

DICTIONARY News

What's in a name?

This is the twentieth issue of this newsletter, and our company is now in its twentieth year. The first issue of the newsletter was published by Kernerman Publishing and Password Publishers in July 1994. It was entitled *Password News*, and had the goal of serving as a “forum for discussion about the semi-bilingual English dictionary.” The title was changed in the next issue to *Kernerman Dictionary News*. Issues No. 2 and 3 appeared in 1995, and since then the newsletter has been published regularly in July each year. Over the years the scope of topics has expanded, and the look and size have changed as well.

Kernerman Publishing was established in 1969, as an independent ELT publishing house. Since the 1980s it has been conducted by my father, Ari (Lionel) Kernerman, who initiated the semi-bilingual dictionary, and became the leading English dictionary publisher in Israel and renowned globally.

Password Publishers was formed in 1993 to coordinate the growing series of Kernerman Semi-Bilingual Dictionaries worldwide. The name of the company was inspired by that of the *Password* dictionary, which has become synonymous with the semi-bilingual English learner's dictionary. Then in 2000 the company name was changed to K Dictionaries.

There were several reasons for this. To begin with, apart from exceptions, Password Publishers was not really a publishing house, but rather a content creator that cooperated with publishers (then with technology firms). As our focus was dictionaries, it seemed appropriate to reflect this in our name. Moreover, *Password* was no longer representative of our wider lexicography activity, which gradually involved other types of dictionaries and other languages. The natural

name replacement seemed to be Kernerman, which by that time had become well known in the dictionary world, but which I personally avoided — because that would be too close to the name of Kernerman Publishing, and because I considered that carrying my surname was somewhat vain. On the other hand, I liked the short form K as an abbreviation of Kernerman (and for 1,000), which was already in our logo, its anonymity, and the nice counterbalance it produced against the long word Dictionaries. Thus, K Dictionaries emerged — and may a thousand dictionaries bloom!

The irony of fate, however, is that on numerous occasions we are referred to as Kernerman Dictionaries, most notably by dictionary professionals... Accepting that there is no escape from your name, and assuming the weight it implies, we began to introduce *Kernerman* to our dictionary titles.

It first appeared in association with a local brand name for a series of French bilingual dictionaries published by Assimil in France in 2009; was co-branded with Houaiss for

Portuguese bilinguals in Brazil since 2010; added in our updated version of *Random House Webster's College Dictionary* last year; and will join Kenkyusha's name for a Japanese bilingual dictionary in Japan this year.

Kernerman was also applied to our 42-language English multilingual dictionary; to the Norwegian bilingual dictionary series published by Vega in Norway; to language versions of our new Global series, such as the Dutch dictionaries available on *Mijnwoordenboek.nl*; and to many mobile and tablet dictionary applications. Meantime, Kernerman Publishing also featured *Kernerman* in its new *Advanced English Dictionary*. ■ IK



- 1 What's in a name?
- 2 Compiling specialized dictionaries differently: A brief overview of terminological projects at the Observatoire de linguistique Sens-Texte (OLST) | **Marie-Claude L'Homme**
- 8 Using a specialized resource to enrich a general language dictionary | **Marie-Claude Demers**
- 12 Practical aspects of the description of terms: contexts, actantial structure and lexical relationships | **Geneviève Camirand**
- 15 Kernerman English French Learner's Dictionary: Adapting the translation from European French to Canadian/Québec French | **Suzanne Desgroseilliers**
- 21 A study of the third-generation Chinese-English dictionaries | **Xia Lixin**
- 26 Pedro A. Fuertes-Olivera and Henning Bergenholtz (eds.). *e-Lexicography. The Internet, Digital Initiatives and Lexicography* | **Adam Kilgariff**
- 29 Olga Karpova. *English Author Dictionaries (the XVIth – the XXIst cc.)* | **Kusujiro Miyoshi**
- 34 Olga Timofeeva and Tanja Säily (eds.). *Words in Dictionaries and History. Essays in honour of R.W. McConchie* | **Michael Adams**
- 39 *Grande Dicionário Houaiss da Língua Portuguesa* | **Mauro de Salles Villar**
- 40 Kernerman Norske Ordbøger

Editor | **Ilan Kernerman**



KDICTIONARIES

© 2012 All rights reserved.

K DICTIONARIES LTD

Nahum 8 Tel Aviv 63503 Israel

Tel: 972-3-5468102

Fax: 972-3-5468103

kdn@kictionaries.com

http://kictionaries.com

Compiling specialized dictionaries differently: A brief overview of terminological projects at the Observatoire de linguistique Sens-Texte (OLST)

Marie-Claude L'Homme

Marie-Claude L'Homme is full professor at the Department of Linguistics and Translation of the University of Montreal, where she teaches terminology and computer tools for translation. She is also the director of the Observatoire de linguistique Sens-Texte (OLST), a research group whose main focus is the lexicon (general and specialized) and the various theoretical and methodological issues related to its modelling. Her main research interests are lexical semantics applied to terminology and corpus-based terminology. Along with researchers in computer science, information science, terminology and linguistics, she develops lexical databases related to the fields of computing, the Internet and the environment.
mc.lhomme@umontreal.ca

1. Introduction

The Observatoire de linguistique Sens-Texte (OLST) is a research group located at the Department of Linguistics and Translation of the University of Montreal. It was created in 1997 with the aim of grouping experts with different backgrounds (linguistics, didactics, information science, terminology) interested in the various aspects of the lexicon. Its main objectives are the following:

- Address theoretical issues that underlie the various and complex properties of the lexicon.
- Develop and disseminate monolingual and multilingual resources (lexical and terminological databases as well as text corpora).
- Devise methods to apply this work to various applications (language teaching and information science, for instance).

The OLST also provides a multidisciplinary environment for training in lexicology, lexicography, and terminology. Several ongoing projects allow students to acquire the knowledge necessary to understand the theoretical issues related to the lexicon and the methodology necessary to compile lexical and terminological resources. In many cases, MA and PhD. students carry out research projects that are directly related to one of the ongoing projects at the OLST.

I focus here on two terminological projects that are in my opinion representative of the kind of environment and training that provides the research group. These projects and many other resources can be found on the OLST website, www.olst.umontreal.ca.

2. DiCoInfo and DiCoEnviro

DiCoInfo¹ and DiCoEnviro² are two lexical databases containing English, French and Spanish terms that are related to the fields of computing and the Internet (DiCoInfo) and to the field of the environment (DiCoEnviro), more specifically climate change and renewable energies. The descriptions provided in both resources are original in the sense that they aim to highlight various linguistic properties of verbs, nouns, adjectives and adverbs that convey a meaning that can be linked to either computing and the Internet (e.g.

administrator, download, graphical, dynamically) or to the environment (e.g. *biomass, reduce, green, globally*). As can be seen in Figures 1 and 2 with the entries devoted to *warm*, the following information is provided:

- Clearly defined semantic distinctions: Figure 1 shows that the inchoative and causative meanings of the verb *warm* appear in two different entries. In fact, an additional entry describes the adjective *warm*: WARM₂, adj: ~ Patient{climate 1}. In DiCoInfo, polysemous lexical items are also described in separate entries. For instance, *download* appears in three different entries:

DOWNLOAD₁, vt: **Agent**{user 1} ~ **Patient**{application, file 1} from **Source**{computer 1, network} to **Destination**{computer 1} (*download hwcLEAR.exe from Hauppauge's website*)

DOWNLOAD_{1,1}, n: ~ of **Patient**{application, file 1} from **Source**{computer 1, network} to **Destination**{computer 1} by **Agent**{user 1} (*a download that never finishes or a similar problem*)

DOWNLOAD_{1,2}, n: ~ of **Patient**{application, file 1} used by **Agent**{user 1} (*the download will be an executable file*)

- A number (status 2, 1 or 0) indicates how advanced the writing of the entry is. Entries with a status 0 (only available in the French version of DiCoInfo) are the most complete. Entries with a status 2 contain valid information, but are lacking some lexical relations and definitions.
- The actantial structure of terms: actants (i.e. arguments) are specified with two different labels (semantic roles such as **Agent, Cause, Patient**) and the typical term that can instantiate an actant (between curly brackets).
- Linguistic realizations of actants (Figure 2): a list of terms that can be found in running text and that can instantiate actants.

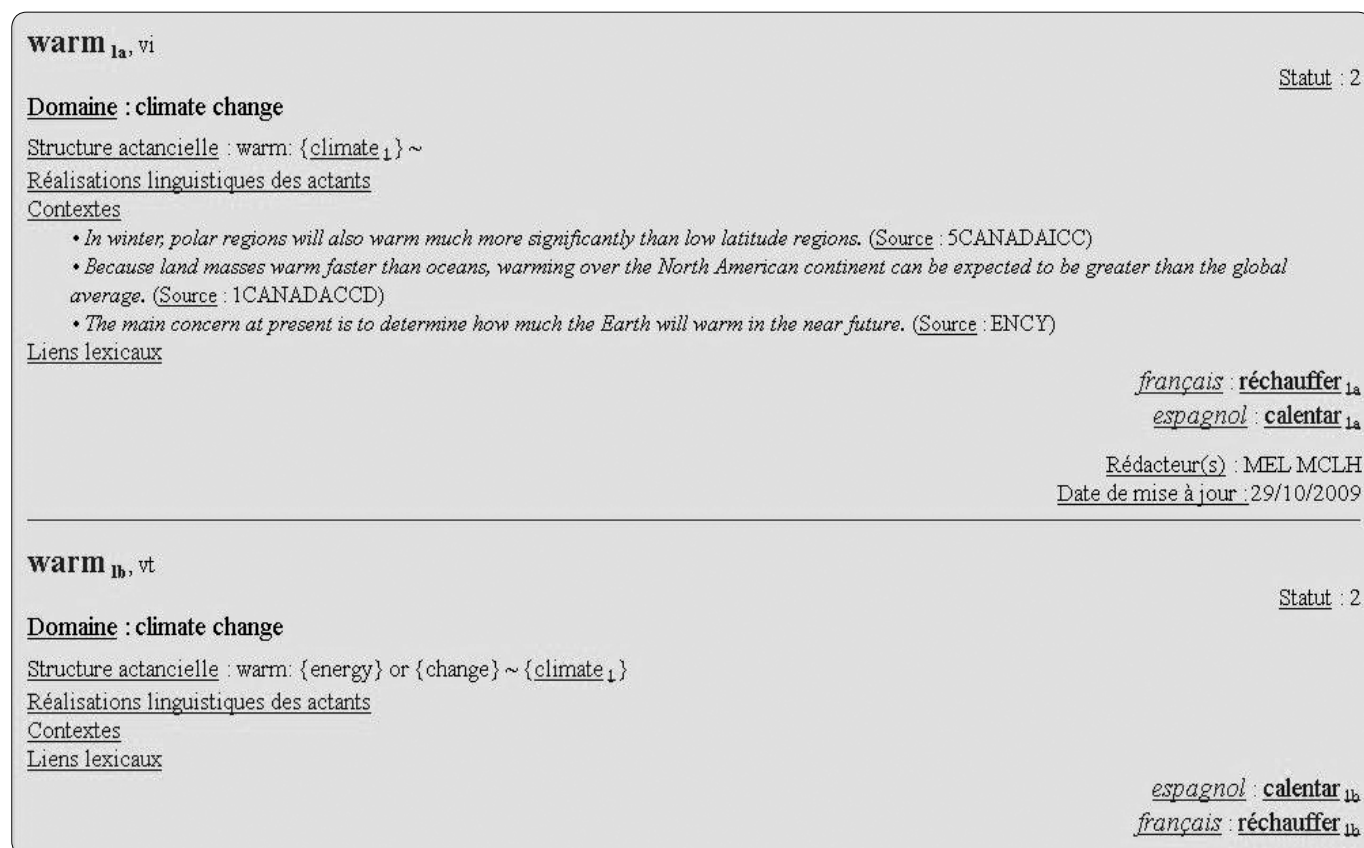


Figure 1: Entries WARM_{1a} and WARM_{1b} extracted from DiCoEnviro.

