

## A customisable dictionary-cum-CALL resource for EAP writing

Sylviane Granger • Magali Paquot

### 1 The Louvain EAP dictionary

The *Louvain EAP dictionary* (LEAD) is a web-based English for Academic Purposes (EAP) dictionary for non-native writers (Granger and Paquot, in press). It contains a rich corpus-based description of c. 900 non-technical words and phrases that express key functions in academic discourse (e.g. contrast, exemplification or cause and effect), with particular focus on their phraseology (collocations and recurrent phrases). The lexical entries provide information derived from an analysis of a large corpus of academic texts (i.e. the academic component of the *British National Corpus*), as well as a range of home-made discipline-specific corpora and English as a Foreign Language (EFL) learner corpora representing a wide range of first language (L1) populations. Its main originality is its customisability: the content is automatically adapted to users' needs in terms of **discipline** and **mother tongue background**.

Before using the dictionary, users have to select a domain (currently business, medicine, linguistics, or general EAP for users working in other disciplines) and specify their L1 background (currently French). Domain selection makes it possible to customise the output and illustrate the phraseological environment of a search word by means of discipline-specific examples extracted from a corpus of either business, medicine or linguistics texts. Access to large generic corpora is a feature that is now regularly included in electronic dictionaries. The LEAD, however, is innovative in giving access to discipline-specific corpora rather than generic corpora.

One of the purposes of L1-background identification is to alert users to potential difficulties that a specific L1 population typically encounters. We are currently focusing on French as an L1 background but are planning to include more languages in the future. To create both the generic usage notes and the L1-specific notes, we make use of the *International Corpus of Learner English* (Granger et al 2009) as well as the *Varieties of English for Specific Purposes dAtabase* (VESPA), a new learner corpus, currently being developed at the Centre for English Corpus Linguistics in collaboration with several international partners. The VESPA corpus contains L2 texts from a wide range of L1 backgrounds (e.g. French, Spanish, Swedish, and Polish), disciplines (linguistics, business, engineering, sociology, etc), genres (papers, reports, MA dissertations) and degrees of writer expertise in academic settings (from

Sylviane Granger

Centre for English Corpus Linguistics, Université catholique de Louvain

E-mail: sylviane.granger@uclouvain.be

Magali Paquot

Centre for English Corpus Linguistics, Université catholique de Louvain

E-mail: magali.paquot@uclouvain.be

first-year students to PhD students) (see <http://www.uclouvain.be/en-258647.html> for further details). Errors and difficulties encountered in a wide range of L1 learner populations are dealt with in generic error notes that are displayed irrespective of the L1 background selected by the user. For example, the lexical entry for “namely” includes a warning against learners’ attested tendency to use this adverb to introduce examples instead of “such as” or “e.g”. Errors that are found exclusively in the writing of French learners are described in notes that only show up if French is selected as L1 background. We also include warnings about common translation mistakes such as the erroneous translation of the French ‘prétendre’ by its false friend ‘pretend’ in English.

Another key feature of the LEAD is that it makes full use of the capabilities afforded by the electronic medium in terms of multiplicity of access modes (see Sobkowiak 2002; Tarp 2009). The dictionary can be used as both a semasiological dictionary (from lexeme to meaning) and an onomasiological dictionary (from meaning/concept to lexeme) via a list of typical rhetorical or organisational functions in academic discourse (cf. Pecman 2008). Selecting one of these functions provides the user with a list of lexical items categorized according to their part-of-speech (nouns, verbs, adjectives, adverbs, conjunctions and prepositions). One of the main advantages of this access mode is that it suggests alternatives and thereby helps users enlarge their academic repertoire. It is also a semi-bilingual dictionary (cf. Laufer & Levitzky-Aviad 2006) as users who have selected a particular mother tongue background can search lexical entries via their translations into that language.

The LEAD dictionary is designed as an integrated tool where the actual dictionary part is linked up to other language resources (in particular, discipline-specific corpora and a corpus handling tool). As described in the next section, we also want to turn the LEAD into a dictionary-cum-CALL resource (Abel, in press) by adding exercises targeting learners’ attested difficulties.

