

# Electronic lexicography and computer-assisted language learning: Breaking down the barriers

Sylviane Granger

University of Louvain  
Centre for English Corpus Linguistics  
[sylviane.granger@uclouvain.be](mailto:sylviane.granger@uclouvain.be)

# Outline

- Introduction
- Some writing-oriented functionalities
- Three priorities for closer integration between electronic dictionaries and CALL
  - Better exploitation of the database structure of electronic dictionaries
  - Integration of electronic dictionaries into Learning Management Systems (LMS)
  - Design of customisable specialised modules
- Conclusion



Links between electronic  
lexicography and CALL

# Computer-assisted Language Learning

- *“the search for and study of applications of the computer in language teaching and learning”*  
(Levy 1997)

# Electronic lexicography

- *“E-lexicography can mean a number of things: using technology for making dictionaries; using dictionaries (and other lexical resources) for high-tech applications; and making (and publishing) dictionaries in electronic form” (Kilgarriff 2009a)*

# Electronic dictionaries (EDs)

- Focus on human-oriented electronic dictionaries; exclusion of strictly computer-oriented NLP lexicons

# EDs and CALL

- Largely two worlds apart
- Signs of rapprochement: cf. Abel (2010):
  - Dictionary-cum-CALL
  - CALL-cum-dictionary
- The most effective integration so far has been achieved in the area of **reading** activities

# Reading: textual glosses

- Textual glosses created by the teacher or materials designer to explain difficult words in a text
- The glosses become visible with a simple click on a word or by hovering the mouse over it
  - 😊 contextualized (....in a little while = period of time)
  - 😊 they work: positive effect on reading comprehension and language learning (Nation 2011)
  - 😞 cannot be used with texts that have not been glossed in advance
  - 😞 time-consuming for teachers/materials designers



# Reading: dictionary lookup

- Direct access to the relevant entry in the electronic dictionary for any word in a text by just clicking on it
  - 😊 can be used with any text
  - 😞 not contextualized, hence requires more skills and time on the part of the user especially for
    - long entries (*to fit*)
    - homographs (*save*: verb, noun, preposition)
    - multiword units (*heavy rain, make heavy weather of sth*)

# Reading: intelligent lookup

- *Sharp Intelligent Dictionary*

- Connectivity: « *provide the user with the correct equivalent for a word used in the passage at hand* » (Whitelock & Edmonds 2000)
- Two technologies:
  - Part-of-speech disambiguation (*while*: noun vs. conj.)
  - Recognition of multiword units (*let go of, in relationship to*)
    - 20%-50% of the word tokens in a text belong to MWUs
    - 89% precision of collocation detection

# Writing in CALL environment

- The link between CALL environment and electronic dictionary is either absent or largely ineffective for writing activities
- Doubly paradoxical
  - Learner needs are particularly acute for writing
  - Remarkable efforts have been made to turn dictionaries into efficient productive tools (Rundell 1999)

# Purposes for dictionary use (Nation 2001)

## Writing

- Look up **unknown** words needed to write
- Look up the spelling, meaning, grammar, constraints on use, collocations, inflections and derived forms of **partly known** words
- Confirm the spelling, meaning, etc. of **known** words
- Check that a word **exists**
- Find a **different word** to use instead of a known one
- Correct an **error**

## Reading

- Look up **unknown** words
- Confirm the meaning of **partly known** words
- Confirm **guesses** from context



**Some writing-oriented functionalities**

## Some writing-oriented functionalities

1. Dictionary lookup
  2. Concordancing
  3. Link to collocator tool
  4. Vocabulary profiling
  5. Error warning
- Focus on monolingual searches

# 1. Dictionary lookup

- The link takes the writer to the full headword entry
- No direct access to the specific type of information s/he might be looking for (formality level, frequency, collocations, etc.)