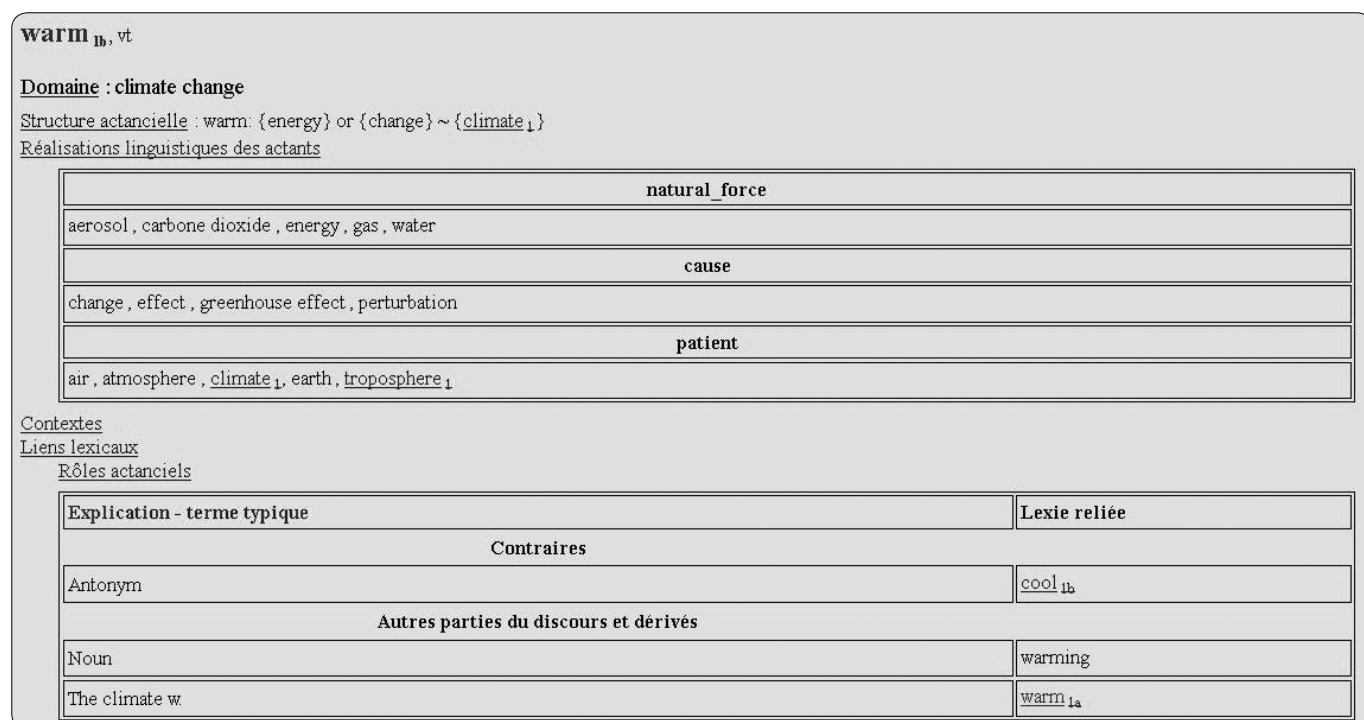


Figure 2: Linguistic realizations of actants and lexical relations in WARM_{1b} (DiCoEnviro).

- Contexts: a selection of sentences extracted from the corpora used by lexicographers to write the entries.
- Links to French and Spanish equivalents when available online.
- A list of lexical relations (Figure 2): other terms that share with the headword a semantic relation such as hypernymy, near synonymy, antonymy, collocation. In some entries, the list of lexical relations is quite long. For instance, in entries for nouns, collocates are listed.
- In some of the entries, definitions are provided. We display them in entries labelled with a status 0.

3. Resources designed according to lexico-semantic frameworks

Both resources are based on lexico-semantic frameworks that were originally designed to account for the general lexicon but that can be adapted to specialized lexical units. The first framework is that provided by Explanatory Combinatorial Lexicology (ECL, Mel'čuk et al. 1995). In DiCoInfo and DiCoEnviro, lexicographers refer to ECL when making semantic distinctions, defining the actantial structure, and listing lexical relationships.

The second theoretical framework applied in our resources is Frame Semantics (Fillmore 1982) and more specifically its

application in FrameNet (Ruppenhofer et al. 2010). A module was recently added to both DiCoInfo and DiCoEnviro that shows how the term appearing as the headword interacts with its participants (actants and circumstants) in running text. Figure 3 shows part of the annotations and the summary table for the term ATTACH₁.

Of course, some adaptations were made to both frameworks when applied to the description of specialized terms.

4. A corpus-based methodology

It can easily be inferred from what has been said up to now that the methodology devised to compile DiCoInfo and DiCoEnviro is

Attach File - Click on this button if you want to send this email with a <i>file ATTACHED</i> to it. [INTERNET 0 JP MCLH 22/02/2009]		
<i>You can ATTACH more than one file to a single email message</i> [INTROCOMP 0 JP 22/02/2009]		
Email documents are created as text files so in order to send a binary file or document via email, <i>it</i> must first BE encoded into a text format and then ATTACHED to the email text <i>message</i> . [INTROCOMP 0 JP 22/02/2009]		
Both methods allow <i>users</i> to sign or ATTACH a digital <i>identification</i> to the email <i>message</i> which verifies, to the recipient, that the message is from the original person or organization and that the information wasn't tampered with in transit. [INTROCOMP 0 JP 22/02/2009]		
Netscape Communicator indicates that a <i>file IS ATTACHED</i> to a <i>message</i> by displaying a paperclip icon in the message window. [WIREDGUIDE 0 JP 22/02/2009]		
ATTACH 1		
Actants		
Patient	Object (NP) (8) Subject (NP) (5)	file (9) document (2) it identification
Destination	Complement (PP -to) (14)	message (11) posting email it

Figure 3: Annotated contexts in ATTACH₁ (DiCoInfo).

the radiative forcings of all the trace gases (CO ₂ , CH ₄ , e: water vapour, carbon dioxide, methane, nitrous oxide, occurring greenhouse gases: water vapour, carbon dioxide, by these trace components: water vapour, carbon dioxide, ffects from changes in concentrations of carbon dioxide, use gases include water vapour, carbon dioxide, methane, ated if other forcing factors such as solar irradiation, ated if other forcing factors such as solar irradiation, one concentrations are also spatially variable. Further, eric components that are rapidly recycled: water vapour, oping countries (section 3.6 in the TS). 2. Halocarbons, quality, and future availability issues. 2. Halocarbons, f the cooling effect and of reducing ODS concentrations, hings, emissions of CH ₄ and pollutants (as noted below). e third most important greenhouse gas after CO ₂ and CH ₄ . etres higher (at about 17 to 20 km) in the stratosphere. ents because they deplete the stratospheric ozone layer. NO _x lead to decreases in the stratospheric ozone layer. everywhere and is independent of where emissions occur. ivities assuming the mid-range IPCC scenario (IS92a). 4.2 with reduced or negligible global warming potentials. 4 all compared to those for greenhouse gases. Explanation:	ozone (03), etc.) and aerosols. The total forcing may be treated as ozone, and halocarbons. Climate Trends An analysis of temperature ozone, methane and nitrous oxide. The amount of radiative forcing of ozone and methane. It is this dependence on the minor components th ozone, methane, water vapour, line-shaped contrails, and aerosols, ozone and nitrous oxide. The magnitude of the natural greenhouse ef ozone depletion and non-sulphate aerosols are added. Once the credi ozone depletion and non-sulphate aerosols are added. Once the credi ozone is not a directly emitted species, but rather it is formed in ozone and aerosols (liquid or solid suspensions in the atmosphere). ozone depletion and climate change 2.1 How do the CFCs and their re ozone depletion and climate change 2.1 What are ozone is likely to increase over much of the stratosphere, but coul Ozone concentrations respond relatively quickly to changes in the e Ozone is formed by photochemical reactions and its future change wi Ozone in the upper troposphere and lower stratosphere is expected t Ozone in the troposphere (the lower part of the atmosphere) is anot Ozone precursor (NO _x) residence times in these regions increase wit Ozone also differs from the other greenhouse gases in that it is no Ozone The NO _x emissions from subsonic aircraft i Ozone within this report refers to stratospheric ozone unless other Ozone depletion allows more ultraviolet radiation to reach the low
------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Figure 4: Concordances for *ozone*.

heavily corpus-based. For each resource, corpora had to be compiled in English, French and Spanish according to criteria adapted to terminology work (Bowker and Pearson 2002). In addition, corpora are enriched periodically to ensure that they are up-to-date.

Lexicographers use an in-house concordancer for obtaining sentences in which relevant terms appear (Figure 4). Between 15 and 20 contexts are selected and placed in entries. Contexts are then annotated (cf. Section 3). Lexicographers find most of the information necessary to fill the data categories of the resources in corpora. However, in some cases, they must also refer to existing specialized dictionaries or experts to validate a piece of information that was found in the corpus or access information that is not readily available in running text.

5. A computer-assisted process

Nearly all steps required to compile entries in DiCoInfo and DiCoEnviro are computer-assisted. We already referred to the use of a concordancer for finding relevant contexts. However, many other stages are partly automated.

The selection of terms to include in the word list is carried out using a term extractor called TermoStat and designed by Drouin (2003). The programme provides a list of candidate terms found in corpora based on a statistical calculation designed to define the specificity of lexical item in a specialized corpus when compared to a corpus of a different nature. The most recent version of TermoStat has many other features that allow lexicographers to have different views on the data (grouping of terms according to their components, identification of potential actants, etc.).

In the DiCoInfo and DiCoEnviro projects, lexicographers use the list of candidate terms generated by TermoStat to make a first selection of terms that will appear in the word lists of each resource. Figure 5 shows some of the results obtained when submitting an English corpus of climate change texts to the program. *Climate*, *change*, and *emission* were identified as the most specific terms in the corpus and now appear in DiCoEnviro.

Terms are then placed in an XML structure containing all the labels for data categories that lexicographers will fill with information based on what can be found in the corpora and their own knowledge of the fields.

6. A step-by-step training process

In order to compile entries in DiCoInfo and DiCoEnviro, lexicographers must become

acquainted with some principles of ECL and Frame Semantics, be comfortable with the different computer programs used throughout the process (term extractor, concordancer, XML structure, etc.) and acquire some confidence as to their own judgements and intuitions about the meaning of terms. All this knowledge cannot be acquired at once, especially in our environment where most people working within our projects are students in translation who have no prior practical training in lexicography. It must also be pointed out that students work part time on the projects.

In order to ease the learning curve, a step-by-step process was defined. It usually flows as follows:

- Read a couple of in-house documents to get a general overview of the methodology.
- Start collecting contexts and place them in relevant entries.³ In doing so, lexicographers must make semantic distinctions and often create new entries to reflect the polysemous nature of some lexical items. Lexicographers can carry out this work for two to three weeks before they start adding information to other data categories.
- Collect true synonyms and graphical variants: these can be found in existing reference works.
- Define the actantial structure: when they become comfortable with semantic distinctions, lexicographers are asked to

The Observatoire de linguistique Sens-Texte (OLST), a research group located at Université de Montréal, began to cooperate with K Dictionaries (KD) in 2010. Four OLST members have contributed articles to this issue of KDN:

- Marie-Claude L'Homme gives a general overview of the work carried out at the OLST and focuses on two specific terminological projects: DiCoInfo, *Dictionnaire fondamental de l'informatique et de l'Internet*, and DiCoEnviro, a database with terms related to the field of the environment.
- Marie-Claude Demers presents a methodology for identifying lexical items and meanings specific to the fields of computing and the



Lemmatized candidate term	Frequency	Weight	Variants
Climate	5020	235,0826658	climate__climates
Change	4556	188,691871	Changes
Emission	3049	182,802937	Emissions
Global	1667	128,6522177	Global
Temperature	1733	126,2666191	temperature__temperatures
Model	1669	117,5733414	Models
Scenario	1228	114,4488449	Scenarios
Carbon	1268	114,0777962	Carbon
Greenhouse	1299	113,2565932	greenhouse__greenhouses
Gas	1654	111,8322141	Gases
Concentration	1088	102,5969814	Concentrations
Ocean	1041	101,0607715	Oceans
Impact	1215	98,62949399	Impacts
Atmosphere	1017	94,90083166	Atmosphere
Warming	853	94,18214054	Warming

Figure 5: Candidate terms extracted from a corpus of climate change.

Internet. This project was carried out within the cooperation framework for the interchange of data between a general language dictionary, the *Random House Kernerman Webster's Collegiate Dictionary* (RHKWCD), and DiCoInfo, and is part of her MA dissertation.

- Geneviève Camirand explains how new lexical items and meanings are described when added to DiCoInfo, within the same framework of interchange with RHKWCD. The format of DiCoInfo is not entirely compatible with what is expected to be found in RHKWCD, but some parts of the entries can later be adapted to the specific requirements of a general dictionary.
- Suzanne Desgroseilliers describes the work she carried out in order to adapt the equivalents provided in an English-French dictionary to the French used in Québec. This project was undertaken within an internship program offered by KD, with the objective of providing a hands-on experience in lexicography.

Another angle of the cooperation between OLST and KD, touching also on the import of entries from RHKWCD to DiCoInfo, is presented in the paper by Demers et al. at Euralex 2012.

Special thanks to Marie-Claude L'Homme for helping with the publication of these articles.

start defining the actantial structures of terms. They normally start with verbs; then they define the structures for adjectives and nouns.

- e) Annotate contexts according to the methodology defined within the FrameNet project: this requires a specific training in order to follow a strict annotation process that encodes semantic and syntactic information about terms and their participants.
- f) Establish equivalence relationships between English, French and Spanish entries.
- g) Collect other semantically related terms: antonyms, near synonyms, collocates, etc.
- h) Encode lexical relations using the system of lexical functions provided by ECL (Mel'čuk et al. 1995).

During each step, students are asked to note all the questions they may have and these are discussed with a more experienced lexicographer. The latter also revises the data categories periodically and decides when an entry can be placed online.

7. And much more

The previous sections offered a quick overview of two terminological projects carried out at the OLST and show how the various linguistic properties of specialized terms can be described in online resources. Recently, our projects have attracted the interest of researchers working in areas different from terminology and of lexicographers working on the general lexicon. Since both resources rely on lexico-semantic frameworks and on a methodology that is very close to the ones used in standard lexicography, they seem to lend themselves to extensions that we did not foresee when we first started compiling them. This section presents some of these extensions.

