

ICAME Web-site

Knut Hofland

The Department of Culture, Language and Information Technology (Aksis)

Up-to-date information on the ICAME archive is found at:

<http://gandalf.aksis.uib.no/icame.html>

Here you will find samples and the order form for the ICAME CD, corpus manuals in electronic form, back issues of *ICAME Journal*, an archive of the corpora discussion list, the ICAME bibliography and many references to useful resources on the Web.

ICAME – CD-ROM Version 2

<http://gandalf.aksis.uib.no/icame/cd>

In addition to the five corpora from the first version, the second version of the ICAME CD-ROM contains an additional 15 corpora. The entire ICAME material is now collected on one single CD-ROM. We no longer distribute single corpora.

The total number of words for all 20 corpora amounts to 17 million (7.9 million words of written modern text, 2.5 million of transcribed speech and 6.6 million of historical text). 3.5 million words have been tagged for part of speech. The modern texts are from the sixties onwards.

The CD comes with a complete version of WordSmith. This program can be used both on the texts on the CD and on the customer's own texts. It can be used for generating and comparing word lists, making concordances and finding collocates and key words for a text.

16 corpora have been indexed with the DOS version of WordCruncher. The search program is included on the CD. WordCruncher can be used directly from the CD. Alternatively, the files can be copied to hard disk or network for quicker access.

One of the corpora, COLT, is also indexed with TACT, a DOS program included in the complete version on the CD.

All manuals are included on the CD. They are also available from the ICAME CD home page on the Internet. This homepage also includes an order form, text samples from the individual corpora, program screen display samples and other useful information.

The ICAME CD is available for non-commercial research. Customers will also receive access to some of the corpora via the Internet with a web-based interface.

Version 3 of the CD-ROM is planned for production in 2005.

Conditions on the use of ICAME corpus material

The following conditions govern the use of the corpus material distributed through ICAME:

1. No copies of corpora, or parts of corpora, are to be distributed under any circumstances without the written permission of ICAME.
2. Printouts of corpora, or parts thereof, are to be used for bona fide research of a non-profit nature. Holders of copies of corpora may not reproduce any texts, or parts of texts, for any purpose other than scholarly research without obtaining the written permission of the individual copyright holders, as listed in the manual or on the record sheet accompanying the corpus in question. (For material where there is no known copyright holder, the person(s) who originally prepared the material in computerized form will be regarded as the copyright holder(s).)
3. Commercial publishers and other non-academic organizations wishing to make use of part or all of a corpus or a printout thereof must obtain permission from all the individual copyright holders involved.
4. Publications making use of the material should include a reference to the relevant corpora, giving the name of the corpus and the distributor.

Use of ICAME texts within an institution

Although ICAME texts cannot be used and distributed outside the institution making the order, they can be freely used within the institution (department, faculty, university) for the purposes of research and teaching. To prevent any use of the material for commercial and profit-making purposes, it is advisable to limit access to registered computer users within the institution. The way this is done may vary depending upon the institution making the order.

Information for contributors

Language. All contributions should be in English. Contributors whose native language is not English should have their manuscripts gone through by a native speaker before submission.

Format. Contributions should be submitted as files on diskette or as attachments in e-mail, together with a printout made from your word-processing system. As regards other possible formats, consult the editors before submission of your manuscript.

Headings. The title of the paper should be followed by the author's name and academic affiliation. Sections and sub-sections should be numbered. Headings should **not** be singled out typographically (by boldface, capitalization, or the like).

Tables and figures should be numbered and titled (tables above and figures below). They should always be referred to by their number, **not** by expressions like "see the diagram below" or "in the following table". Tables and figures can be included in the main text file or submitted as a separate file (only black and white reproductions are possible). Drawings, graphs, and other illustrations must be reproducible originals.

Quotations. When quoting, use double quotation marks. Use single quotes *within* a quotation. Long quotations should be indented and given without quotation marks. Use single quotes for 'terms'.

Abbreviations. Periods should be used in abbreviations such as "e.g.", "i.e.", "cf."

Examples should normally be numbered and set apart from the text following standard linguistic practice. Short examples in the running text (words or phrases) should be in italics.

Notes should be placed at the end of the paper. References to notes in the text should be indicated as follows: 1, 2, etc. (in superscript numbers, by using the word-processor function).

References should conform to standard linguistic practice. References in the text should follow this pattern: Francis (1979: 110) defines a corpus as "a collection of texts assumed to be representative of a given language, dialect, or other subset of a language, to be used for linguistic analysis". The list of references at the end of the paper should be presented as shown by these examples:

Altenberg, Bengt. 1984. Causal linking in spoken and written English. *Studia Linguistica* 38: 20–69.

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University Press.

Renouf, Antoinette. 1987. Corpus development. In J. M. Sinclair (ed.). *Looking up: An account of the COBUILD Project in lexical computing*, 1–40. London & Glasgow: Collins ELT.

Tottie, Gunnel and Ingegerd Bäcklund (eds.). 1986. *English in speech and writing: A symposium* (Studia Anglistica Upsaliensia 60). Stockholm: Almqvist & Wiksell.

Authors should be given with their full first names, unless they always use the initials themselves.

Web addresses should follow the location format, e.g., <http://www.abo.fi/fak/hf/enge/icle.htm>, <http://www.TeleNex.hku.hk>.

Reviews. The heading of a review should contain the information shown in the following example:

Roger Garside, Geoffrey Leech, and Geoffrey Sampson (eds.). *The computational analysis of English: A corpus-based approach*. London: Longman, 1987. 196 pp. ISBN 0-582-29149-6. Reviewed by Gunnel Källgren, University of Stockholm.

Manuscripts for articles should be sent to one of the editors:.

Merja Kytö
Department of English
Uppsala University
P.O. Box 527
SE-751 20 Uppsala, Sweden
merja.kyto@engelska.uu.se

Anna-Brita Stenström
Nedanvägen 7
SE-291 35 Kristianstad, Sweden
ab.stenstrom@telia.com

Books for review and correspondence on reviews and abstracts should be sent to:

Jan Aarts
Department of Language and Speech
Erasmusplein 1
University of Nijmegen
NL-6525 HT Nijmegen, Holland
j.aarts@let.kun.nl

The editors are grateful for any information or documentation which is relevant to the field of concern of ICAME.