# *Oxford iWriter* : pop-up dictionary

*They were anxious to get the **support** from each State.*

THE IMAGE CANNOT BE SHOWN  
DUE TO COPYRIGHT REASONS



*iWriter* : go to entry

THE IMAGE CANNOT BE SHOWN  
DUE TO COPYRIGHT REASONS

Extra information menu



THE IMAGE CANNOT BE SHOWN  
DUE TO COPYRIGHT REASONS

# Longman Writing Assistant

- Direct access to four types of information
  - Thesaurus
  - Grammar
  - Collocations
  - Errors



THE IMAGE CANNOT BE SHOWN  
DUE TO COPYRIGHT REASONS

## 2. Concordancing

*Concord Writer* (Cobb 2007)

THE IMAGE CANNOT BE SHOWN  
DUE TO COPYRIGHT REASONS

# Concordancing vs. dictionary

- Gabel (2001): “*it is argued that **concordancers** are superior to traditional grammar books, **dictionaries** and coursebooks, because they allow easy access to huge amounts of `real' language in use (...)*”.
- Kilgarriff (2009b): “*Most learners do not want to be corpus linguists, and **concordances** are unfamiliar and difficult objects. But **dictionaries** are familiar from an early age (...)*”.

## 3. Link to collocator tool

- Milton's (2006) *Check my Words* writing aid tool
  - Lexical aid: link to *Word Neighbors* to identify the preferential lexical patterning of words

Word Neighbors: *support* (n.)

THE IMAGE CANNOT BE SHOWN  
DUE TO COPYRIGHT REASONS



# [ 4. Vocabulary profiling ]

- Highlight some words in the text
  - Cobb's *VocabProfile* (different colours for different frequency bands)
  - Oxford *iWriter*: 'Highlight Academic Word List' icon

# [ Highlight Academic Word List ]

THE IMAGE CANNOT BE SHOWN  
DUE TO COPYRIGHT REASONS

# [ 5. Error warning ]

- Requires prior analysis of learner corpus data
- Error notes in learners' dictionaries
  - *Get it right boxes* in Macmillan Dictionary based on the International Corpus of Learner English
  - '*Errors*' icon in Longman Writing Assistant based on Longman Learner's Corpus
- Automatic highlighting of potential errors in learner texts
  - *Check my Words* (Milton 2006): highlights words and phrases that are often used incorrectly by Chinese learners of English.
  - Limitation: focus on grammatical errors

# Overall picture

- The **coverage** of writing needs is patchy. Not all user needs are covered (cf. Nation's list)
- **Abundance of tools** with each tool catering for a portion of the needs
- A lot of **time and effort** is required from writers to access the production-oriented information they need
- The user has **no control** over the type of information provided
- Most tools target **generic learners**; no **customization** in function of
  - learners' mother tongue background
  - targeted language variety
  - task



**Three priorities for closer integration  
between EDs and CALL**

## Three priorities for closer integration

- 1) Enhanced exploitation of the database structure of dictionaries**

# Electronic dictionaries

- *“they are collections of structured electronic data that can be accessed with multiple tools, enhanced with a wide range of functionalities, and used in various environments” (de Schryver 2003)*

# Structured database

- Selected facts about words are stored in different fields in present-day dictionaries

Etymology	Formality label
Inflected forms	Frequency
Derived forms	Synonyms/antonyms
Part-of-speech category	Example bank
Grammatical patterning	Usage notes
Collocates	Error notes
Geographical label	Cultural notes



# Access

- « *This highly structured format means that much of the information in the database is accessible to computerized searching and filtering* » (Atkins & Rundell 2008)
- In principle it should be possible for writers to access the different types of information separately in function of their needs
- In reality: this is not implemented or poorly implemented
- But cf. Verlinde 2011: different access possibilities for writing, reading & translating

# Direct differential access

- Direct access to different types of productively relevant information via icons that users can turn on or off
- Example:
  - Collocation
  - Formality level
  - Frequency
  - Geographical variety
- **Adaptable** system (manual customization of the system by the user) (Gamper & Knapp 2002)
- Dictionary logs can progressively help to turn the adaptable system into an **adaptive** one (system adapts automatically to the user)

## Two consultation modes

### ■ **On-the-fly**

- The student clicks on the relevant tab while writing.
- The teacher does the same while marking.

### ■ **Batch mode**

- Student: post-writing
- Teacher: pre- or post-marking

# Student: post-writing checks

- Students have the option of highlighting a range of features in their text after they have written it
  - Words in different frequency bands or words in curriculum-/exam-specific vocabulary lists
  - Stylistically marked words (formal, informal, taboo)
  - Potential errors
  - Etc.