It soon became obvious that our resources, especially in the case of DiCoInfo, which has a larger coverage than DiCoEnviro, could be compared to general lexical resources in order to find lexical items or meanings that are specific to specialized domains and that might be lacking in more general repositories. We first carried out a comparison with a lexical resource that is located at the OLST, namely the DiCo (Jousse and Polguère 2005)⁴ that is also compiled according to the theoretical and methodological principles of ECL. This comparison led to a series of criteria that can be taken into account when adding meanings related to specialized fields of knowledge to a general resource (L'Homme and Polguère 2008). Another comparison was carried out between the English

FrameNet and the English version of DiCoInfo to find meanings that could be missing in the general language resource. We found that most meanings covered in DiCoInfo could be considered for inclusion in FrameNet (Pimentel et al. 2012).

The “lexicographic” potential of DiCoInfo also led to another interesting project. With K Dictionaries, we devised a method for interchanging data taken from the English version of DiCoInfo and from the *Random House Kernerman Webster's College Dictionary* (RHKWCD). The wordlists of each resource were compared semi-automatically and this comparison led to the identification of missing lexical items or meanings in each resource (Camirand, herein; Demers, herein; Demers et al. 2012). First, terms such as *avatar*, *artificial intelligence* and *google* (verb) were added to DiCoInfo. Other terms or specific meanings, such as *arrow key*, *attach* (as in *attach a file to an email*) and *data-driven*, are considered for inclusion in RHKWCD. When introduced in each resource, entries must be written according to their respective style guides.

We recently designed an interface to access the varied general and specialized resources available at the OLST in order to provide a first glance at the different meanings that may have a lexical item in general language as well as in specific subject fields. The interface, called Olster⁵, searches the various resources and extracts the entries and an example for each entry. Users can then access the entry as it appears in the resources.

The last project that will be presented here concerns the various adaptations that were made to DiCoInfo and DiCoEnviro to make them more accessible to users who do not have a background in ECL or Frame Semantics. The resources were initially designed as research environments allowing researchers and students to carry out different kinds of analyses on terminological data. However, some colleagues pointed out that some aspects of the presentation of the data in the online versions of the resources and the access to their various data categories could be modified so as to make them more compatible with specific user needs (L'Homme et al. 2012). This work led to changes in the display of information on-screen and to the addition of new search functions. For instance, users can now access translations of collocations (e.g. *send sth as an attachment* -> *envoyer qqch. en pièce jointe*). A browsing module was also introduced in the French version of DiCoInfo that allows users to access a collocate that expresses a specific meaning. More changes will be introduced in the near future.

Notes

¹ DiCoInfo can be accessed here: <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi/>.

² DiCoEnviro can be accessed here: http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search_enviro.cgi/.

³ In both projects, a preliminary list of terms is provided to lexicographers. When they start working on the projects, they are not asked to select terms themselves. However, in the course of their work, they may find that some terms are missing and add them to the word list.

⁴ <http://olst.ling.umontreal.ca/dicouebe/>.

⁵ Olster was designed by Benoît Robichaud, research assistant at the OLST: <http://olst.ling.umontreal.ca/olster/>.

References

- Bowker, L. and Pearson, J. 2002.** *Working with Specialised Languages: A Practical Guide to Using Corpora*: London: Routledge.
- Demers, M.-C., Kernerman, I. and L'Homme, M.-C. 2012.** Lexicographic interchange between a specialized and a general language dictionary. In *Euralex 2012 Proceedings*.
- Drouin, P. 2003.** Term Extraction Using Non-Technical Corpora as a Point of Leverage. *Terminology*, 9.1: 99-117.
- Fillmore, C.J. 1982.** Frame Semantics. In *The Linguistic Society of Korea (ed.), Linguistics in the Morning Calm*. Seoul: Hanshin, 111-137.
- Jousse, A.L. and Polguère, A. 2005.** *Le DiCo et sa version DiCouèbe. Document descriptif et manuel d'utilisation*. Montréal: Observatoire de linguistique Sens-Texte (OLST).
- L'Homme, M.-C. and Polguère, A. 2008.** Mettre en bons termes les dictionnaires spécialisés et les dictionnaires de langue générale. Maniez, F. and Dury, P. (dir.), *Lexicographie et terminologie : histoire de mots. Hommage à Henri Béjoint*. Lyon: Presses de l'Université de Lyon, 191-206.
- L'Homme, M.-C., Robichaud, B. and Leroyer, P. 2012.** Encoding collocations in DiCoInfo: from formal to user-friendly representations. In Granger, S. and Paquot, M. (eds.), *Electronic Lexicography*. Oxford: Oxford University Press, 209-234.
- Mel'čuk, I., Clas, A. and Polguère, A. 1995.** *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Duculot / Aupelf - UREF.
- Pimentel, J., L'Homme, M.-C. and Laneville, M.E. 2012.** General and Specialized Lexical Resources: A Study on the Potential of Combining

Efforts to Enrich Formal Lexicons. *International Journal of Lexicography*, 25.2: 152-190.

Ruppenhofer, J., Ellsworth, M., Petruck, M.R.L., Johnson, C. and Scheffczyk, J. 2010. *FrameNet II: Extended Theory and Practice*. 12 November 2011. <http://framenet.icsi.berkeley.edu/>.

ASIALEX 2013 in Bali

The 8th International Conference of the Asian Association for Lexicography (AsiaLex) will be held in Bali, Indonesia, on 20-22 August 2013. Preparations are underway, the conference website Asialex2013.org is up and running, and the Call for Papers will be published shortly.

The conference theme is Lexicography and Dictionaries in the Information Age. It is expected to draw the participation of not only lexicographers but also linguists, translators, teachers, and others interested in lexicography and dictionaries. The featured speakers include Dr. Diah A. Arimbi (Indonesia), Prof. Henning Bergenholtz (Denmark), Dr. Adam Kilgarriff (UK), Prof. Robert Lew (Poland), and Prof. Yukio Tono (Japan). In addition to the parallel paper sessions, there will be special sessions for software developers and for publishers interested in presenting the innovative features of their products.

The conference program is organized by the Secretary of AsiaLex, Dr. Deny Kwary, and his team from Airlangga University, in Surabaya, with the professional assistance of a local event organizer. Bali is well-known for its beautiful beaches, so the conference venue will be a beach-front hotel, and the conference will include a rich social program in addition to the academic presentations.

It is both a challenge and an opportunity to conduct a lexicography conference in Indonesia. On the one hand, lexicography is under-developed. For example, the biggest national language dictionary, *Kamus Besar Bahasa Indonesia* (4th edition, 2008), contains only 41,250 lemmata and 48,799 sublemmata, and consists of a single volume with the total of 1,701 pages. On the other hand, Indonesia is a vast country, with a large number of languages, and many foreign visitors: it has over 260 million inhabitants, 746 local languages, and approximately 600,000 tourists coming from abroad every month (more than 200,000 to Bali). Therefore, dictionaries should play an important role here.

Most Indonesians still use old versions of dictionaries, and are not aware of the latest developments and innovative features of modern dictionaries. Dictionaries are not compiled with the help of corpora and dictionary writing systems, but are usually written by using simple word processors. Therefore, ASIALEX 2013 will offer special sessions for software developers to promote their dictionary software and for publishers to promote their up-to-date dictionaries.

Deny A. Kwary

Using a specialized resource to enrich a general language dictionary

Marie-Claude Demers



Marie-Claude Demers has a Bachelor's degree in German studies and a Master degree in Translation from Université de Montréal. During the course of her studies, she contributed to enriching the English version of the DiCoInfo at the Observatoire de linguistique Sens-Texte (OLST). She now works as a translator for a Canadian retailer and pursues translation as a freelancer. marie-claude.demers.2@umontreal.ca

Abstract

This paper describes a method for comparing a specialized lexicographic resource and a general one, thus evaluating the extent to which the former can contribute to increase the coverage of the latter. Concretely, it compares the contents of the English wordlist of the *Dictionnaire fondamental de l'informatique et de l'Internet* (DiCoInfo), developed at the Observatoire de linguistique Sens-Texte (OLST), and the wordlist of the *Random House Kernerman Webster's College Dictionary* (RHKWCD), of K Dictionaries. Firstly, the entries from both resources were automatically extracted and compared. Then, we carried a manual analysis of every lexical item that we classified in different categories according to their presence in RHKWCD and the way they are described in it. Based on this research, recommendations regarding ways of improving the integration of specialized units in a general language dictionary were made. Overall, this paper concludes that both lexical resources are compatible and that it is possible to incorporate information recorded in a specialized resource into a general one. (Parallel research to extract lexical units from RHKWCD and record them in DiCoInfo has demonstrated that the reverse is possible as well.)

Keywords: terminology, lexicography, general language, computing language

1. Introduction

Over the last three decades, computational linguistics has evolved constantly, providing an increasing number of tools—such as term extractors, database management software, concordancers, and translation memories—that accelerate, automate, and ease the work of linguists, terminographers, lexicographers, and translators. There is undoubtedly an infinite amount of data compiled by organizations, companies, institutions, and individuals, creating the possibility of sharing research findings and information. This issue is at the core of this project on bilingual lexicography that studies the compatibility between lexicography and terminology, examined by different researchers such as

Cabré (2007) and Béjoint (2007). More precisely, by focusing on the integration of terms in a general language dictionary, we compare the wordlist of the *Dictionnaire fondamental de l'informatique et de l'Internet* (DiCoInfo), developed at the Observatoire de linguistique Sens-Texte (OLST), with that of the *Random House Kernerman Webster's College Dictionary* (RHKWCD)¹, of K Dictionaries. We begin by presenting each resource. Then, we explain the different steps taken to extract, analyze and select relevant data from these dictionaries. Finally, we conclude with some observations and recommendations as to how information found in a specialized resource can be incorporated correctly in a general language dictionary.

2. Presentation of DiCoInfo and RHKWCD

RHKWCD was originally published in 1947 under the name of *American College Dictionary* (Demers et al. 2012). The dictionary was revised and updated annually, was eventually retitled *Random House Webster's College Dictionary*, and K Dictionaries acquired the last version published in 2005. RHKWCD is intended for college students and the general public, native English speakers and advanced non-native users. Today, it comprises approximately 130,000 words and expressions from all language ranges. Common meanings are ordered before specialized ones and frequent units appear before older ones. The entries include pronunciation, definitions, and examples of usage, as well as information on etymology and usage. Although RHKWCD and the DiCoInfo compile different types of information, they share a common encoding system both using a markup language to record data.

The DiCoInfo is a specialized dictionary created by the Observatoire de Linguistique Sens-Texte at the Université de Montréal. It is a free online resource, focusing on terms related to the fields of computing and the Internet. Its objectives are to describe fundamental terms, such as *email*, *bug* and *network*, as well as to list and explain the relations between the terms of the field. When compiling the entries, terminographers refer

to a corpus that has more than a million words containing mainly pedagogical texts dealing with topics such as the Internet, networks, programming, micro-computing, and operating systems.

The records in the DiCoInfo are divided into sections. The sections *headword*, *part of speech*, *status*, *actantial structure*, *written by*, and *last update* appear in every entry. The sections *synonym(s)*, *linguistic realization of actants*, *contexts*, *variant(s)*, and *French* are shown by default, but only in records for which the information is available. The section *definition* is only provided for records of status 0. Figure 1 illustrates how records are written in the DiCoInfo.

Based on the Explanatory Combinatorial Lexicology method (Melčuk 1999), this resource is still under construction and enriched on an ongoing basis. Some records are complete and available online, while others appear in the wordlist but still need to be compiled (presenting only a few contexts and including no actantial structure). The achievement level of records is indicated by a status number that ranges from 0 to 3. Completed records are attributed the number 0.

3. Extraction, analysis and integration of data

We started the project by automatically extracting all entries from both dictionaries and comparing them. This was facilitated by the fact that both resources are encoded

in XML. The first step consisted of identifying lexical units that were not recorded in RHKWCD. Since many items are polysemous, we also had to carry out a manual analysis of every meaning defined under each dictionary entry of RHKWCD. The entries of the DiCoInfo were classified in one of the following categories depending on how they were taken into account in RHKWCD. We give below (Tables 1 to 12) the description and an example of a lexical unit for each of the six categories.

Table 1. Category A1

Category	A1
Criterion	A lexical item that is listed in RHKWCD with a clear indication that it belongs to the field of computing, such as a usage label or the presence of the word <i>computer</i> in its definition.
Quantity	302

Table 2. Example of a lexical unit from category A1: *batch1*

Lex. Unit	batch1
Part of sp.	noun
Definitions in RHKWCD	<ol style="list-style-type: none"> 1. a quantity or number coming at one time or taken together; group; lot: <i>a batch of prisoners</i>. 2. the quantity of bread, dough, etc., made at one baking: <i>a batch of cookies</i>. 3. the quantity of material prepared or required for one operation: <i>to mix a batch of concrete</i>. 4. a group of jobs, data, programs, or commands treated as a unit for computer processing. 5. a. a quantity of raw materials mixed in proper proportions and prepared for fusion into glass. b. the material so mixed.

software ₂ , n		Status : 2
Actantial structure : software: ~ used by { user ₁ } to act on { task ₁ } on { computer ₁ }		
Linguistic realizations of actants		
Contexts		
Lexical relations		
Actantial roles		
Explanation - Typical term	Related term	
Opposites		
Contrastive	hardware ₁	
Types of		
That is created for commercial purposes	commercial ~	
That is free	freeware ₁	
That is shared by a number of users	shareware ₁	
Combinations		
Someone or something creates the s.	develop ₁ ~	
The user prepares the s. to allow the s. to operate	install ₂ ~	
-> NOUN	installation ₂ of ~	
The s. operates on a specific	the ~ runs ₁₂ on ...	
French : logiciel ₂		
Written by : LPD AB MCLH		
Last update: 06/12/2010		

Figure 1: Record of SOFTWARE₂

Table 3. Category A2

Category	A2
Criterion	A lexical item that appears in RHKWCD with a meaning that clearly belongs to computing but without any clear indication, such as a usage label or the presence of the word <i>computer</i> in its definition.
Quantity	43

Table 4. Example of a lexical unit from category A2: *bot1*

Lex. unit	bot1
Part of sp.	noun
Definitions in RHKWCD	1. the larva of a botfly. 2. a device or piece of software that can execute commands or perform routine tasks, as electronic searches, usually without user intervention (often used in combination): <i>intelligent infobots</i> ; <i>shopping bots</i> .