## **2 A dictionary-cum-CALL resource for academic writing**

Abel (in press) uses the term ‘dictionary-cum-CALL’ to refer to electronic dictionaries that serve as the basis for the development of CALL modules (e.g. Verlinde et al 2007). As such, they constitute yet another type of mixed dictionary genres that are the result of ‘hybridisation’ with one or more types of reference work (Hartmann, 2005). The development of ‘dictionary-cum-CALL’ resources reflects a general tendency to strengthen the amount and quality of productively-oriented information in learner dictionaries (Rundell 1999). For example, most English monolingual learner dictionaries on CD-ROM now offer exercises and study pages focussing on topics as varied as common learner errors, phrasal verbs, pragmatics, and business English (e.g. Rundell 2002). They also sometimes include exam practice and lesson plans (e.g. Summers 2005).

We are currently adding a CALL component to the LEAD dictionary to help learners master academic vocabulary. We make use of the Hot Potatoes software (<http://hotpot.uvic.ca/index.php>) to create exercises such as fill-in-the-blanks, word building, error detection and collocation exercises. In the same way as the dictionary is customisable to users’ needs, our aim is to develop an adaptable learning environment that directs users to exercises that address their specific needs. Thus, a French-speaking user will have access to a number of exercises that deal with French learners’ specific difficulties (e.g. overuse of ‘in fact’ as a direct translation of ‘en fait’, heavy reliance on ‘let us’ to structure their academic texts). Exercises aimed at French users include, for example, error detection exercises that focus on L1-induced errors and translation quizzes in which the user must choose between several alternative English translations of a French sentence.

## Acknowledgements

We gratefully acknowledge the support of the Belgian National Fund for Scientific Research (FNRS), which funded this research within the framework of the FRFC project “Lexicography and phraseology: onomasiological and semasiological approach to English for Academic Purposes” (2.4.501.08.F).

## References

- Abel, A. (in press) Towards a systematic classification framework for dictionaries and CALL. In Granger, S. and Paquot, M. (eds), *eLexicography in the 21<sup>st</sup> century: Proceedings of eLex2009, Cahiers du Cental*, 7.
- Granger, S., Dagneaux, E., Meunier, F. and M. Paquot (2009) *The International Corpus of Learner English. Handbook and CD-ROM. Version 2*. Louvain-la-Neuve: Presses universitaires de Louvain.
- Granger, S. and Paquot, M. (in press) Customising a General EAP Dictionary to Learner Needs. In Granger, S. and Paquot, M. (eds), *eLexicography in the 21<sup>st</sup> century: Proceedings of eLex2009, Cahiers du Cental*, 7.
- Hartmann, R. R. K. (2005) Pure or hybrid? The development of mixed dictionary genres. *Facta Universitatis* 3(2): 193-208. Available from <http://facta.junis.ni.ac.rs/lal/lal2005/lal2005-06.pdf>
- Laufer, B. and Levitzky-Aviad, T. (2006) Examining the effectiveness of ‘bilingual dictionary plus’ – a dictionary for production in a foreign language. *International Journal of Lexicography* 19(2): 135-155.
- Pecman, M. (2008) Compilation, formalisation and presentation of bilingual phraseology: problems and possible solutions. In Meunier, F. and Granger, S. (eds), *Phraseology in Foreign Language Learning and Teaching*. Amsterdam & Philadelphia: Benjamins, 203-222.
- Rundell, M. (1999) “Dictionary use in production”. *International Journal of Lexicography* 12(1). 35-53.
- Rundell, M. (2002) (ed.) *Macmillan English dictionary for advanced learners* (CD-ROM). Oxford: Macmillan Education.
- Sobkowiak, W. (2002) What can be but is not (and why) in learners’ MRDs. *Teaching English with Technology* 2(3). Available at [http://www.iatefl.org.pl/call/j\\_article9.htm](http://www.iatefl.org.pl/call/j_article9.htm)
- Summers, D. (2005) *Longman dictionary of contemporary English. Writing Assistant Edition CD-ROM* Harlow: Pearson Education Limited.
- Tarp, S. (2009) Beyond lexicography: new visions and challenges in the information age. In H. Bergenholtz, S. Nielsen & S. Tarp (eds), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang, 17-32.
- Verlinde, S., Binon, J., Ostyn, S., and Bertels, A. (2007). La Base lexicale du français (BLF) : un portail pour l’apprentissage du lexique français. *Cahiers de lexicologie*, 91(2), 251-266.