## Teacher: pre-/post-marking checks

- Teachers can use the same functionalities to get additional information on learners' texts
  - before they start marking
  - when they have finished marking

# Three priorities

**2) Integration of electronic dictionary  
into Learning Management Systems**

# Learning Management Systems (LMS)

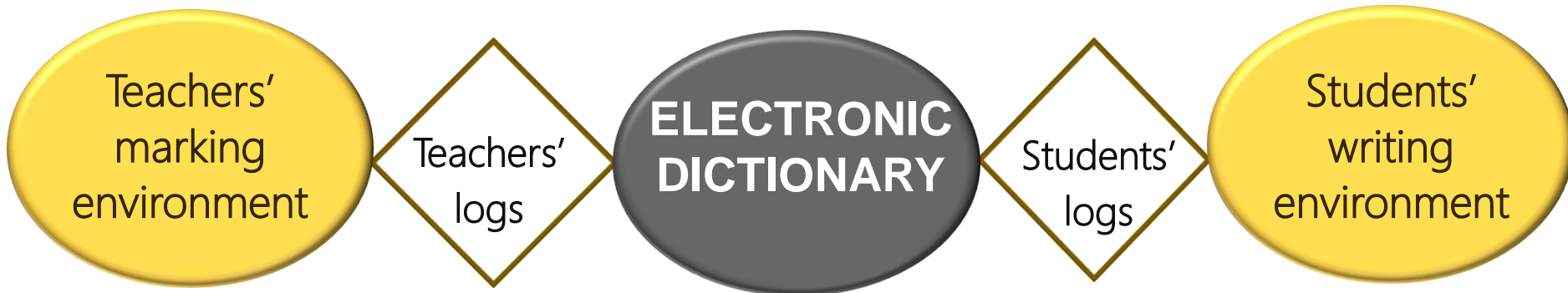
- *“LMSs provide teachers with many features to create, manage and administrate online courses, allowing them to include different kinds of learning objects/activities such as learning material, forums, quizzes, examples, and so on, and facilitating administrative issues such as enrolment, grading and monitoring the learners’ progress and performance” (Graf et al 2010)*
- Very popular
  - 2011 survey national study of undergraduate students and information technology; 73 % of students use an LMS

# Moodle

- Open source LMS (Nagel 2010, Tsun-Ju 2011)
- Used by about 1.1 million teachers; more than 38 million users
- Latest version:
  - Integration with third-party and external tools (e.g. Flickr, Google Docs, etc.)
  - Mobile app provides access to the Moodle LMS

