Addressing the Training Resource Deficit
Utilizing Open Corpora and OER

Alannah Fitzgerald

http://www.flickr.com/photos/opensourceway/5538036344
Overview

• Key concepts within data-driven language learning
  – Meaning, frequency, use, derivatives, collocations
• Training with corpora and text analysis tools
  – Open vs proprietary content & tools
• FLAX – the Flexible Language Acquisition project
  – Digital infrastructure and enhancement
  – Copyleft dances with copyright
• OpenSpires
  – Creative commons podcast content (OER)
  – iTunesU as an OER channel
• Further resources
Typical concordance interface
Definitions

Goalless

Adapted from Hoffmann et al., 2008
Where do you start?

goalless

adjective UK /ˈɡoʊl.əs/ US /ˈɡoʊl.əs/ /ˈgoʊl-/  

Definition  
without any goals being scored  
The match ended in a goalless draw.  

(Definition of goalless adjective from the Cambridge Advanced Learner's University Press)
goalless - definition of goalless by the Free Online Dictionary at www.thefreedictionary.com/goalless

adj. 1. having no points scored; - of games. ThesaurusLegend: Synonyms Related Words Antonyms. Adj. 1. goalless - having no points scores; "a scoreless ...

Goalless draw hurts Liverpool, Tottenham

IBNLive.com - 19 hours ago
We've come and got a point, kept a clean sheet and that's most thing." Goalless draw hurts Liverpool, Tottenham.

1126 related articles
The **FLAX** *(Flexible Language Acquisition)* project aims to automate the production and delivery of practice exercises for overseas students who are learning English. Exercise material comes from digital libraries, which can supply a virtually endless supply of collaborative and competitive language activities. *(All software produced by this project is open source, issued under the GNU General Public License.)*
<table>
<thead>
<tr>
<th>Research Development Demo collections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Web Phrases</td>
</tr>
<tr>
<td>Web Collocations</td>
</tr>
<tr>
<td>Wikipedia Collocations</td>
</tr>
<tr>
<td>Collocation Activities</td>
</tr>
<tr>
<td>The Best of Password (with more collocation activities)</td>
</tr>
</tbody>
</table>

- Other FLAX collections
- Text Checking

Ongoing research:
- Web Pronoun Phrases
- eJELTS
- Web Phrases
- Educational Leadership Abstracts
- Web Collocations
- Curriculum and Change Abstracts
FLAX - web phrases - single word search

Ongoing research:
- Web Pronoun Phrases
- eIFLTS
- Web Phrases
- Educational Leadership Abstracts
- Web Collocations
- Curriculum and Change

Type in your word(s) goalless (up to 4 words)

use * or - to find words inbetween, for example go * river
search for phrases following
return up to 100 phrases

group by word type on

show phrases

Goalless
Meaning + Usage + Frequency

- goalless + *Singular Noun* (28,000)
- goalless + *Preposition* (2,900)
- goalless + *Adverb* (2,100)
- goalless + *Plural Noun* (950)
- goalless + *Coordinating Conjunction* (560)
- goalless + *Adjective* (250)
- goalless + *Singular Proper Noun* (170)
- goalless + *Verb* (140)
- goalless + *Wh-adverb* (44)
- goalless + *To* (40)

*goalless* search in the FLAX Web Phrases collection
Derivatives (word families)

- economy
- economical
FLAX - web collocations

Ongoing research:
- Web Pronoun Phrases
- eIELTS
- Web Phrases
- Educational Leadership Abstracts
- Web Collocations
- Curriculum and Change Abstracts

Economy

Search
Compare

Type in your word: economy
return up to 50 collocations
with frequency cut-off 1000
grouping off
show collocations
FLAX - web collocations

related words: economic economical economically economics economies economist economists uneconomical

Which do you want:

economy used as Adjective

- economy + Noun: economy fancy dress

economy used as Noun

- Adjective + economy: the global economy
- Adjective + economy: economic economics economy

economy used as Noun: Adjective + economy

the global economy 170,000 the n
the local economy 66,000 the p
FLAX - sources

Web samples

Global Economic Prospects 2011: Washing developing countries is over. Efforts must now focus on problems, the Bretton Woods conference created... Data & Research > Prospects > Prospect should prepare for further downside risks, as...