Table 5. Category B

Category	B
Criterion	A lexical item that is listed in RHKWCD, but the computing meaning is not recorded.
Quantity	273

Table 6. Example of a lexical unit from category B: *script1*

Lex. unit	script1
Part of sp.	noun
Definitions in RHKWCD	1. the letters or characters used in writing by hand; handwriting. 2. a manuscript or document. 3. the written text of a play, motion picture, television program, or the like. 4. any system of writing. 5. Print. a type imitating handwriting. 6. a plan.

Table 7. Category B-C

Category	B-C
Criterion	A lexical item that belongs to both categories B and C.
Quantity	193

Table 8. Example of a lexical unit from category B-C: *partition1*

Lex. unit	partition1
Part of sp.	noun
Definitions in RHKWCD	1. a division into or distribution in portions or shares. 2. a separation, as of two or more things. 3. something that separates or divides. 4. a part, division, or section. 5. an interior wall or barrier dividing space into separate areas. 6. Logic. the separation of a whole into its integral parts. 7. Math. a mode of separating a positive whole number into a sum of positive whole numbers.

Table 9. Category C

Category	C
Criterion	A lexical item that is recorded in RHKWCD whose general language meaning is applicable to its meaning in the field of computing.
Quantity	111

Following this analysis, two types of lexical units were considered for inclusion into RHKWCD: units absent from RHKWCD and units that are present in RHKWCD but that do not convey a meaning related to the field of computing. Therefore, we selected units labeled B and D. After having selected which units could be added, we had to determine how these could be integrated based on their presence in RHKWCD and how they are described in this resource.²

4. Observations and recommendations

A few scenarios were identified based on the analysis in section 2. We made recommendations for the inclusion of lexical units depending on the category in which they were classified.

When a term is used exclusively in a specialized context, L'Homme and Polguère (2008) recommend adding a label indicating the field to which it belongs. Lexical units from category D, which belong exclusively to the field of computing, should be accompanied by such a label. However, if the field is clearly indicated in the definition, for example when the word *computer* is mentioned, as in cases A1, the label becomes superfluous (Josselin-Leray and Roberts 2004).

Although lexical units B appear in RHKWCD, no meaning from the field of computing is recorded. To integrate these lexical units into the dictionary, lexicographers could simply add a new meaning.

Lexical units in groups B-C and C are listed in RHKWCD and their definition could also apply to computing. In many cases, in addition to conveying a general meaning, those units also cover a terminological usage. To illustrate how these lexical units are used in the field of computing, sentences from that domain may be added in the form of examples after the general language definition.

Figure 2 presents an example of how this latter scenario applies. In the field of computing as well as in general language, the verb *decipher* means “to decode a message”. In computing, that message is in an electronic format and is decoded with a key. This figure shows the three meanings of *decipher* as listed in RHKWCD. We added an example (in boldface) to illustrate the usage of the verb in computing.

Three criteria motivate the inclusion of specialized units in general language dictionaries: the level of specialization of the term, the nature of the term (single-word or multi-word unit) and morphological relations between lexical units (Josselin-Leray and Roberts 2004). First, lexicographers prefer less specialized terms than specialized ones since the former

de•ci•pher/dr'sai fər/ *v.t.*

1. to make out the meaning of (something obscure or difficult to read or understand): *I couldn't decipher his handwriting.*

2. to interpret by the use of a key, as something written in cipher: *to decipher a secret message, to decipher data with an algorithm.*

3. Obs. to depict; portray.
[1520–30; MF *déchiffrer*]

Figure 2: *decipher* recorded in RHKWCD

are more likely to be relevant for a vast audience. Furthermore, they prefer single terms over multi-word units. Lastly, they not only consider the meaning of the unit but also its formal resemblance to other units and would rather work on a group of units than on individual units.

5. Conclusion

In this study, we evaluated whether it is possible to use a specialized lexical resource, the DiCoInfo, to enrich a general language dictionary, namely RHKWCD. We compared the wordlist of the DiCoInfo with that of RHKWCD. We then proceeded by classifying each lexical unit found in the DiCoInfo into six different categories according to the way they were taken into account in RHKWCD. Lexical units in categories A1 and A2 were described in RHKWCD and it was obvious that they belonged to the field of computing. Only lexical units B and D were considered for inclusion into RHKWCD. The former were present in the dictionary but a computing meaning had to be added, while the latter were completely absent from it. For lexical units in B-C and C, examples could be added after the definitions to show that the lexical units are used in the field of computing, as demonstrated by the term *decipher*. We thus showed that it is possible to use an existing specialized resource to increase the coverage of a general language dictionary. We also provided a few guidelines on how to proceed based on the presence and the type of definition of units in the general language dictionary.

According to L'Homme and Polguère (2008), lexical units should be selected based on the target audience of the resource. However “different users require different things from their dictionaries, but even where dictionaries set out to address similar userships, there are discrepancies between the levels of information and kinds of detail for scientific and technical words or meanings.” (Moon 2008) How can we determine which units are relevant to the target audience and which information

Table 10. Example of a lexical unit from category C: *decipher1*

Lex. unit	decipher1
Part of sp.	transitive verb
Definitions in RHKWCD	1. to make out the meaning of (something obscure or difficult to read or understand): <i>I couldn't decipher his handwriting.</i> 2. to interpret by the use of a key, as something written in cipher: to decipher a secret message. 3. Obs. to depict; portray.

Table 11. Category D

Category	D
Criterion	A lexical item that is not listed in RHKWCD.
Quantity	421

Table 12. Example of a lexical unit from category D: *server machine1*

Lex. unit	server machine1
Part of sp.	noun
Definitions in RHKWCD	Not applicable.

should be added in the definition? To answer this question, we suggest selecting the units that are to be added to the general language resource based on their occurrences in a general language corpus. This allows us to objectively determine which specialized unit is now part of the general language vocabulary. For instance, in this study, we could have selected which computing lexical units labeled as B and D are part of the general language and should be added to RHKWCD based on their level of occurrence in a general language corpus. The following step would be to decide which information the definitions of those computing lexical units should provide.

Notes

¹ RHKWCD consists of the core of *Random House Webster's College Dictionary* (Random House, New York, 2005) updated by K Dictionaries.

² A total of 1,353 lexical units were analyzed, although DiCoInfo does not contain as many records. This occurred because many lexical units are synonyms or variants, and many terms have a compositional meaning. Multi-word terms containing the word *internet*, such as *internet site*, *internet access*, *internet browser* and *internet network*, can be quoted as examples. Thus, although 273 lexical units pertained to category B and 421 units were labeled as D, not as many units could be added to RHKWCD. By excluding variants, the number of potential entries that could be added to RHKWCD decreased.

References

- Béjoint, H. 2007.** Nouvelle lexicographie et nouvelle terminologie : convergences et divergences. In L'Homme, M.-C. and Vandaele, S. (eds.), *Lexicographie et terminologie : compatibilité des modèles et des méthodes*. Ottawa: Les Presses de l'Université d'Ottawa, 29-78.
- Cabré, T. 2007.** La terminologie : une discipline en évolution : le passé, le présent et quelques éléments prospectifs. In L'Homme, M.-C. and Vandaele, S. (eds.), *Lexicographie et terminologie : compatibilité des modèles et des méthodes*. Ottawa: Les Presses de l'Université d'Ottawa, 79-109.
- Demers, M.-C., Kernerman, I. and L'Homme, M.-C. 2012.** Lexicographic interchange between a specialized and a general language dictionary. In *Euralex 2012 Proceedings*.
- DiCoInfo: Le dictionnaire fondamental d'informatique et de l'Internet.** 2011. <http://olst.ling.umontreal.ca/cgi-bin/DiCoInfo/search.cgi/>.
- Josselin-Leray, A. and Roberts, R.P. 2004.** Le traitement des termes dans les dictionnaires généraux. In Béjoint, H. and Maniez, F. (eds.), *De la mesure dans les termes*. Lyon: Presses de l'Université de Lyon, 322-348.
- L'Homme, M.-C. and Polguère, A. 2008.** Mettre en bons termes les dictionnaires spécialisés et les dictionnaires de langue générale. In Maniez, F. and Dury, P. (eds.), *Lexicographie et terminologie : histoire de mots. Hommage à Henri Béjoint*. Lyon: Presses de l'Université de Lyon, 191-206.
- Mel'čuk, I. (ed.). 1999.** *Dictionnaire explicatif et combinatoire du français contemporain : Recherches lexico-sémantiques IV*. Montréal: Presses de l'Université de Montréal.
- Moon, R. 2008.** Technicalities of definition. In Maniez, F. and Dury, P. (eds.), *Lexicographie et terminologie : histoire de mots. Hommage à Henri Béjoint*. Lyon: Presses de l'Université de Lyon, 83-98.
- Random House Kernerman Webster's College Dictionary.** 2011. <http://kdictionariesonline.com/DictionaryPage.aspx?ApplicationCode=18/>.
- Random House Webster's College Dictionary.** 2005. New York: Random House.

Practical aspects of the description of terms: contexts, actantial structure and lexical relationships

Geneviève Camirand

As a translation student, my contribution to Marie-Claude Demers's directed study on the enrichment of a general dictionary's wordlist with the relevant contents of a specialized dictionary gave me the opportunity to investigate hidden aspects of some of the resources I will likely use extensively in a professional setting. My role has been to participate, as a research assistant, in developing terminological dictionary entries related to the computer field and contained in a terminological dictionary, the English version of the *DiCoInfo* developed at the Observatoire de linguistique Sens-Texte (OLST), that had been selected with the aim of supplying a general dictionary, the *Random House Kernerman Webster's College Dictionary* (RHKWCD), with new entries and meanings. And indeed, specialized and general resources being some of the main tools for translators, I took a particular interest in the various aspects

of the project. The specific challenge of this project was to write entries that could be added to RHKWCD while respecting the guidelines usually applied in *DiCoInfo*. My role was to add data categories compatible with *DiCoInfo* (contexts, actantial structure, lexical relationships). Once added, these data categories could be used to write a definition and select examples that could be incorporated into RHKWCD.

The criteria for the selection of the terms to be included in the project were basically the following: among the terms whose meaning relative to the computer field was not already described in RHKWCD, only those that were not too specialized to be part of the general language were accepted. It is worth mentioning here that, since *DiCoInfo* is in constant evolution, as is the computer field, the list of terms established the first time is open to new additions.

Figure 1 is a screenshot of part of the list

with which I worked; as can be seen, each lexical unit represents a unique meaning. *Alias*, for example, can refer either to: 1) a kind of pseudonym, 2) a shortcut for a command, or 3) to create a shortcut for a command. Initially, *analyst*₁ was not part of the list; it was added since it was considered that one of its meanings definitely belongs to the computer field, and serves as an actant for other terms (e.g. *analysis*₁), and it is not too specialized to be listed in a general dictionary. *DDR*₁ (double data rate), on the contrary, was removed from the list because it was decided it was not common enough in everyday language.

The contexts, the actantial structure, and the lexical links are the three most important data categories of the *DiCoInfo*'s structure on which I was brought to work. Since the *DiCoInfo* is based on a lexico-semantic perspective, which puts forward a semasiological method, contexts extracted from corpora are the basis of the description. In effect, all other data categories are developed according to the data found there, which means that the quality of the descriptions depends mostly on the quality of the chosen contexts. The search and selection of relevant contexts is thus a core step, which must be given much attention.

Since the computer corpus used for searching lexical items does not contain all the terms that were identified for description,

nor, in some cases, a sufficient number of occurrences for specific terms, it had to be "enriched" with new texts, all of them found on the Internet. I had to develop some skills allowing me to choose, among abundant sources, the ones that would be useful, and to be careful to select recent texts (more likely to present up-to-date information) pertaining to a variety of specialization levels. Between 15 and 20 contexts had to be chosen for each meaning; they were then organized according to the quantity and nature of the information they presented. My main goal was to allow dictionary users to access additional information, so various elements were considered: the presence of actants, of synonyms and antonyms, aspects of definition, etc.