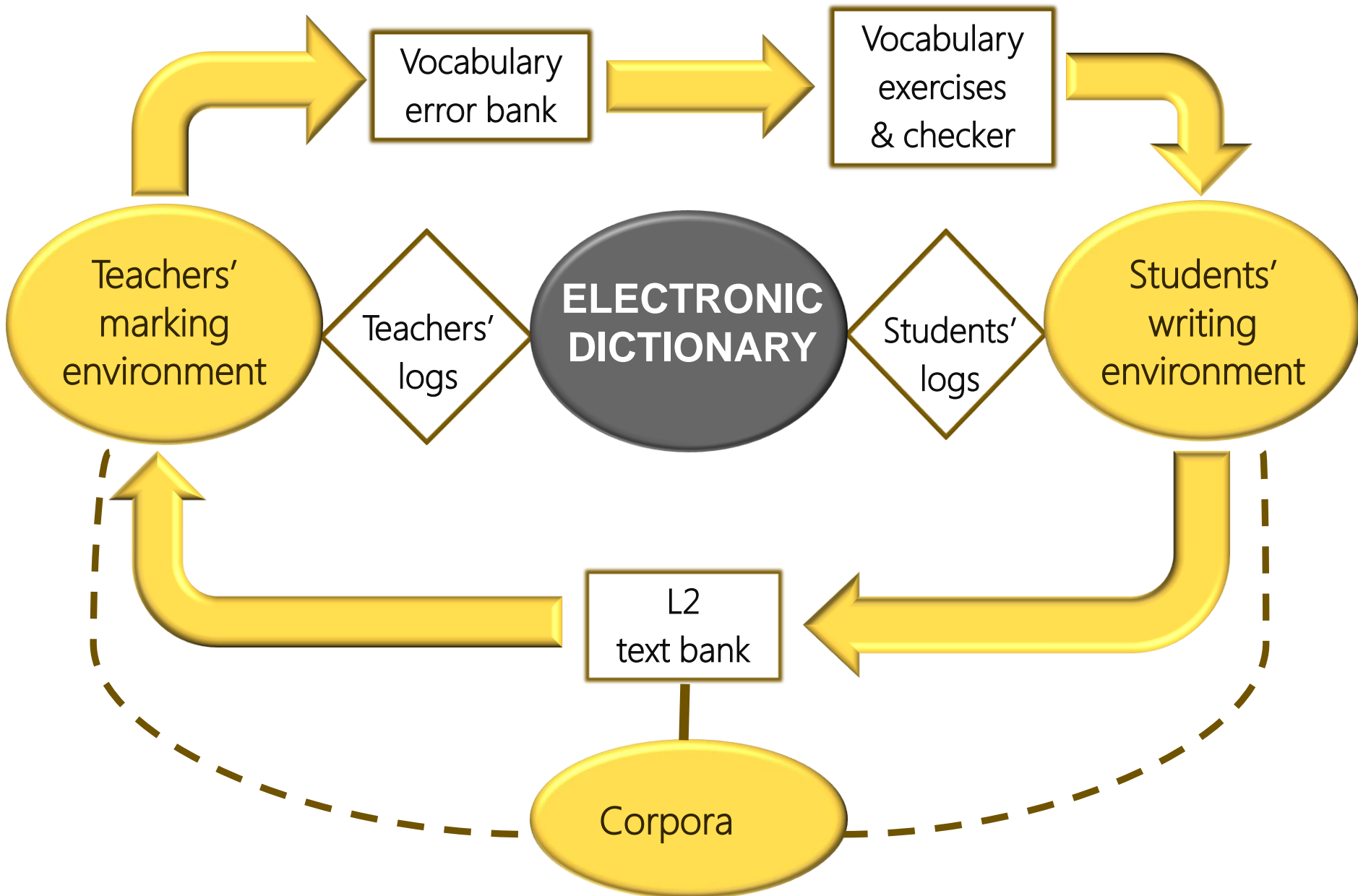


# Lexical component of LMS-based L2 writing environment



- Student and teacher environments
  - Wible et al (2001): *iWill* (Intelligent Web-based Interactive Language Learning)
  - Milton (2006): *Check My Words* and *Mark My Words*

# Lexical component of L2 writing environment



# Three priorities

**3) Design of customisable specialised modules**

# General vs specialised

- Most tools target the 'generic learner'
- Need for specialised modules
- Massive needs for English for Academic Purposes (EAP)

# Macmillan English Dictionary (2007)

- Louvain contribution: extended 'Improve your writing skills' section
  - 12 major rhetorical functions (introducing a topic, contrasting, exemplifying, etc.)  
+ specific EAP words and phrases used to express them
  - Focus on phraseology (collocations, patterns of use)
  - Use of large learner corpus (ICLE) to highlight learners' difficulties (errors; over- and underuse)

# Pros and cons

## ■ Pros:

- Rich material based on authentic difficulties shared by many learners
- Highly fruitful collaboration between
  - trained lexicographers (Macmillan)
  - experienced ELT specialists (Louvain)

## ■ Cons:

- no real integration of the writing section
  - CD-ROM: no links between the writing skills sections and the dictionary
- forced to reject many learner difficulties which were restricted to one language population/family (e.g. false friends) because of the generic nature of the dictionary.