About.com Economics

The Global Economy

The Global Economy

From U.S. Department of State,

See More About: international trade international finance

To help countries with unmanageable balance-of-payments conference created the International Monetary Fund (IMF) nations unable to meet their debts through conventional n
The UN General Assembly convened a special session from April 23 to May 1 to discuss ways of revitalizing economic growth and relieving external debts in the Third World. The changes in Eastern Europe were also raised in the context of the difficulties in managing the global economy.
FLAX - web collocations - activity

related words: economic economical economically economics economies economist economists uneconomical

Which do you want:

economy used as Adjective

- economy + Noun: economy fancy dress

economy used as Noun

- Adjective + economy: the global economy
- Adjective + economy: economic economics economy

Find...
- most common use
- most common collocation
The FLAX (Flexible Language Acquisition) project aims to automate the production and delivery of practice exercises for overseas students who are learning English. Exercise material comes from digital libraries, which can supply a virtually endless supply of collaborative and competitive language activities. (All software produced by this project is open source, issued under the GNU General Public License.)
Standard Collections

The Best of Password
Learning Collocations
British Academic Written English
Image Guessing activity demo collection

Learning Collocations

your name: 

BAWE
FLAX – Learning Collocations (BNC, BWE, Wikipedia)

Learning Collocations

ageing population in BNC

- The age structure of the population was changing.
- Figure 5.1 The age structure of the population of England and Wales, 1851-2021
- Figure 6.7 presents data for the age structure of the population in more detail.
- The most usual way of predicting future service needs is by the forward projection of current prevalence data to the predicted age structure of the population in the future.
- Bill said: Within ten to 20 years, changes in the age structure of the population, it is going to be vital to attract women to ensure that the complement of applicants is kept up.
FLAX – Learning Collocations
(related collocations, definitions, topics)

**related collocations**

<table>
<thead>
<tr>
<th>ageing</th>
<th>population</th>
<th>fertility</th>
<th>age</th>
<th>longevity</th>
<th>pension</th>
<th>country</th>
<th>increase</th>
</tr>
</thead>
<tbody>
<tr>
<td>health</td>
<td>overpopulated</td>
<td>decline</td>
<td>global</td>
<td>expenditure</td>
<td>elderly</td>
<td>rise</td>
<td>immigration</td>
</tr>
<tr>
<td>aging</td>
<td>reweighing</td>
<td>well-being</td>
<td>people</td>
<td>sustainability</td>
<td>demographic</td>
<td>world</td>
<td>sub-replacement</td>
</tr>
<tr>
<td>old</td>
<td>benefit</td>
<td>fast</td>
<td>proportion</td>
<td>migration</td>
<td>myth</td>
<td>counter</td>
<td>reduce</td>
</tr>
<tr>
<td>median</td>
<td>ratio</td>
<td>rate</td>
<td>develop</td>
<td>transition</td>
<td>future</td>
<td>experience</td>
<td>arise</td>
</tr>
</tbody>
</table>

**definitions**

Population ageing or population aging (see [English spelling differences](https://en.wikipedia.org/wiki/Ageing)) occurs when the median age of a country or region rises. [Wikipedia]

extended definitions from [wiktionary](https://en.wiktionary.org/wiki/Population_ageing)

**related topics in Wikipedia**

- Sub-replacement fertility
- Population pyramid
- C. D. Howe Institute
- Demographic transition
- Pension Benefit Guaranty Corporation
Brainstorm keywords from your subject/topic area and fill your cherry basket!

- They experience less significant *economies of scale*.
- Growth may provide *economies of scale* (decreases MC) and prevent take-over in some circumstances.
- 2, supposedly small inefficient farmers, would exit the industry and release land for larger farms achieving *economies of scale*.
- They have made the large initial investment to overcome the barriers of starting up, acquiring technical knowledge and gaining *economies of scale*.
- Acquiring O 2 gives Telefónica the *economies of scale* in the industry and enables them to retain and increase their supplier power.
- In economies with imperfectly competitive markets and *economies of scale*, the country's ability to grow can totally depend on access to a larger market.
Collocation guessing game in FLAX
What can you do with this?

Creative commons content
This is an example of how a podcast series can be combined with other freely available resources to enhance the user experience and offer additional support to non-native speakers and those from other subject areas. The resource offers links to the audio of eight conversations and transcriptions of these, as well as material to support the learning and understanding of the vocabulary involved.

All the material is being made available under the Creative Commons BY-NC-SA licence.

In this site

- Home
- OpenSpires Blog
- About
- Resources
- Events
- FAQs
- Contact

Links

- University of Oxford Podcasts
- JISC Open Content Strand
- Creative Commons
- Creative Commons for Learning
- JISC Steeple Project wiki

On this page:

- About the series
- Participants
- Content
- Links to the material

The Credit Crunch and Global Recession

About the series: A podcast series (eight talks) about the credit crunch and global recession featuring Oxford academics. This series will examine how the current crisis developed, analyse market and government responses to it, and look at what might happen next.