Below, I present contexts that were found for *microcontroller*₁:

A microcontroller is a complete system, consisting of the CPU (computing unit/microprocessor), the programming memory (FLASH or EPROM), working memory (RAM) and in/output on a chip. (Source: MEMORY_CHIPS)

Also inside the mouse are a switch for each button, and a microcontroller which interpret the signals from the sensors and the switches, using its firmware program to translate them into packets of data which are sent to the PC. (Source: INPUTDEVICE)



Geneviève Camirand is currently undertaking a bachelor's degree in Translation at Université de Montréal, where she participated in the enrichment of the English version of *DiCoInfo*, developed at the Observatoire de linguistique Sens-Texte (OLST). She has a background in Modern Languages, which she studied at Universidad Autónoma de Querétaro, in Mexico.
gencamirand@gmail.com

olst.ling.umontreal.ca/dicoinfo/rhwcd/kdproject.php#part3					
<u>administrator 2</u>	n	2			
<u>agenda 1</u>	n	2			<ul style="list-style-type: none"> • calendar program • electronic agenda • electronic calendar • agenda software • scheduling program • scheduling software • digital agenda • agenda application
<u>algorithmic 1</u>	adj	2			
<u>alias 1</u>	n	2			<ul style="list-style-type: none"> • alias name • user alias • pseudonym
<u>alias 2</u>	n	2			
<u>alias 3</u>	vt	2			
<u>analyst 1</u>	n	2			<ul style="list-style-type: none"> • systems analyst
<u>analyzer 1</u>	n	2		<ul style="list-style-type: none"> • analyser 	
<u>analyzer 2</u>	n	2		<ul style="list-style-type: none"> • analyser 	<ul style="list-style-type: none"> • network analyzer • network protocol analyzer • packet analyzer • protocol analyzer

Figure 1: A screenshot of a portion of the list of terms.

The 8051, being a communications-oriented microcontroller, gives the user the ability to access a number of bit variables. (Source: TYPMEM)

Interesting information can be found in the above contexts: the components of a microcontroller, an example of hardware it can be installed in (*mouse*), a related meaning (*chip*), and so on. Attention must be given to the diversity of the sources and the complementary nature of the information contained in the contexts.

During this first step of the descriptive work, I had to deal with two main difficulties. The first was the fact that some terms that unquestionably pertain to the computer field seldom appear in specialized texts. For example, *computerization*, which refers to a rather abstract reality, appears more often in governmental or journalistic texts referring to the computerization of an organization than in an academic article or a user's guide... The second had to do with an opposite problem: the profusion of texts on the Internet, which complicated the identification of serious and relevant sources.

As regards the actantial structure, it allows one to identify which participants of a given term are necessary in order to understand its meaning, and how they interact with it, i.e. which actantial roles they fill. The actantial structure has multiple functions: to identify new terms among the actants, to help construct a definition of the term, to draw the line between different meanings of a lexical unit, to contribute to the explanation of how given terminological units behave within language, and many others. I established actantial structures by analyzing contexts and observing already existing structures in the dictionary. I will illustrate the process with the actantial structure of the verb *bounce*:

bounce: { email1 } ~ from { address3 }
to { sender }

It was decided, for this term, that three actants are necessary in order to understand its meaning: the patient (what bounces), the source (where it bounces from), and the destination (where it bounces to). Each actantial role is replaced on the online *DiCoInfo* by a typical term, i.e. the lexical unit that is most likely to play that role in context, or the generic that better represents all the possible realisations of those units. Choosing the most adequate typical term is often a difficult task.

Finally, the analysis of the contexts allowed me to become more familiar with the lexical relationships section and enrich it. This part of the entry contributes to the

establishment of a network between terms, thus determining the position of the term within a semantic network. It provides paradigmatic relations (hypernyms and hyponyms, antonyms, derivatives), as well as syntagmatic ones (collocates that participate in the description of a term's behaviour within language). Finding new lexical relationships also resulted in introducing new terms to the *DiCoInfo*'s wordlist, and thus additional candidates for the list provided to RHKWCD, since according to the lexico-semantic perspective behind the *DiCoInfo*, most lexical units surrounding a term are also likely to be terms, as are derivatives, synonyms and antonyms. For example, *case insensitivity*₁ was added because *case sensitivity*₁, its antonym, was already part of the wordlist; also, the verb *crack*₁ led to the inclusion of the noun *crack*₂, which designates the same notion, and the noun *cracker*₂, whose function is to crack something.

In brief, my participation in the description of terms allowed me to become more familiar with many steps of terminological work, from supplying a corpus to the establishment of a semantic network. And since, from a translation point of view, I consider general and specialized tools as complementary, I believe that the collaboration between the *DiCoInfo* and RHKWCD is a rich source of investigation themes aiming to demonstrate the inexhaustible bonds between lexicography and terminology.

References

- Demers, M.-C. 2011.** *Travail de recherche en lexicographie bilingue : enrichissement de la nomenclature d'un dictionnaire général à partir de celle d'un dictionnaire spécialisé.* <http://olst.ling.umontreal.ca/pdf/DemersMarieClaude2011.pdf>.
- DiCoInfo : Le dictionnaire fondamental d'informatique et de l'Internet.** Observatoire de linguistique Sens-Texte. 12 May 2012. <http://olst.ling.umontreal.ca/cgi-bin/DiCoInfo/search.cgi/>.
- L'Homme, M.-C. 2004.** *La terminologie : principes et techniques.* Montréal: Les presses de l'Université de Montréal, 32-36.
- Random House Kernerman Webster's College Dictionary.** 12 May 2012. <http://www.kdictionariesonline.com/DictionaryPage.aspx?ApplicationCode=18/>.

Kernerman English-French Learner's Dictionary: Adapting the translation from European French to Canadian/Québec French

Suzanne DesGroseilliers

Introduction

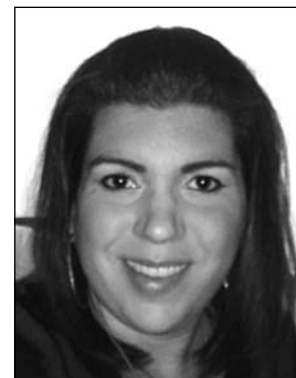
This article describes the process behind the adaptation of an English-French dictionary for a French-speaking audience in Québec. The varieties of French spoken in Canada, and in Québec in particular, have evolved differently than in Europe, where the community of French speakers is much larger. As French is one of the official languages in Canada, and is the official language in Québec, one would assume there are many dictionaries suited to speakers of French in both places. However, the situation in practice is somewhat different. The vast majority of dictionaries used in schools, universities, or businesses are those edited and compiled in France: mainly, *Le Robert*, for French only, and the *Robert-Collins* or *Harraps* for bilingual use. Some exceptions come to mind: *Le Multidictionnaire de la langue française*, *le Dictionnaire Franqus*, *Le dictionnaire québécois d'aujourd'hui*, *Termium*, etc. But these types of dictionaries are mainly intended to explain the specificities of the French language spoken in Québec and few are bilingual. Others are databases offering only equivalents, without examples of usage, and are thus not useful enough for learners. The translated content of the bilingual dictionaries produced in Europe is certainly understood by French speakers in Québec, who are regularly exposed to French culture outside of Québec. Although the French language in Québec is different in its verbal form, in its written form it tends to follow the rules originating from France. That being said, when compiling a dictionary there is no justification not to take into consideration the differences between two communities that speak the same language when compiling a dictionary. Every language has its own peculiarities, making each language distinctive, unique, and special to those who speak it, and is the reflection of the community in which it takes place. It is then clear that some work had to be done: a proper French adaptation of a bilingual dictionary for Québec should be undertaken in order to offer to locals a dictionary they can rely on and in which they can recognize their own variety of French. Mainly, the task was to add French

equivalents (from Québec) to an English/French dictionary that had already been translated by lexicographers from France. The resulting dictionary would then be the reflection of the French spoken by native speakers in Québec.

I adapted the dictionary in Montréal, Canada, in the summer of 2011 as part of my internship as a third year translation student at the University of Montréal. The purpose of this article is to explain in detail the process of this adaptation and to give more information about the general and the specialized references that I used to do so. After analyzing the main modifications noted in the process, I was able to identify categories of the main differences found between the France version and the Québec version. This not only shows that differences do exist in French between France and Québec, but also the importance of adapting a dictionary to a specific target audience.

Working method

As mentioned above, the main task was to add Québec equivalents to an English/French dictionary in which the initial English nomenclature had already been translated by lexicographers in France.



Suzanne DesGroseilliers

obtained a bachelor degree in Psychology in 2006 and a bachelor degree in Translation in 2011. During her translation studies, she worked as a student researcher in terminology at the Observatoire de linguistique Sens-Texte (OLST) at Université de Montréal. She currently works as a translator at SDL International in Montréal.
sdesgroseilliers@sdl.com

FRANCE VERSION

ice-cream [ais'kri:m] *n* 1 a sweet, frozen food made from cream and sugar

{fr} - crème glacée *f*

◇ chocolate/vanilla/strawberry ice cream

{fr} - de la crème glacée à la vanille/fraise/ au chocolat

2 *Brit* a serving of ice cream

{fr} - crème glacée *f*

◇ Can I have an ice cream?

{fr} - Je peux avoir une crème glacée ?

QUÉBEC VERSION

ice-cream [ais'kri:m] *n* 1 a sweet, frozen food made from cream and sugar

{fr} - glace *f*

◇ chocolate/vanilla/strawberry ice cream

{fr} - de la glace à la vanille/fraise/ au chocolat

2 *Brit* a serving of ice cream

{fr} - glace *f*

◇ Can I have an ice cream?

{fr} - Je peux avoir une glace ?

Figure 1: Example of the translation of *ice-cream* in the France and Québec versions.



Figure 2: Example of the term *crème glacée* in *Antidote*.

The example of the lexical unit ‘ice cream’ is shown in Figure 1. After analyzing the different parts of the entry (definition, equivalents, examples, etc), it was clear that the France equivalent was not properly suited for a Québec audience. Indeed, the lexical unit « glace » in Québec means ‘ice’ and since this term is strongly related to the Québec reality (winter season), it is unlikely that native French speakers in Québec would relate the term « glace » to another term than ‘ice’. The decision to change « glace » for « crème glacée » was then taken. Moreover, the term « crème glacée » was validated both by the equivalent found in the reference documents and by other native speakers.

Reference documents

The general references that were used for the adaptation were the *Harrap’s Shorter*, mainly to validate the equivalents, *Le Grand Robert de la langue française* (online version), to add the International Phonetic Alphabet, and *Le Petit Robert*, to validate the French definitions already in place. *Merriam-Webster’s Collegiate Dictionary* was used to validate the English definitions. As for European users, these types of reference documents are the ones that are usually used by French speakers in Québec.

I began by compiling a list of specialized references in order to properly adapt and validate the modifications made to the dictionary. This list had to be complete, the references well known in the terminology and lexicography fields, and most importantly, they had to be properly adapted to the local audience. The software *Antidote* was used, since it was created in Québec and is well known in the academic community. This application brings together a French grammar checker, 12 different dictionaries (synonyms, antonyms, co-occurrences,

idioms, etc), 11 linguistic guides (grammar, conjugation, definitions, anagrams, family, etc.) and many other linguistic revision tools. It considers the Québec reality and the specific usages of French language in this culture. It also compares the different lexical units used in Québec against those in France. A mention of *Québec* or *France* is shown in the majority of entries, telling the user if the word is adapted to either community. This software is not only useful to French speakers in Québec but also to learners, since the differences between France and Québec are well exposed. Figure 2 is a screenshot of *Antidote* referring to the same example presented in Figure 1. It validates the final decision to replace « glace » by « crème glacée ».

The *Dictionnaire Franqus*¹ was used to evaluate the French equivalents. This dictionary is the first French general dictionary to be completely compiled outside of France. It is also the first dictionary to be elaborated strictly based on a Québec linguistic corpus. This dictionary was useful but since it is still being developed, the results were not always conclusive. The wordlist is still limited and more specific or technical terms are not included. The *DicoInfo* (L’Homme 2011) is a specialized dictionary listing and explaining the myriad connections of terms from various domains of computer science as well as Internet usage. Since this resource is compiled in Québec, and is available online, it was used to validate the equivalents of the field of computing. Also available online, *Le Grand dictionnaire terminologique* is a terminological data bank gathering terms and their English equivalents from specialized fields. Mainly used to validate French equivalents, *Termium* is another online terminological and linguistic data bank. It was elaborated by the Translation Bureau, which is the

federal government's centre of expertise in translation and linguistic services and one of the world's leading translation organizations, and gathers almost 4 million English and French terms. I used both to confirm French equivalents. Finally, I chose two printed dictionaries to validate either the equivalents already in place or the new ones. The first, *Dictionnaire québécois d'aujourd'hui*, is based on *Le Robert d'aujourd'hui* and tends to reflect the usage of French in Québec. Secondly, *Le Multidictionnaire de la langue française* was often consulted. This dictionary was compiled by Marie-Éva de Villers and is based on the enquiries to the *Office québécois de la langue française*. It takes into consideration the present use of language in Québec and is a complete guide adapted to the particular case of French in Québec. Finally, besides the consultation of these dictionaries and data banks, I applied my own knowledge of Québec French, and often referred my questions also to other local native speakers to validate the use of certain terms.

Results

During the process of adapting the dictionary to a Québec public, which implies the revision of the entries and the French proposed equivalents, I noted all the modifications between the two French wordlists that I had made. These notes were then divided into two categories: differences and errors. The compilation of the main differences strengthened the importance of having a bilingual dictionary that is well-adapted to a specific public, providing a better idea of the differences between two groups of native speakers from different countries.

Main differences

After reviewing all the main differences, I divided them into seven categories: Spelling, Anglicism, Usage, Unknown, Intercultural, Idiomaticity and No equivalent. Each of these categories is explained below and an example is provided.

First, some differences were noted as **Spelling** differences. This implies that the spelling of a certain term in Québec is different, or that the France spelling is not usual in Québec. Figure 3 shows an example.

light² [lait] *adj*
6 (lite) containing less fat or sugar than usual {fr} - allégé/-ée
 ◇ *light yogurt* {fr} - *yoghourt allégé*

Figure 3: Example of the entry *light*.

