

Text Analysis Software and Corpus

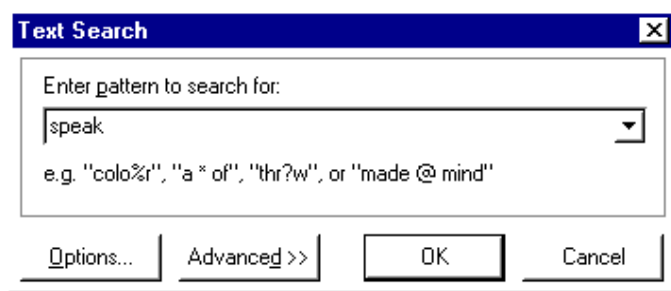
MP 2.2 -- MonoConc Pro 2.2

Corpus of Spoken Professional American English, 2 million words with part-of-speech tags.

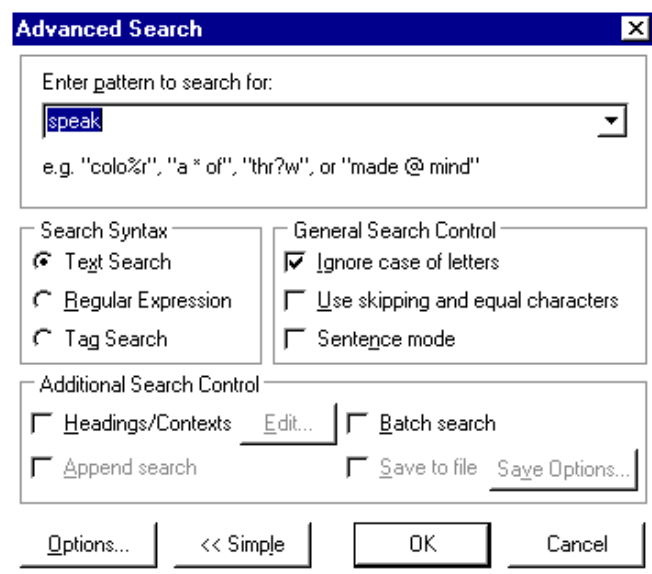
MonoConc Pro is widely used in universities and schools for teaching and research. The program has a well-designed user interface that allows novice users to load a text or corpus and start analysing language patterns with ease, and yet the software provides expert users with the powerful and configurable commands needed to work with large, complex, annotated texts.

Search

A corpus or text can be searched very quickly and easily. By choosing the basic **Search** command, the user can retrieve all examples of a word or phrase (or part of a word) from the corpus (one or more text files). Below, the search term **speak** has been entered. All instances of *speak* will be collected from the text-only files comprising the loaded corpus.

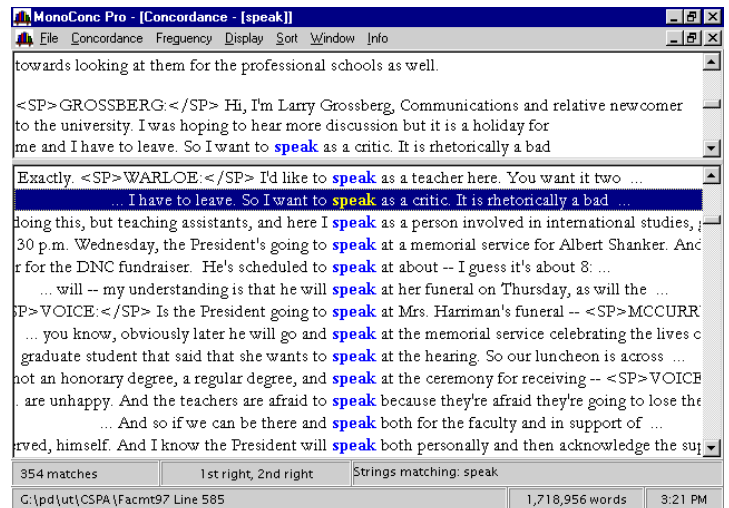


More complex search commands are also available when necessary, as the screen shot below reveals. Some of the possible search options are: Text search, Regular expression search, Tag (part-of-speech) search, Batch search, and various heading-sensitive and context-sensitive searches.



Results

The retrieved examples of the search term are typically displayed in a Keyword in Context format such that if the lines are sorted, patterns in the data can very easily be identified **visually**. (Once unwanted examples have been deleted, the results can be displayed and saved in a sentence format, with accompanying source file information if required.) Above the results window is a **context window** (reduced in size in the screen shot below). This context window provides up to 30K of context for each “hit.” Various kinds of information relating to the results are displayed in the lower panels of this window.



Various frequency calculations can be performed, including collocate frequency, collocation frequency, as well as word frequency for the entire corpus (in frequency order or alphabetical order).