## **The *Louvain EAP Dictionary (LEAD)***

An **integrated** web-based tool to help non-native speakers write **academic texts** in English (Granger & Paquot 2010 a,b)

- **New features**
  - **Multiplicity of access modes**
  - **Customisation (discipline and L1)**
  - Integrated EAP-focused exercises
  - Direct access to EAP corpora (cf. M. Paquot's presentation)

# Multiplicity of access modes

- Word search
- Search by translation
  - Bilingualised dictionary
- Search by function



# Search by function

**C E C L** The Louvain EAP dictionary

Add information  Search

Adjectives	Adverbs	Conjunctions	Prepositions
<a href="#">contrary</a> <a href="#">contrasting</a> <a href="#">different</a> <a href="#">differing</a> <a href="#">opposite</a> <a href="#">reverse</a> <a href="#">unlike</a>	<a href="#">by comparison</a> <a href="#">by contrast</a> <a href="#">conversely</a> <a href="#">in comparison</a> <a href="#">in contrast</a> <a href="#">on the contrary</a> <a href="#">on the other hand</a>	<a href="#">whereas</a> <a href="#">while</a>	<a href="#">as against</a> <a href="#">as opposed to</a> <a href="#">by comparison with</a> <a href="#">contrary to</a> <a href="#">in comparison with</a> <a href="#">in contrast to</a> <a href="#">in contrast with</a> <a href="#">unlike</a> <a href="#">versus</a>

Add information  
 Add information  
 Compare and contrast: Describing similarities  
 Compare and contrast: Describing differences  
 Conclude  
 Condition  
 Express cause and effect: Cause  
 Express cause and effect: Effect  
 Express cause and effect: Linking cause and effect  
 Introduce a concession  
 Introduce a topic: Introduce the main topic  
 Introduce a topic: Introduce related ideas  
 Introduce a topic: Change topic  
 List and sequence: Introducing the first item  
 List and sequence: Introducing the second and following items  
 List and sequence: Introducing the last item  
 List and sequence: Leaving the list unfinished  
 Summarize  
 Quoting and reporting  
 Refer to information in other parts of the text (chapter, section and paragraph)  
 Refer to tables, figures and graphs  
 Give examples  
 Explain  
 Define terms  
 Reformulate: Paraphrase and clarify  
 Express possibility and certainty  
 Express personal opinion

# Customisability

- The content is automatically adapted to users' needs in terms of
  - mother tongue background
    - French, Dutch (more to come!)
  - discipline
    - Business, linguistics, medicine (more to come!)



# The Louvain EAP dictionary

## EAP dictionary

[Welcome](#)

[Dictionary](#)

[Concordancer](#)

[Exercises](#)

[References](#)

Please select a discipline:

Business

What is your mother tongue?

French

## Links

[Centre for English Corpus Linguistics](#)

## Intranet

[Lexicographer's corner](#)

Contact : [Magali Paquot](#) & [Sylviane Granger](#)

