.it
Dossena Marina	Lingue, Bergamo	" marina @ unibg.it
Dury Richard	Economia, Brescia	" richard @ pop.spm.it
Fabbro Mariateresa	Statale, Milano	" feke2a @l fin.it
Facchinetti Roberta	Lettere, Verona	" faro @ chiostro.univr.it
Fodde Luisanna	Economia, Cagliari	" fodde @ vaxcal.unica.it
Gotti Maurizio	Lingue, Bergamo	" m.gotti @ mediacom.it
Haarman Louann	Dip.LLSM, Bologna	" haarman @ lingue.unibo.it
Iamartino Giovanni	Statale, Milano	" giamar @ tin.it
Iottini Laura	Scienze pol., Cagliari	" iottini @ unica.it
Lombardo Linda	Sociologia, Roma I	" lombardo @ uniroma1.it
Maggioni Maria Luisa	Cattolica, Milano	" dipling @ unicatmi.it
Nuccorini Stefania	Lettere, Roma III	" nuccorin @ uniroma3.it
Pantaleo Nicola	Lingue, Bari	" n.pantaleo@ lingue.uniba.it
Pinnavaia Laura	Cattolica, Milano	" allar @ tin.it