



A Freeware, Multi-platform Corpus Analysis Toolkit for the Classroom

Laurence Anthony

Center for English Language Education in Science and Engineering (CELESE)
School of Science and Engineering
Waseda University

anthony@antlab.sci.waseda.ac.jp
<http://www.antlab.sci.waseda.ac.jp/>

IPCC '05 (July 13th, 2005)



Outline

- Background to AntConc
- Getting Started (Installation)
- Tools
 - KWIC Concordancing
 - Search Term Distribution Plot
 - File View / File Search
 - Word Clusters / N-grams
 - Word List / Keyword List
- Classroom Example
- Future Developments

-2-

Background

- **AntConc 1.0 created in 2002**
 - Designed for use in the 'Scientific & Technical Writing Program' at the Graduate School of Engineering, Osaka University, Japan
 - Commercial software is **expensive**
 - Commercial software is **designed for researchers**
 - Commercial software is **designed for Windows**
 - Developed using the PERL 5.8 programming language
- **Distributed as freeware for Windows and Linux via my website**

-3-



Background

- **Rapid development from 2002 ~**
 - 19 updates (3 major upgrades)
 - Additions based on requests, comments, and feedback from teachers/students in Germany, Netherlands, Spain, UK, US, ...
 - Bug fixes
 - GUI design improvements
 - New tools / features
- **AntConc 3.0.1 released in 2005**

-4-



Features

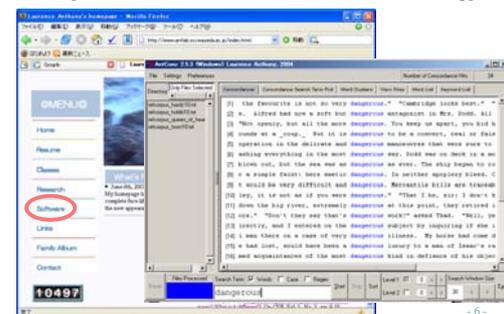
- **Freeware**
- **Multiplatform**
 - Windows 95/98 / NT / XP
 - Unix / Linux
- **Extensive set of text analysis tools**
 - KWIC Concordance
 - Search Term Distribution Plot
 - File View / File Search
 - Word Clusters / N-grams (Lexical Bundles)
 - Word lists
 - Keyword lists
- **Powerful Search Features**
 - Extensive Wildcards
 - Regular Expressions (REGEX)
- **Multiple-Level Sorting**
- **HTML/XML Tag Handling**
- **Unicode Support**
- **Learner/Classroom friendly GUI**

-5-



Getting Started (Installation)

- <http://www.antlab.sci.waseda.ac.jp/>



-6-



KWIC Concordance

File Window

Progress Report

Search Term

Search Options

Sort Levels

Save Current Window

Total Hits

KWIC

Hit File Name

Search Term Distribution Plot

Plot Results

Plot Statistics

Zoom Options

File View / File Search

Selected File

File View Results

Search Term

Search Options

Word Clusters / N-Grams

Cluster/Bundle Results

Search Term

Cluster/Bundle Parameters

Word List

Wordlist Options

Stoplist / Target Words

Keyword List

Keyword Results

Keyword List Options

Reference Corpus

Classroom Example



- **Appropriate language**
 - We **did** an experiment and...
 - We conducted **a lot of** experiments...
 - We obtained many **good** results...
- **Improvements?**
 - We **conducted** an experiment and...
 - We conducted **numerous/many** experiments...
 - **Our results show that...**
 - **Experimental results show that...**

- 13 -



Future Developments

- Improved stability (bug fixes)
- Improved XML / Unicode support
- Improved statistics data
- Improved documentation
- New web concordance feature

Your feedback is always welcome!

- 14 -



Acknowledgements

- Grant-in-aid for Scientific Research by the Japan Society for the Promotion of Education, Science, Sports and Culture, Japan (No. 16700573)
- Waseda University Grant for Special Research Projects, Japan (No. 2004B-861)

- 15 -



**A Freeware,
Multi-platform
Corpus Analysis
Toolkit for the
Classroom**

Laurence Anthony

Center for English Language Education in Science
and Engineering (CELESE)
School of Science and Engineering
Waseda University

anthony@antlab.sci.waseda.ac.jp
<http://www.antlab.sci.waseda.ac.jp/>

IPCC '05 (July 13th, 2005)