In this entry, the spelling of the French equivalent *yoghourt* is not usual in Québec. In another entry ('fat free'), the same French equivalent is spelled *yaourt*. Although the two forms of spelling are known in Québec, they are certainly not usual. The spelling is the first form presented in the *Multidictionnaire de la langue française* and this form of spelling was also confirmed in *Antidote*. Furthermore, the pronunciation of the words is different in France and in Québec. Most of the time, native speakers in Québec do not pronounce the final "t" while this is not the case in the French-speaking communities in Europe. I thus modified this equivalent to adapt the spelling of the word *yogourt* to the French speakers in Québec and consequently, I changed the IPA to reflect the pronunciation.

The second category concerns differences between words that I define as **Anglicism**. The terms in this category are those considered to be Anglicism in French-speaking communities, particularly in Québec. There seems to be a general tendency in France to use Anglicism in the common language. French speakers in Québec are widely exposed to English (from Canada and the United States) and, like French speakers in France, tend to use English terms in all kinds of situations. Furthermore, as a result of being exposed to English grammar, French speakers in Québec tend to make grammatical errors in French. For example, the use of the verb « identifier », which is a proper French word, is most of the time incorrect. The correct use of the term is influenced by its English equivalent (*identify*). Dictionaries and other reference books then tend to promote a proper usage of French terms and to banish from the vernacular all forms of Anglicism. I changed all Anglicisms found in the France wordlist to proper French terms that are actually used in Québec. For example, *email* was changed to *courriel* (which is a neologism proposed by the Office québécois de la langue française) and *weekend* to the French equivalent *fin de semaine*.

Some differences can be referred to the question of **Usage**. Usually, the term is well understood by the French speaker in Québec, but it is not usual. I validated such terms with the 'frequency rate' (called *indice de fréquence*) proposed in *Antidote*. For each word, the rate of usage frequency is indicated on a scale of 100. For example, the entry 'oatmeal' initially had the French equivalent *porridge*. However, based on my native knowledge of French, and as a professional translator, I considered this was most certainly not a word largely used in Québec. In *Antidote*, the frequency rate

for *porridge* is 24, whereas that for *gruau* is 30, so I used the latter. Other examples of terms that I decided to change based on the *Antidote* frequency rate, on results found in the reference tools (*Multidictionnaire*, *OQLF*, *Colpron*, etc), and on my own knowledge of French, include: canular (40) / blague (46), troquet (30) / brasserie (44), mesure (35) / cabane (46), myrtille (34) / bleuets (39) and trouillard (27) / peureux (37).²

For certain entries, the equivalent proposed was not known in Québec. I gathered these few examples in the category **Unknown**. In the entry ‘bribe,’ the equivalent was « bakchich ». Since I did not know this word I did some research and found that it is not used in Québec. Furthermore, I validated this with other local native French speakers and no one knew what this word meant. I therefore decided to change the proposed equivalent to a more commonly-used word in Québec, which is « pourboire ».

The next category is **Intercultural**, which is closely related to that of **Usage**.

I gathered here the terms that represent a different reality in both societies or that often have a different meaning. Examples of such terms are those used in the domain of education and in the proper terminology related to the different meals in a day. In the field of education, the terminology usually used in France and in Québec varies largely. For example, in France the terms « jardin d’enfant » and « lycée » are used as opposed to « maternelle » and « secondaire » in Québec. Those differences could be categorized both in usage or intercultural. However, if we take the example of « baccalauréat » we see that the term is used both in France and in Québec, but the meaning is different in each community. In France it refers to the diploma required to pursue superior studies (received at the end of the secondary studies), whereas in Québec it refers to the diploma received at the end of the first cycle in university (bachelor’s degree). Those usages have to be taken into consideration in the adapted nomenclature of the bilingual dictionary. Another example refers to the different meals. Table 1 illustrates the main differences.

As seen here, the same term, « déjeuner », refers to two different meals: in France it is the second meal of the day (known as ‘lunch’ in English), while in Québec it is the first meal of the day (known as ‘breakfast’

Table 1. Example of intercultural differences between France and Québec.

English	France	Québec
breakfast	petit-déjeuner	déjeuner
lunch	déjeuner	dîner
supper/dinner	dîner	souper

lai	adj.	pot au lait
laid	adj.	fromage au lait
laid	n.	pinte de lait
laie	v.	biberon de lait
laie	n. f.	bouteille de lait
laie	n. f.	frère de lait
lais	n. m.	agneau de lait
lé	n. m.	consommation de lait
lei	n. m.	boîte de lait
let	n. m.	
let	adi.	

Figure 4: Examples provided for the term *lait* in *Antidote*.

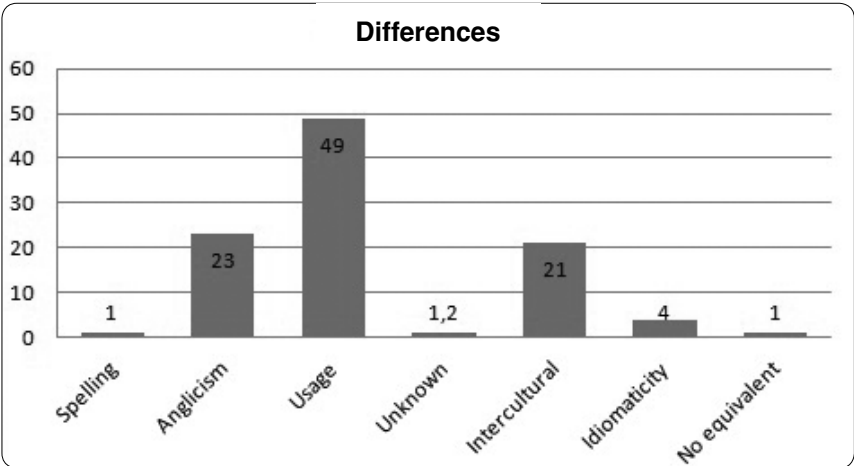


Figure 5: Graphic representation of the main differences in the France/Québec translations.

in English). It is important to note such differences not only in order to adapt the wordlist to the target public, but also to alert learners and prepare them to use the proper terminology in a certain domain.

Certain terms are considered to be more idiomatic in a particular society. This is what we call co-occurrence, and the category I created for it is **Idiomaticity**. To verify the level of co-occurrence of a certain equivalent I used once more the tools provided in *Antidote*. One of the examples provided in the entry ‘expiration’ of the France equivalent is « brique de lait ». Since this is not used in Québec, I verified the co-occurrence for « lait » in *Antidote*. It appears that « pinte de lait » or « carton de lait » are widely used, while I did not

find any concurrence for « brique de lait » (Figure 4).

Therefore, I changed the example to « pinte de lait », since it is more idiomatic and more usual in Québec, as validated in *Antidote*.

Finally, for certain terms, I did not find any equivalent since the realities in Québec and France are not the same. I aggregated these words in the category **No equivalent**. The example provided for the term ‘junior’ is a ‘junior high school student’. In Québec, there is no such differentiation of students in high school. Therefore, no equivalent can be provided. I left the French equivalent proposed by the France lexicographers, which is « en classe de première ». In the Québec version of the dictionary there could be a “lexicographer’s note” to explain the difference in cultural realities.

Figure 5 represents the main differences found in the adaptation process of the dictionary. The main difference concerns the category of Usage, followed by Anglicism and Intercultural differences. This is helpful data for a lexicographer, since it emphasizes the type of equivalent that should be adapted and should be more carefully analyzed in the process of translating and adapting a dictionary to different speakers of the same language.

Main errors

Besides the differences noted in the adaptation process, I also noted errors and gathered them in four categories: *Inflection*, *Spelling*, *Equivalent*, and *Definition*.

In the category **Inflection**, some entries did not provide the feminine inflection. I added it, since feminizing terms is important in Québec, as stated by Larivière (2000): “how can we be equal if invisible.” For example, the term ‘coach’ only provides the French equivalent « entraîneur ». I added « entraîneuse », which is commonly used in Québec.

I also noted certain **Spelling** errors, of grammar, spelling, typography, or obsolete use of terms. Errors of grammar were common in the examples. In the example of ‘lift¹’, the following sentence was provided: « Sa père l’a fait passer par dessus la barrière. » Since « père » is a masculine word I changed « sa » for « son ». For the entry ‘flight attendant’ the French equivalent is « hôtesse de l’air ». After verification (Figure 6), I found that this was an obsolete use of that word (although still correct), and changed it to « agent de bord », both because it tends to become more common in Québec and because it is the official term.

I also noted **Equivalent errors**. For example, in ‘decade’ the French equivalent

was « décade », while the proper one is « décennie ». Finally, from my personal view, I thought that certain **Definitions** were not well adapted to the Québec reality. For example, the following definition of ‘grape’: “a small green or purple fruit used to make wine”. Although it is well known in Québec that grapes are used to make wine, such activity is less common than in France. Therefore, this definition does not fit the Québec reality, where a grape is more often a “fruit that is eaten” and a native speaker would not have immediately the mental representation of a “fruit used to make wine”. However, I did not change the definition since that was not within my task, but noted it for further adjustments or updates. Figure 7 shows graphically the main errors noted during the adaptation process. It is clear that these consisted mainly of spelling mistakes, which are common in the making of a dictionary. Surprisingly, 35% of errors were in the nature of equivalents. The processes of verifying, updating, or even

HOTEL EMPLOYES ASSIMILES NON SPECIALISES ALIMENTS BOISSONS	1. Subject Field(s)	
	- Occupation Names (General) - Air Transport Personnel and Services	
HOTEL		Domaine(s) - Désignations des emplois (Généralités) - Personnel et services (Transport aérien)
HOTEL ACCUEIL		
HOTEL AIR	cabin crew member	membre de l'équipage de cabine
HOTEL CHEF	Source CORRECT, OFFICIALLY APPROVED	Source CORRECT, MASC, OFFICIALLY APPROVED
HOTEL CONGRES	flight attendant	agent de bord
HOTEL GITE FAMILIAL	Source CORRECT	Source CORRECT, MASC, OFFICIALLY APPROVED
HOTEL INFOVARIETES	cabin attendant	agente de bord
HOTEL JEUX TELEVISES	Source CORRECT, OBSOLETE	Source CORRECT, MASC, OFFICIALLY APPROVED
HOTEL PRINCIPALE	stewardess	agente de bord
HOTEL RESTAURANT	Source CORRECT, SEE OBS, OBSOLETE	Source CORRECT, FEM
HOTEL RESTAURANT BAR SALON	air hostess	hôtesse
HOTEL SALLE MANGER	Source CORRECT, SEE OBS, OBSOLETE	Source CORRECT, SEE OBS, FEM, OBSOLETE
HOTEL SOL	flight hostess	hôtesse de l'air
HOTEL VILLAGE VACANCES	Source CORRECT, SEE OBS, OBSOLETE	Source CORRECT, SEE OBS, FEM, OBSOLETE
HOTEL ACCUEIL		
HOTEL EMPLOYES ASSIMILES NON SPECIALISES ALIMENTS BOISSONS		

Figure 6: Terminological entry of *flight attendant* in *Termium*.

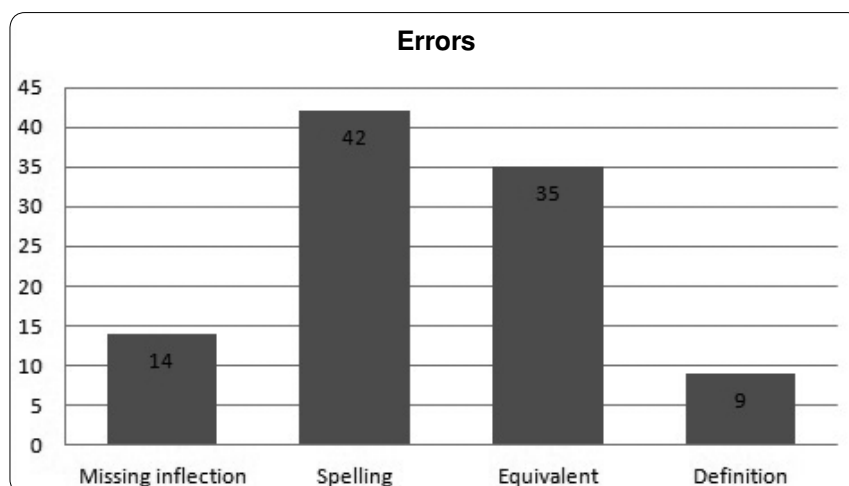


Figure 7: Graphic representation of the main errors.

Conclusion

Libro ingabiato [*Book in a cage*], Elisabetta Gut, 1981
Wood, Wire, and French-Italian dictionary. 4 ¾ x 4 ½ x 5 ¼ in.
National Museum of Women in the Arts, Washington, DC



¹ Authorization to use the online version of *Franqus* dictionary was graciously granted by H  l  ne Cajolet-Lagani  re.