MonoConc Pro will work with any size of corpus and with different languages. Various features of the software make it easy to work with both mark-up and part-of-speech tags.

MP 2.2 New Features: Frequent collocates are highlighted in colour in the concordance results window. Corpus frequency lists can be compared to identify overused and underused words. And the concordance results can be saved in the form of html files, which can then posted on the web or inserted into Word documents etc..

Pricing (Educational)

Single user: \$85 (Includes paperback manual, program disk and sample corpus disk.)

Site licence: \$550 (2-year licence for up to 15 users)

Athelstan

5925 Kirby Drive, Suite E 464
Houston, TX 77005
U.S.A.

Email: info@athel.com

Fax: (713) 524-1159

Web: www.athel.com

Corpus of Spoken Professional American English (CSPAЕ)

The *Corpus of Spoken Professional American English* includes transcripts of conversations of various types occurring between 1994 and 1998. The corpus consists primarily of short interchanges by approximately 400 speakers that are centred on professional activities broadly tied to academics and politics, including academic politics. The seventeen files comprising the corpus contain over 2 million words. For comparison the Brown Corpus and LOB Corpus, popular corpora from the sixties, each contain around 1 million words.

The corpus, which has been constructed from a selection of existing transcripts of interactions in professional settings, contains two sub-corpora of a million words each. One sub-corpus consists mainly of academic discussions such as faculty council meetings and committee meetings related to testing. The second sub-corpus contains transcripts of White House press conferences and is almost exclusively question-and-answer sessions. The corpus is suitable for lexical and grammatical investigations of spoken usage, but not for close discourse analysis since backchannel, overlap and pause length information is not available in the corpus.

Both tagged and untagged files are included in the corpus. MonoConc Pro will allow the tags to be suppressed for easier analysis once a search has been performed.

CSPAЕ is available for \$79 (\$295 for a site licence for up to 15 users.) The untagged version of the corpus is \$49. (\$179 for the site licence.)

Catalogue:

MonoConc Pro (MP 2.2)

Concordance/Text analysis software (Windows). Wordlists, Corpus comparison, Collocate frequency, Regex search. Concordance. \$85 (\$550 2-year Site Licence)

Collocate

Shows collocations along with statistical measures such as MI, log likelihood. N-grams. Extracts collocations. Requirements: Windows, \$45. (\$350 2-year Site Licence)

ParaConc

Multilingual Concordance/Text Analysis software that is similar to MP 2.2 but is designed to analyse translation texts. \$95 (\$750 2-year Site Licence)

Order form: Educational pricing

Athelstan, 5925 Kirby Drive, Ste. E 464, Houston TX 77005

Fax: (713) 524 1159

Web orders: www.athel.com

Guarantee. Items found to be unsatisfactory for any reason can be returned within 30 days for a full refund.

Please send the following:

___ MP 2.2	@ \$85/\$550*	___
___ Collocate	@ \$45/\$425*	___
___ ParaConc	@ \$95/\$750*	___
___ MonoConc Easy	@ \$45/\$290*	___
___ CSPAЕ (tagged + untagged)	@ \$79/\$295*	___
___ CSPAЕ (untagged)	@ \$49/\$179*	___
___ Learning with Corpora	@ \$24.95	___
___ Concordances in Classroom	@ \$21.95	___
___ Concordance pack II	@ \$159	___
___ Concordance pack III	@ \$179	___

Shipping and Handling: \$8

Total: _____

Ship/Bill to:

Payment:

P.O. Number _____ Cheque number _____

Credit card payment

Visa/MC/Amex Number _____

Exp. Date _____ Name _____

Credit Card Billing Zip _____

*Site licence price (2-year licence for up to 15 users)

MonoConc Easy

Concordance/Text analysis software for casual users, languages learners. Does not include corpus comparison, Advanced Sort etc.. Requirements: Windows. Wordlists, Collocate frequency, chunks. Concordance. \$45 (\$290 2-year Site Licence)

Learning with Corpora

Contains articles by a variety of authors on the use of corpora in linguistic research, language teaching and translation studies. Edited by Guy Aston. 2001. \$24.95

Concordances in the Classroom

Using concrete examples, the authors Chris Tribble and Glyn Jones show how a concordancer can be used to significantly enrich language and literature classes. 1997. \$21.95

Corpus of Spoken Professional American English

The basic, tagged version includes both tagged and untagged files. 2000. \$79 (\$295 Site licence). Untagged version. \$49 (Site Licence \$179)

Concordance pack II

MP 2.2, Corpus (CSPAЕ – both tagged and untagged files) and *Learning with Corpora* Edited by Guy Aston. \$159

Concordance pack III

MP 2.2, Collocate, and Corpus (CSPAЕ – both tagged and untagged files). \$179