# Discipline customisation

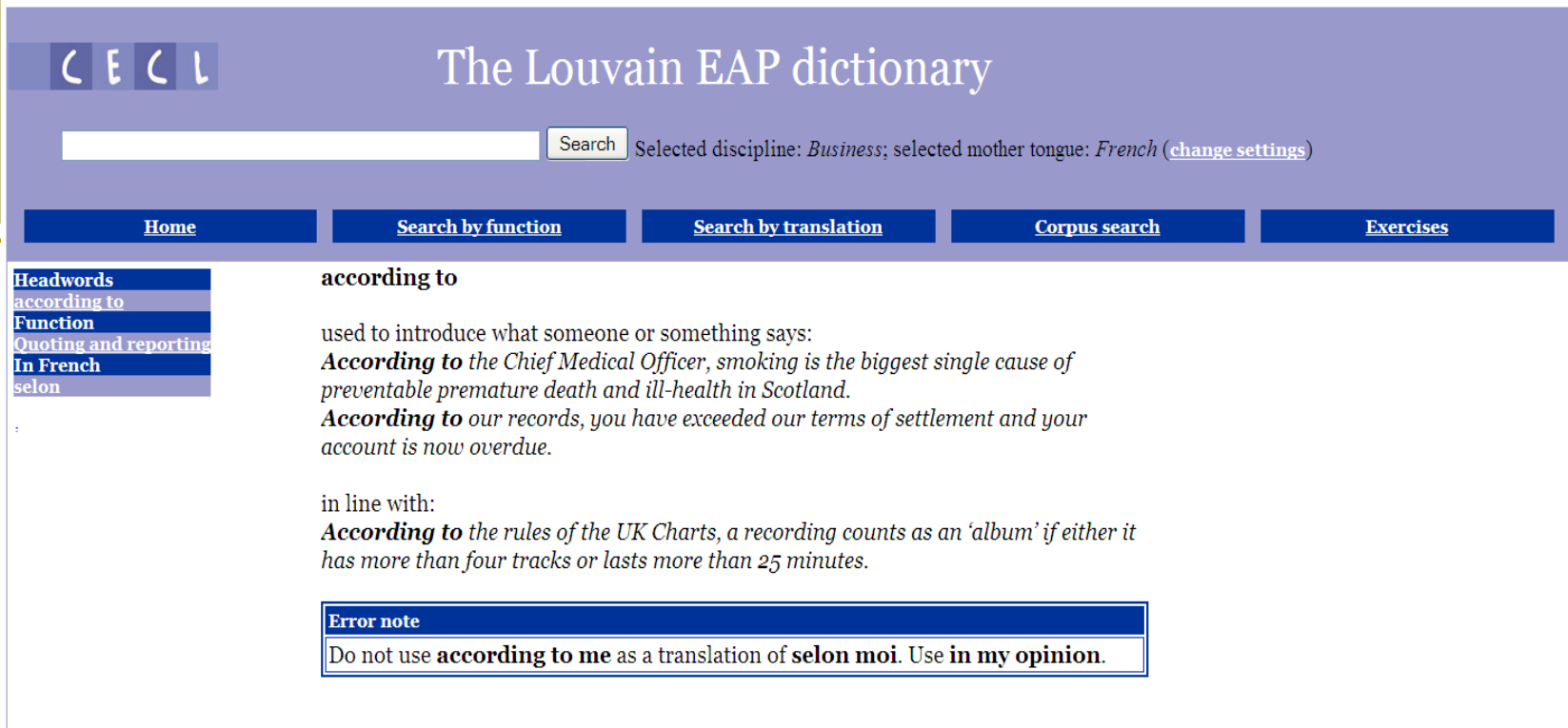
**C E C L** The Louvain EAP dictionary

Selected discipline: *Business*; selected mother tongue: *French* ([change settings](#))

[Home](#) [Search by function](#) [Search by translation](#) [Corpus search](#) [Exercises](#)

<b>Headwords</b> <a href="#">contrast (n.)</a> <a href="#">contrast (vb.)</a> <a href="#">contrasting (adj.)</a> <a href="#">in contrast (adv.)</a> <a href="#">by contrast (adv.)</a> <a href="#">in contrast with (prep.)</a> <a href="#">in contrast to (prep.)</a>	<b>contrast (n.)</b> a difference between people, ideas, situations or things that you can see clearly when they are compared or put close together: <ul style="list-style-type: none"><li>• There was <b>a marked contrast between</b> rural and urban settings.</li></ul>	<b>Collocations</b> Adj + <i>contrast</i> <i>direct, marked, sharp, stark, striking</i> V + <i>contrast</i> <i>draw, provide</i>
<b>Function</b> Compare and contrast: Describing differences	<b>Examples in Business...</b>	
<b>In French</b> contraste différence	<div style="border: 1px solid black; padding: 5px;"><p>The actual deficit was \$155.1bn, <b>a sharp contrast to</b> the revised target of \$144bn and the original target of \$108bn.</p></div>	

# L1 customisation



The screenshot displays the 'The Louvain EAP dictionary' interface. At the top, there is a search bar with a 'Search' button and a dropdown menu showing 'Selected discipline: Business; selected mother tongue: French (change settings)'. Below the search bar are five navigation buttons: 'Home', 'Search by function', 'Search by translation', 'Corpus search', and 'Exercises'. On the left side, there is a vertical menu with the following items: 'Headwords', 'according to', 'Function', 'Quoting and reporting', 'In French', and 'selon'. The main content area shows the search results for 'according to'. It includes a definition: 'used to introduce what someone or something says: **According to** the Chief Medical Officer, smoking is the biggest single cause of preventable premature death and ill-health in Scotland. **According to** our records, you have exceeded our terms of settlement and your account is now overdue.' It also includes a note: 'in line with: **According to** the rules of the UK Charts, a recording counts as an 'album' if either it has more than four tracks or lasts more than 25 minutes.' At the bottom, there is a blue-bordered box with the text: 'Error note: Do not use **according to me** as a translation of **selon moi**. Use **in my opinion**.'