References

- Antidote. Correcteur-Dictionnaires-Guides.** 2010. Montréal: Druide Informatique.
- DicoInfo. Le Dictionnaire fondamental de l'informatique et de l'Internet.** 2011. <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi/>.
- Dictionnaire Québécois d'aujourd'hui.** 1992. Boulanger, J-C. Montréal: Dicorobert.
- Franqus. Dictionnaire de la langue française. Le français vu du Québec. Édition BÊTA.** 2011. Université de Sherbrooke. Faculté des lettres et sciences humaines. Sherbrooke. <http://132.210.238.8/dictio/accueil.jsp;jsessionid=1iqxrdqs4955q/>.
- Grand dictionnaire terminologique.** 2011. Québec: Office québécois de la langue française. <http://www.oqlf.gouv.qc.ca/ressources/gdt.html/>.
- Harrap's Shorter. English-French/French-English Dictionary. 7th edition.** 2004. Edinburgh: Chambers Harrap Publishers.
- Larivière, L.** 2001. *Pourquoi en finir avec la féminisation linguistique ou À la recherche des mots perdus.* <http://id.erudit.org/iderudit/000577ar/>.
- Le Grand Robert de la langue française. Dictionnaire alphabétique et analogique de la langue française.** 2011. <http://www.bib.umontreal.ca/>.
- Le Petit Robert. Dictionnaire alphabétique et analogique de la langue française.** 1993. Paris: Dictionnaires Le Robert.
- Le Robert & Collins Senior. Dictionnaire français-anglais. Fourth Edition.** 1995. Great Britain and Paris: HarperCollins Publishers and Dictionnaires Le Robert.
- Merriam-Webster's Collegiate Dictionary. Tenth Edition.** 2002. Springfield, MA: Merriam-Webster.
- Multi dictionnaire des difficultés de la langue française.** 2008. De Villers, M-E. (ed.). Montréal: Éditions Québec/Amérique.
- Termium. The Government of Canada's terminology and linguistic data bank.** 2011. Ottawa: Translation bureau. Public Works and Government Services Canada. <http://www.termiumplus.gc.ca/>.

A study of the third-generation Chinese-English dictionaries

Xia Lixin

Abstract

This paper introduces a survey of the general Chinese-English dictionaries published in Mainland China over the past 30 years. The author argues that whereas these dictionaries have distinctive Chinese characteristics, they are not made for any specific user category nor are they made for specific uses. Instead they are intended for all types of users and linguistic activities. Besides this, they follow traditional dictionary-making practice, introducing few innovations. On the basis of the survey and its analysis, suggestions are put forward to improve the compilation of Chinese-English dictionaries.

Keywords Chinese-English dictionary, dictionary-making, lexicography

1. Introduction

The history of Chinese-English (C-E) dictionary-making in China in the last 200 years can be divided into three phases (Zeng 2003). The first phase was from 1815 to 1911. The first title was *A Dictionary of the Chinese Language: in Three Parts*, compiled by Robert Morrison from 1815~1823. It was followed by Samuel Williams's *A Syllabic Dictionary of the Chinese Language* (1874), Herbert Giles's *A Chinese-English Dictionary* (1892), and Frederic Baller's *An Analytic Chinese-English Dictionary* (1900). Most of these lexicographers were Western priests aiming to facilitate communication in their religious work in China.

Not being professional lexicographers, these compilers selected lemmas without applying systematic rules and provided little information about them except equivalents. Therefore, these titles, in modern terms, were at most bilingual glossaries, but not dictionaries. Besides, due to their limited knowledge of the Chinese language, the authors made a great number of mistakes in their dictionaries, especially concerning codifying pronunciation and explaining the lemmas.

The second phase started in 1912 and ended in 1977. During this period, a dozen C-E dictionaries were compiled mainly by Chinese lexicographers rather than by Western missionaries. The first one was probably *A Chinese-English Dictionary* by Zhang Zaixin and Ni Shengyuan (1912), which included about 3,800 single characters. Six years later, Li Yuwen (1918) published

A New Chinese-English Dictionary, which contained about 10,000 single characters and over 50,000 phrases. Other influential works in this phase were Liang Shih-chiu's *The New Practical Chinese-English Dictionary* (1971) and Lin Yutang's *The Chinese-English Dictionary of Modern Usage* (1972). Compared with those made in the first phase, these dictionaries pioneered such major innovations as distinguishing free morphemes from bound morphemes and arranging headwords according to the radicals. They also housed substantial improvements in their lemma selection, definitions, and ways of representation.

The third phase began in 1978 and has been going on to the present day. Due to the open policy adopted by the Chinese government, the demand for C-E dictionaries increased dramatically. As a result, a great number of new titles appeared in this period. Our survey is intended to cover the main problems of the C-E dictionaries published in this period in Mainland China.

2. The status quo of the third-generation Chinese-English dictionaries

In general, the C-E dictionaries published in China since 1978 have clear compilation purposes and aims with distinctive Chinese characteristics. Most of them are arranged in phonetic order, are rich in illustrative examples, and have double macro-structures — that is to say, an entry is usually comprised of a single character as the main entry and a list of multi-character phrases as its subentries. Based on the *Contemporary Chinese Dictionary* (2005), they select entry words according to their types and sizes, with particular attention to new words and new senses of existing words. Some dictionaries explain and illustrate the usage of the equivalents by providing synonym discrimination and collocations.

A dozen titles were selected from this category, as listed in Table 1. These dictionaries are widely used in China, and are representative in one way or another. For example, *A Chinese-English Dictionary*, edited by Wu Jingrong (1978), is the first one of the third-generation Chinese-English dictionaries. It won popularity soon after it was published, and became one of the most influential C-E dictionaries in China. Yao Xiaoping, editor of the third edition, claims



Xia Lixin is professor and MA candidate supervisor at the Centre for Lexicographical Studies, Guangdong University of Foreign Studies. His main research interests are lexicography, translation, and semantics. He is general secretary of the Chinalex Bilingual Committee, author of more than 20 papers on lexicography and translation, and collaborator on several major research projects. cdhuiyi@yahoo.cn

that “[t]his first step is extremely important in the history of the Chinese-English dictionary compilation in China. Every other Chinese-English dictionary coming after it, no matter whether it is a new edition or a revised one, benefits more or less from this first dictionary (2010: 3).” So this dictionary and its following editions are included in this research.

The Chinese-English Dictionary (1998) and its two revised editions, *A Comprehensive Chinese-English Dictionary* (2004) and *The New Chinese-English Dictionary* (2003), were all written by Wu Guanghua. His works are famous for their large size of entry lists and a wealth of technical lemmas. *A New Century Chinese-English Dictionary* (2003) was compiled by Hui Yu. He claims that his dictionary is practical and reader-centered. In particular, it outnumbers any other C-E dictionary in providing collocations. *A New Age Chinese-English Dictionary* (2000) is another dictionary made by Wu Jingrong. It is highly acclaimed by critics and readers for its novelty, accuracy, and practicality.

3. Problems of the third-generation Chinese-English dictionaries

Although the past 30 years have been a period of prosperity for the publication of C-E dictionaries, there still are some problems.

3.1 Mixed user categories

In Table 1, the type of each dictionary is given according to the statement in the preface of the dictionary. According to Svensén, a monodirectional dictionary is a bilingual dictionary “intended only for native speakers of one of the language” (2009: 28). In this case, the C-E dictionary designed only for Chinese users is monodirectional, such as *A Chinese-English Dictionary* (1978), *A Modern Chinese-English Dictionary* (2001), and so on. In this spirit, a bidirectional dictionary is a bilingual dictionary “intended for native speakers of both languages” (Svensén 2009: 28). If a C-E dictionary is made for both Chinese users and foreign users, it is a bidirectional bilingual dictionary. Thus, *The Chinese-English Dictionary* (1998),

Table 1. A list of the dictionaries under investigation

Editor and Year	Dictionary Title	Published by	Type
Wu Jingrong (1978)	A Chinese-English Dictionary	The Commercial Press	monodirectional
Wei Dongya (1995)	A Chinese-English Dictionary (Revised Edition)	Foreign Language Teaching and Research Press	bidirectional
Yao Xiaoping (2010)	A Chinese-English Dictionary (Third Edition)	Foreign Language Teaching and Research Press	monodirectional
Wu Guanghua (1998)	The Chinese-English Dictionary	Shanghai Jiao Tong University Press	bidirectional
Wu Guanghua (1999)	The Chinese-English Dictionary (Second Edition)	Shanghai Jiao Tong University Press	bidirectional
Wu Guanghua (2010)	The Chinese-English Dictionary (Third Edition)	Shanghai Translation Publishing House	bidirectional
Wu Guanghua (2001)	A Chinese-English Dictionary (New Century Edition)	Shanghai Jiao Tong University Press	bidirectional
Wu Guanghua (2003)	The New Chinese-English dictionary	Shanghai Jiao Tong University Press	bidirectional
Wu Guanghua (2004)	A Comprehensive Chinese-English Dictionary	Dalian University of Technology Press	bidirectional
Wu Jingrong (2000)	A New Age Chinese-English Dictionary	The Commercial Press	bidirectional
Dictionary Division of FLTRP (2001)	A Modern Chinese-English Dictionary	Foreign Language Teaching and Research Press	monodirectional
Hui Yu (2003)	A New Century Chinese-English Dictionary	Foreign Language Teaching and Research Press	bidirectional
Wu Wenzhi (2001)	A Practical Chinese-English Dictionary for Translation	Lijiang Publishing House	monodirectional
Gao Ling (2010)	A Practical Chinese-English Dictionary for Students	Foreign Languages Press	monodirectional

A New Age Chinese-English Dictionary (2000), *A New Century Chinese-English Dictionary* (2003), etc. belong to this type of dictionary.

As appears in Table 1, most of the dictionaries claim to be bidirectional, meant both for speakers of Chinese who are learning English and for foreign learners of Chinese. In order to meet the needs of these two different user categories, the dictionaries have to include more information in the entries, which makes their structure complicated. For example, *A Comprehensive Chinese-English Dictionary* (2004), *The Chinese-English Dictionary* (1998) and *The New Chinese-English dictionary* (2003) define the headwords of a single Chinese character in both Chinese and English. The definitions in the Chinese part are intended for users whose native language is not Chinese. However, because of the complicated structures of these dictionaries, it would be very hard for foreign users to use them. According to another user research (Xia 2009), non-native learners of Chinese seldom use CFL (Chinese as a Foreign Language) dictionaries, or the monodirectional C-E dictionaries made especially for CFL learners. Thus, the Chinese definitions provided especially for CFL learners are useless, or at least superfluous, since both user categories would not consult them.

Let's take another example. *A Comprehensive Chinese-English Dictionary* (2004) and *The New Chinese-English Dictionary* (2003) set up 6,000 and 4,000 synonym discrimination columns, respectively. The objects to be discriminated are the semantic differences between Chinese headwords. The information is apparently provided for CFL learners. However, since almost no CFL learners would use these dictionaries, the information is unnecessary, and even becomes an obstacle to Chinese users. As we know, the major users of these dictionaries are Chinese native speakers. They need badly to know the semantic nuisance and usage of the English equivalents, but not the differences between the Chinese headwords. Therefore, the information of Chinese synonym discrimination might hamper their consultation, or at least lower the speed of their searches.

Based on the above analysis, we might conclude that the information added for CFL users in these bidirectional C-E dictionaries neither finds favor with CFL users nor wins the praise of Chinese native speakers. On the contrary, the added information may make dictionary searches more difficult. As a result, some of the Chinese users may give up using these dictionaries.

3.2 Mixed uses

In theory, a C-E dictionary may be used for translation or learning, or both. The dictionaries listed in Table 1 all claim in their prefaces that they can be used for both purposes. However, the purposes of using a C-E dictionary are quite different when used for translation than for learning. A translator looks up a C-E dictionary mainly for the equivalents of the headword. Thus, the dictionary should include as many headwords as possible. But an English learner would like to know more about the usage of the equivalents rather than the equivalents only. The learner needs grammatical, collocation, and pragmatic information of the equivalents. Thus, the dictionary should provide full and detailed explanations of the equivalents. To meet the requirements of both these dictionary users, the dictionary will inevitably become bulky and complicated. That's why current C-E dictionaries grow bigger and bigger.

Such oversized dictionaries may cause problems for users. First, they are too heavy to carry. For example, *A Comprehensive Chinese-English Dictionary* (2004) has more than 6,000 pages in three volumes. It's very inconvenient for a student to carry them to and from classroom. Second, they are too complicated to use for common dictionary users. As these dictionaries tend to include as many headwords and as much information as possible, they become very complex in their structures. This adds the difficulty of dictionary searches. In fact, some scholars (Zeng 2005; Zhang 2007) have pointed out that it is infeasible to compile a C-E dictionary for all uses.

3.3 Lack of innovation

Although the C-E dictionaries have experienced rapid developments in their quantity and quality in recent years, the same cannot be said for efforts to introduce innovation in their features. Zeng Taiyuan argues that "Looking at the Chinese-English dictionaries published across the Straits, they have little difference in their contents. The structure of the entry is generally made up of two parts: definitions (equivalents and explanations) and illustrative examples (expressions, phrases, and sentences). The only difference lies in the number of headwords and examples in the dictionaries, the inclusion of new words and new senses and the accuracy of the definition" (2005: 81).

We will look at *A Chinese-English Dictionary* (1978) and its revised edition (1995) first. The first change made is the addition of 800 single character headwords and 18,000 multi-character headwords (including proverbs, idioms, and colloquial

expressions). The added headwords are mainly new words. At the same time, the revised edition drops some obsolete entries. The next change is that more explanations are given to polysemous headwords and equivalents. Another major change is to use English labels instead of the Chinese ones in the first edition.

A New Age Chinese-English Dictionary (2000) advocates novelty, accuracy, and practicality. However, its notion of “novelty” restricts only to the inclusion of new words and new senses. According to its statement in the preface, this is embodied in the following two aspects. One is to include as many new words as possible. The other is to give detailed explanations of the new senses of the old words (Wu 2003: 3).