Participants:

- Dr. Linda Yueh, Fellow in Economics, University of Oxford.
- Prof. Jonathan Michie, Professor of Innovation and Knowledge Exchange, Director of the Department for Continuing Education, University of Oxford.
- Martin Slater, Fellow and Tutor in Economics, St Edmund Hall, University of Oxford.
- Dr. Outi Aarnio, College Lecturer in Economics, Lincoln College and St. Edmund Hall, University of Oxford.

http://http://openspires.oucs.ox.ac.uk/crunch/

Open Educational Resources (OER)
Combining open tools & open content
Creative Commons Licensing Suite

http://www.flickr.com/photos/peterleth/
What is OER?

OER Commons:

“Open Educational Resources are teaching and learning materials that you may freely use and reuse, without charge. OER often have a Creative Commons or GNU license that state specifically how the material may be used, reused, adapted, and shared.”

http://www.slideshare.net/tbirdcymru/itunes-u-corporate-channel-of-free-educational-resources
Creative Commons

1-51 of 51

1909 - The People's ... Oxford University

1909 - The People's ... Oxford University

The Beazley Arc... Oxford University

The Beazley Archive ... Oxford University

Censorship in Litera... Oxford University

The Credit Crunch a... Oxford University

Critical Reasoning f... Oxford University

Critical Reasoning f... Oxford University

Crossing Borders: H... Oxford University

Crossing Borders: H... Oxford University

The Dragon and The... Oxford University

General Philosophy Oxford University

Getting to Zero: Mic... Oxford University

Global Economic Go... Oxford University

Global Economic Go... Oxford University

The God Delusion W... Oxford University

Health, Law and Em... Oxford University

History of Tropical M... Oxford University

http://www.slideshare.net/tbirdcymru/itunes-u-corporate-channel-of-free-educational-resources
Under the category “OER Tools and Services”

<table>
<thead>
<tr>
<th><strong>iTunes U</strong></th>
<th>**iTunes U puts the power of the iTunes Store to work for colleges and universities, so users can easily search, download, and play course content just like they do music, movies, and TV shows. **iTunes U contains free audio &amp; video downloads from universities across the United States including Stanford, Duke, MIT, Arizona State. More European universities are now joining and putting lectures on iTunes, such as University College London, the Open University and Trinity College Dublin</th>
</tr>
</thead>
</table>


http://www.slideshare.net/tbirdcymru/itunes-u-corporate-channel-of-free-educational-resources
Elements of a successful OER channel

• Attractive to contributors
• Usable
• Useful
• Used (and re-used)
• Sustainable

http://www.slideshare.net/tbirdcymru/itunes-u-corporate-channel-of-free-educational-resources
# iTunesU OER Success Factors

<table>
<thead>
<tr>
<th>Attractive to contributors</th>
<th>Usable</th>
<th>Useful</th>
<th>Used</th>
<th>Sustainable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profile ✔</td>
<td>User Experience ✔</td>
<td>Quality material ✔</td>
<td>Download numbers ✔</td>
<td>Over 800 universities ✔</td>
</tr>
<tr>
<td>‘Apple gloss’ ✔</td>
<td>Search function ✔</td>
<td>Consistency ✔</td>
<td>Teachers ✔</td>
<td>Apple ✔</td>
</tr>
<tr>
<td>International reach ✔</td>
<td>Apple mobile ✔</td>
<td>Copyright ☑️</td>
<td>Personal ✔</td>
<td>Benefit to contributors/ institution ✔</td>
</tr>
<tr>
<td></td>
<td>Linux, Android ☑️</td>
<td>Feedback ☑️</td>
<td>Not very repurposeale ☑️</td>
<td></td>
</tr>
<tr>
<td>Discoverability ☑️</td>
<td>Community ☑️</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[http://www.slideshare.net/tbirdcymru/itunes-u-corporate-channel-of-free-educational-resources](http://www.slideshare.net/tbirdcymru/itunes-u-corporate-channel-of-free-educational-resources)
It’s all in the downloads

<table>
<thead>
<tr>
<th>University</th>
<th>Downloads</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open University, UK</td>
<td>Over 34 million since June 2008</td>
</tr>
<tr>
<td>University of Oxford</td>
<td>Over 9 million since June 2008</td>
</tr>
<tr>
<td>Coventry University</td>
<td>2.5 million in 2010 alone</td>
</tr>
<tr>
<td>University of Warwick</td>
<td>1 million Jan ‘09 – June ‘10</td>
</tr>
</tbody>
</table>

http://www.slideshare.net/tbirdcymru/itunes-u-corporate-channel-of-free-educational-resources
Lextutor