CECL

## The Louvain EAP dictionary

Search Selected discipline: *Business*; selected mother tongue: *French* ([change settings](#))

[Home](#) [Search by function](#) [Search by translation](#) [Corpus search](#) [Exercises](#)

Headwords  
according to  
Function  
Quoting and reporting  
In French  
selon

**according to**

used to introduce what someone or something says:  
**According to** the Chief Medical Officer, smoking is the biggest single cause of preventable premature death and ill-health in Scotland.  
**According to** our records, you have exceeded our terms of settlement and your account is now overdue.

in line with:  
**According to** the rules of the UK Charts, a recording counts as an 'album' if either it has more than four tracks or lasts more than 25 minutes.

**Error note**  
Do not use **according to me** as a translation of **selon moi**. Use **in my opinion**.

## From stand alone to integrated tool

- Our objective: integrate LEAD as a module into a wider writing environment
- Rundell's mix and match scenario:
  - *“A possible scenario is to see our reference materials as a set of components which customers can mix and match according to their needs. For example, a learner from China doing a Masters in agriculture at a British university could have access to a ‘core’ **ELT dictionary** with the option of Chinese translations, **academic-writing aids**, and subject-specific terminology”.* (Rundell 2007).

# LEAD module in L2 writing environment

- **On-the-fly:** direct link to EAP module

- **Post-check:**


- automatic highlighting of all the EAP words/phrases and link to entries in LEAD
- warning for error-prone items (in function of learners' L1)





Conclusion





- 
- A large black left bracket and a yellow right bracket are positioned at the top of the slide, with a horizontal line extending between them. The line is light green on the left and yellow on the right.
- « *One major challenge for CALL practitioners is the need to harness the vast assortment of technology resources in a manner that capitalizes on the opportunities they present to meet pedagogical expectations* »  
(Butler-Pascoe 2011).

- 
- A large black left bracket and a large yellow right bracket are positioned at the top of the slide, with a horizontal olive-green line passing through them.
- Electronic dictionaries have a major role to play in this search for greater pedagogical effectiveness.
  - But to be truly effective EDs - and more particularly MLDs - need to
    - make full use of their database structure in order to facilitate and speed up information access
    - be integrated in the type of environment that the vast majority of learners and teachers are using, viz. Learning Management Systems

- 
- A large black left square bracket and a large yellow right square bracket are positioned at the top of the slide. A horizontal line with a light green-to-yellow gradient runs across the slide, starting from the left edge and ending at the right edge, partially overlapping the brackets.
- In addition, there is a need for customisable modules like LEAD which cater for learners' more specific needs
  - Last but not least, to ensure pedagogical effectiveness, it is imperative to rely on multidisciplinary expertise
    - IT, lexicography, corpus linguistics, CALL, SLA and language pedagogy



**Thank you very much for your attention!**

# References

- Abel, A. (2010). Towards a systematic classification framework for dictionaries and CALL. In Granger, S. & Paquot, M. (eds.) *eLexicography in the 21st century: New challenges, new applications*. Presses universitaires de Louvain: Louvain-la-Neuve, 3-11.
- Atkins, B.T.S. & Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Butler-Pascoe, M.H. (2011). The History of CALL: The Intertwining Paths of Technology and Second/Foreign Language Teaching. *International Journal of Computer-Assisted Language Learning and Teaching*, 1(1), 16-32.
- Cobb, T. (2007). Computing the vocabulary demands of L2 reading. *Language Learning and Technology* 11/3, 38-63.
- de Schryver, G.-M. (2003). Lexicographer's dreams in the electronic-dictionary age. *International Journal of Lexicography*, Vol. 16, No. 2, 144-199.
- Gabel, S. (2001). Over-indulgence and under-representation in interlanguage: Reflections on the utilization of concordancers in self-directed foreign language learning. *Computer Assisted Language Learning*, 14, 269-288.
- Gamper J. & Knapp J. (2002), Adaptation in a Vocabulary Acquisition System, *KI - Zeitschrift Künstliche Intelligenz*, Vol. 3(2), 27-30.