Compleat Lexical Tutor v.6.2
For data-driven language learning on the Web

CURRENT PROJECTS: De-Homynizer; Dedicated lexicon
NEAR FUTURE: To-Do List

TUTORIAL

Test your lexis

List Learn
List-Learn BNC20

Group Lex v.7.7
*New* Input Security
Multi-Edit

RA-Reading
Resource Assistant, 13 stories

Corpus
Grammar

RESEARCH

Range

Vocabprofile
Faster, bigger classic May '11

Concordance
New! Output Refining
Complete July '11

N-Gram v.4.2
10MB capacity on upload
July 2011

RT Builder v.3

Vocab Stats

Text Tools
Better T-Units "NEW"

Familiarizer

Frequency

Text Lex Compare

RaNdOmlCiTy
MCCs are ABC @ Compleat

KeyWords

Morphology

SUMMER TARGETS - FULL MAC COMPAT, VP HOMOGRAPHS, MULTIFILE CONCS + N-GRAMS, PHONO-MEMSPAN

http://www.lex tutor.ca
VocabProfile + Web Collocations OER

http://www.lectutor.ca/vp/

http://www.youtube.com/watch?v=iYzgZgZHMovI
AntConc freeware – build your own corpus

<table>
<thead>
<tr>
<th>Software</th>
<th>AntConc</th>
<th>A freeware concordance program for Windows, Macintosh OS X, and Linux.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>The AntConc Homepage (including previous versions, tutorials, and help)</td>
</tr>
<tr>
<td>Platform</td>
<td>Download</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AntConc3.2.4w</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AntConc3.2.4 - Readme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Screenshots</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Tested on Windows 98, 2000, ME, XP, Vista, Win 7</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AntConc 3.2.4m</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AntConc3.2.4 - Readme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Essential guide to installing AntConc on Macintosh OS X</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download X11 (for OS X 10.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Screenshots</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Tested on Macintosh OS X 10.6.8</em></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AntConc3.2.4u</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AntConc3.2.4 - Readme</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AntConc3.2 - Linux Quick Install Guide Written by Laurence Anthony</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Brief guide to installing AntConc on Linux Written by Graham Ranger</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Screenshots</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Tested on Ubuntu 10 (Linux Mint 10)</em></td>
</tr>
</tbody>
</table>

http://www.antlab sci.waseda.ac.jp/software.html
AntConc video tutorials

Here are some video tutorials to help you learn how to use AntConc.

- Getting Started
- Tutorial 1: Concordance Tool - Basic Features
- Tutorial 2: Concordance Tool - Advanced Features
- Tutorial 3: Concordance Plot Tool - Basic Features

http://www.antlab.sci.waseda.ac.jp/antconc_index.html
David Lee’s Corpus-based Linguistics

LINKS

Corpora, Collections, Data Archives*

---

* What are the different types of corpora?
* Ok, so how do I actually get my hands on these corpora, and what do I do with them?
* For freely accessible, on-line corpora...

---

Major English Language General Corpora


- **Pre-electronic Corpora**: (biblical & literary studies, early dictionaries, etc.)

- **1st-generation Major Corpora**:
  - Brown, LOB, LLC, Kolhapur, Wellington, etc.

- **2nd-generation Corpora**:
  - Mega Corpora:
    - British National Corpus (BNC); Corpus of Contemporary American English (COCA); COBUILD Bank of English
  - Not-so-mega Corpora:
    - ICE-GB, American National Corpus (ANC), etc.

[Some people call the major "general" corpora above "balanced corpora", but I personally avoid this term. To say something is "balanced" suggests that linguists have agreed on different genres (patently untrue) in order to achieve this "balance", while for others "balanced" just means an equal sample of spoken & written. For me, the best terms for the "general language", though I personally avoid generalizations about "language in general".]

Thank you!

**Blog:** Technology for Open English – Toying with Open E-resources
[http://www.alannahfitzgerald.org](http://www.alannahfitzgerald.org)

**Twitter:** @AlannahFitz

**Slideshare:**
[http://www.slideshare.net/AlannahOpenEd/](http://www.slideshare.net/AlannahOpenEd/)

**Email:**
alannah.fitzgerald@education.concordia.ca
a.t.d.fitzgerald@durham.ac.uk