- Graf, S., Liu, T.-C & Kinshuk (2010). Analysis of learners' navigational behaviour and their learning styles in an online course. *Journal of Computer Assisted Learning*, 26, 116–131.
- Granger, S. & M. Paquot (2010a) Customising a general EAP dictionary to meet learner needs. In Granger, S. & M. Paquot (eds) (2010) *eLexicography in the 21st century: New challenges, new applications*. Louvain-la-Neuve, Presses universitaires de Louvain, 87-96.
- Granger, S. & M. Paquot (2010b) The Louvain EAP Dictionary (LEAD). In Proceedings of the XIV EURALEX International Congress, Leeuwarden, The Netherlands, 6-10 July 2010, 321-326.
- Grunwald Associates LLC (2011). ECAR National Study of Undergraduate Students and Information Technology, 2011—Survey Questionnaire (Survey Instrument). Boulder, CO: EDUCAUSE Center for Applied Research, available from <http://net.educause.edu/ir/library/pdf/ERS1103/ERS1103W.pdf>
- Kilgarriff, A. (2009a). How to monetise a web presence (and hoover a moose). A report on the e-lexicography conference at Louvain-la-Neuve, Belgium, 22-24 October 2009. <http://www.uclouvain.be/en-307398.html>

# References

- Kilgarriff, A. (2009b). Corpora in the Classroom Without Scaring the Students. Proceedings from the 18th International Symposium on English Teaching, Taipei. [www.kilgarriff.co.uk/Publications/2009-K-ETA-Taiwan-scaring.doc](http://www.kilgarriff.co.uk/Publications/2009-K-ETA-Taiwan-scaring.doc)
- Levy, M. (1997). *Computer-Assisted Language Learning: Contexts and Conceptualization*, Oxford: Oxford University Press.
- *Longman Dictionary of Contemporary English* (2009). 5th edition (with Writing Assistant)
- *Macmillan English Dictionary for Advanced Learners* (2007). Second edition.
- Milton, J. (2006). Resource-Rich Web-Based Feedback: helping learners become independent writers. In Hyland, K & Hyland F. (eds), *Feedback in Second Language Writing: Contexts and Issues*. Cambridge University Press, 123-137.
- Nagel, N. (2010). Moodle 2.0 Boosts Integration and Web 2.0 Features. Campus Technology
- Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- *Oxford Advanced Learner's Dictionary* (2010). 8th edition (with Oxford iWriter) Oxford: Oxford University Press.

# References

- Rundell, M. (1999). Dictionary Use in Production. *International Journal of Lexicography* 12/1, 35-53.
- Rundell, M. (2007). The dictionary of the future. Granger S. (2007) (ed.) *Optimizing the role of language in Technology-Enhanced Learning*. Proceedings of the expert workshop organized by the Integrated Digital Language Learning (IDILL) seed grant project, Louvain-la-Neuve (Belgium), 4-5 October 2007, 49-51.
- Tsun-Ju, L. (2011). Review of Moodle 2.0. *Language Learning & Technology*, 15/2, 27-33.
- Verlinde, S. (2011). Modelling interactive reading, translation and writing assistants. In Fuertes-Olivera, P.A. & Bergenholtz, H. (eds.) *e-Lexicography. The internet, digital initiatives and lexicography*. London & New York: Continuum, 275-286.
- Whitelock, P. & Edmonds, P. (2010). The Sharp Intelligent Dictionary. In Proceedings of the Ninth EURALEX International Congress (EURALEX 2000), 871-876.
- Wible, D., Kuo, C-H., Chien, F-Y., Liu, A. & Tsao, N-L. (2001). A web-based EFL writing environment: integrating information for learners, teachers, and researchers. *Computers and Education* 37, 297-315.



# Electronic resources

- Check My Words/Mark My Words
  - <http://mws.ust.hk/cmw/index.php>
  - <http://mws.ust.hk/mmw/index.php>
- ConcordWriter
  - [http://conc.lex tutor.ca/concord\\_writer/index.pl?lingo=English](http://conc.lex tutor.ca/concord_writer/index.pl?lingo=English)
- Moodle
  - <http://moodle.org/>
- VocabProfile
  - <http://www.lex tutor.ca/vp/eng/>
- Word Neighbors
  - <http://wordneighbors.ust.hk/>

# Centre for English Corpus Linguistics

- <http://www.uclouvain.be/en-cecl.html>