A New Century Chinese-English Dictionary (2003) is hailed as opening up a new generation of C-E dictionaries by some critics (Su 2004). However, the dictionary hasn't achieved any significant breakthroughs and innovations in the field of lexicography. In modern lexicography, corpora have been made full use of. According to Rundell (2010), now few serious English dictionaries (or UK-published bilingual dictionaries) are compiled without reference to corpora. And all the major publishers of English learners' dictionaries were using corpora as their primary source of linguistic data. However, *A New Century Chinese-English Dictionary* (2003) might be an exception. It is not compiled on any modern corpora.

3.4 Absence of Chinese-English learners' dictionaries

C-E dictionaries enjoy a big market among learners of English in China. According to Yong (2003), up to 73% of English majors at universities own a C-E dictionary. And the rate of owning a C-E dictionary among English learners at universities ranks second, just next to bilingualized dictionaries (Yu 2001). These data show the popularity of this dictionary type among English learners.

However, there are no C-E dictionaries in the market that are especially made for English learners. Yuan (1996) argues that no C-E learner's dictionaries are available in China. Although many dictionaries claim to be “a Chinese-English dictionary for students,” they are only the shortened versions of the general C-E dictionary. The situation has remained unchanged up until now. That is to say, no C-E dictionaries are compiled especially for English learners, despite the great need for them in the market.

4. Conclusions

Based on the above analysis, we may conclude that C-E dictionaries need improving badly. First of all, they should be compiled according to the needs of different user categories. In other words, C-E dictionaries should be designed and made especially for users whose native language is Chinese and for users whose native languages are not Chinese. For the dictionaries aimed at foreign users, they should focus on providing information on the meanings and usage of Chinese headwords. But for the dictionaries aimed at Chinese users, the emphasis should be placed on explaining the meanings and usage of the English equivalents.

Secondly, the C-E dictionary should be made specifically for different linguistic activities. That is to say, a dictionary for translation and a dictionary for English learners should be compiled separately instead of having an integrated one for both purposes. While the former shall be centered on providing adequate and accurate equivalents to help users in their production of English texts, the latter will aim at creating an environment that facilitates the acquisition of the English language.

Thirdly, latest research findings in linguistics and Second Language Acquisition (SLA) should be introduced into the compilation of the C-E learner's dictionary. For example, English learners' dictionaries benefit vastly from research findings in SLA. Modern lexicography is shifting from the lexicographer-centered to user-centered approach, which accords with the cognitive laws and the needs of foreign language learning. Therefore, C-E lexicographers must study the interlanguage characteristics of learners and their English linguistic features in order to compile their dictionaries accordingly. For a C-E learner's dictionary, the Chinese part can only be used as an index for dictionary users, and no further information should be given to it. This would avoid distracting the users. But the English part should be explained in great detail.

Finally, C-E dictionaries must be made with the aid of modern corpora. Ever since the first electronic corpus was used in the COBUILD dictionary in 1987 (Rundell 2010), large electronic corpora have become indispensable for dictionary-making, including monolingual as well as bilingual dictionaries. They can help lexicographers make editorial decisions based on actual language usage rather than on intuition or on second sources. Furthermore, they can provide lexicographers not only with the frequency of word uses and linguistic

features, but also collocation and usage of particular words. Nowadays English learners' corpora are available in China. Through careful analysis of their data, lexicographers can predict possible difficulties that learners might encounter in their process of learning English, highlighting them in their dictionaries accordingly. This could greatly improve the quality of C-E dictionaries.

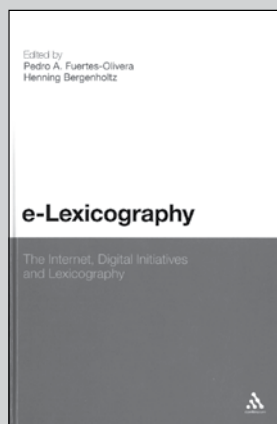
Acknowledgements

The research presented in this paper is sponsored by the innovation project of Guangdong University of Foreign Studies (No. TD1102) and the innovative talents project of the Department of Education of Guangdong Province (WYM10033). I would also like to thank the anonymous reviewers for their valuable comments.

References

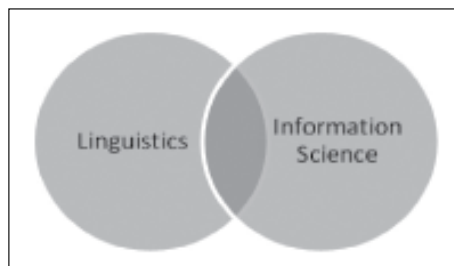
- Baller, F. 1900.** *An Analytical Chinese-English Dictionary*. Shanghai: China inland mission and American Presbyterian mission press.
- Giles, H. 1892.** *A Chinese-English Dictionary*. London: B. Quaritch.
- Li, Y. 1918.** *A New Chinese-English Dictionary*. Shanghai: The Commercial Press.
- Liang, S.-C. 1971.** *The New Practical Chinese-English Dictionary*. Taipei: The Far East Book Company.
- Lin, Y. 1972.** *The Chinese-English Dictionary of Modern Usage*. Hong Kong: Chinese University of Hong Kong.
- Morrison, R. 1822.** *A Dictionary of the Chinese Language: in Three Parts*. Macau: Printed at the Honorable East India company's press, by P.P. Thoms.
- Rundell, M. 2010.** Taking corpus lexicography to the next level: explicit use of corpus data in dictionaries for language learners. In Zhang Y. (ed.), *Proceedings of the First International Symposium on Lexicography and Second Language Acquisition*. Shanghai: Shanghai Foreign Language Education Press, 367-386.
- Su, X. 2004.** New concepts of making the fourth-generation Chinese-English dictionaries: A review of *A New Century Chinese-English Dictionary*. *Foreign Language Teaching and Learning*, 3: 6-8.
- Svensén, B. 2009.** *A Handbook of Lexicography: The Theory and Practice of Dictionary-Making*. Cambridge: Cambridge University Press.
- The Contemporary Chinese Language. 2005.** Dictionary Department, Institute of Linguistics, Chinese Academy of Social Sciences. Beijing: The Commercial Press.
- Williams, S. 1874.** *A Syllabic Dictionary of the Chinese Language*. Shanghai: American Presbyterian Mission Press.
- Wu, J. 2000.** Preface to *A New Age Chinese-English Dictionary*. Beijing: The Commercial Press
- Xia, L. 2009.** An investigation into the publication of CFL (Chinese as a Foreign Language) learners' dictionaries from user perspective. *Publishing Science*, 1: 65-69.
- Yao, X. 2010.** Preface to *A Chinese-English Dictionary (Third Edition)*. Beijing: Foreign Language Teaching and Research Press.
- Yong, H. 2003.** A comparison of users needs for English-Chinese bilingual dictionaries and English monolingual dictionaries. *Lexicographical Studies*, 6: 123-133.
- Yu, W. 2001.** An investigation into the uses and needs for English-Chinese and Chinese-English dictionaries. In Huang, J. (ed.), *Proceedings of the First Asialex Regional Symposium*. Shanghai: Shanghai Lexicographical Publishing House, 171-182.
- Yuan, K. 1996.** On the compilation of Chinese-English dictionaries for students. *Lexicographical Studies*, 2: 98-103.
- Zeng, D. 2003.** An introduction to the history of Chinese-English dictionaries. In Zeng, D. (ed.), *Studies on Bilingual Dictionaries*. Shanghai: Shanghai Lexicographical Publishing House, 3-15.
- Zeng, T. 2005.** Corpora and the compilation of Chinese-English dictionaries. *Lexicographical Studies*, 1: 79-87.
- Zhang, Y. 2007.** *Contemporary Lexicography*. Beijing: The Commercial Press.
- Zhang, Z. and Ni, S. 1912.** *A Chinese-English Dictionary*. Shanghai: The Commercial Press.

Pedro A. Fuertes-Olivera and Henning Bergenholtz (eds.). *e-Lexicography: The Internet, Digital Initiatives and Lexicography*



e-Lexicography.
The Internet, Digital
Initiatives and Lexicography
Pedro A. Fuertes-Olivera and
Henning Bergenholtz (eds.)
London/New York:
Continuum.
2011. xiv, 341 pp.
ISBN 978 1 4411 2806 5

A lexicographer is a divided soul, part scientist, part tool-builder. The scientist is a linguist, wanting to describe the language. The tool-builder wants to help the user find the information they want, the territory of information science. Lexicography is in the intersection.



One might divide the lexicographic process into two parts: analysis, in which we aim to determine the facts, and synthesis, in which entries are prepared. Analysis is linguistics; synthesis is information science.

In this book information science reigns.

Description

The book is the outcome of a symposium in Valladolid, Spain, in 2010, and comprises fifteen chapters by different authors, most but not all associated with the Aarhus School and its function theory of lexicography. Many of the chapters had their seeds in presentations at the symposium.

The book opens with an introduction by the editors Fuerte-Olivera and Bergenholtz. It starts with a conversation with Wiegand, about whether there should be different theories for print and electronic dictionaries, answering that we need one theory that covers all. It then borrows from Gouws's article four agenda items for the book:

- Using databanks from which different types of entry can be extracted
- The mistake of including too much information
- The broadening of lexicographic theory beyond dictionaries
- E-dictionary users are familiar with the internet and the potential it offers: what implications does that have?

They then summarise the remaining articles.

Gouws calls his chapter 'Learning, Unlearning and Innovation' and addresses a colleague's question, "does all the research in theoretical lexicography lead to an improvement in the quality of dictionaries?" He answers yes, for several reasons: because it will now be based on a sound theory;

because lexicography will no longer be a sub-discipline of linguistics; and because, with bold planning, they will embrace the potential of user-generated content.

Henning Bergenholtz, in his chapter, makes the case for research into user needs and describes some results from logfile analysis. His presentation of the time it took users to find information, and its relation to whether they had found the information they wanted, is thought-provoking: in particular the paper dictionaries gave faster access than the electronic ones! While electronic dictionaries potentially allow fast searching, whether they actually do depends on their design, and electronic media introduce many new ways to get distracted, confused and lost. He then describes a set of four monofunctional dictionaries derived from a single database: the pluri-monofunctional model. This was clearly successful, pointing the way ahead for user experiences tailored to information needs. He looks forward to the time when these dictionaries have been more extensively used, so the logfiles will be a large enough body of data to support extensive user research.

In this chapter I did find the review of the literature partial: he says

"... lexicographic interest in the needs of the users ... has been astonishingly scant" (p 31)

not acknowledging the substantial volume of work on the theme, and only picking out one article, by Bogaards, from 1990, to criticise it. One might have expected him to view it as an early, if modest, attempt to start an enterprise that he and others were continuing.

Tarp, in his chapter, provides an appealing vocabulary, already widely adopted, for talking about e-dictionaries: *copycats* (paper dictionaries copied onto digital media), *faster horses* (as copycats, but faster searching), *model T Fords* (first attempts at using what digital media offer) and *Rolls Royces*. The subtitle is 'Towards the Individualisation of Needs Satisfaction' and this is the key to moving Rolls-Royce-wards. He makes a useful distinction between interactive, active and passive methods for individualization, according to whether the user (or the system) takes the driving seat.

This theme is taken further in the following chapter by Bothma. His discipline is information science. He surveys the methods and techniques that modern information technology offers to

lexicography, for filtering and adapting data as held in a database according to user needs. The chapter is full of examples of how particular online dictionaries use particular methods, so provides lots of examples of good (and not-so-good) practice. Online dictionaries are shown as belonging in the same sphere as Google, Booking.com and Amazon, with methods pioneered in those places available for dictionary database and interface designers.

A finding referred to twice in the book is from Leroyer's chapter, that only one quarter of lexicographic works published in 2008-2009 are general language dictionaries, whereas three quarters are

"made up of wordlists and language data organized in dictionary articles, but which nonetheless have nothing to do with language as a scientific object of study" (Leroyer p124)

They are special language dictionaries of one kind and another. This gave me pause for thought. Most lexicography, it suggests, is not linguistic at heart.

But then:

- When ordinary people refer to dictionaries, they mean general language dictionaries like the *Oxford English Dictionary*, *Le Grand/Petit Robert*, *Duden*, *Webster*, etc.
- Purchases and sales of general language dictionaries dwarf those of special language dictionaries.
- Almost all substantial dictionaries (more than 4cm thick, if we take print as a reference point) are general language dictionaries.
- Almost all large lexicographic projects (comprising, say, more than ten people over more than three years) are for general language dictionaries.

The comparison is like noting that there are more local airstrips than international airports in the world, so basing an account of aviation on local airstrips. Numbers of publications alone do not give a good overall picture, and I remain convinced that general language dictionaries are central to the lexicographical firmament.

For lack of space I'll take the chapters by Spohr, Nielsen and Almind, Fuertes-Olivera and Niño-Amo, Bergenholtz and Bergenholtz, and Anderson and Almind together. They present technical challenges, and present examples, of the approach to dictionary-making where a single database meets a range of user needs by selecting only the appropriate information to show in a particular case. The dictionaries referred to are monolingual and multilingual, general language (English phrasal verbs) and special language: music, in the *Danish Music Dictionary* (Bergenholtz and Bergenholtz),

and the accountancy dictionaries, for English, Danish and Spanish (with English as the hub), which Fuertes-Olivera and Niño-Amo's and Nielsen and Almind's chapters describe. The project looks good, though it was disappointing to find no references to the extensive discussion of issues relating to hub-and-spokes models and to translation mismatches, for example in Janssen (2004). Also their example displays the hazards of lexicography moving into encyclopedic territory, with Fig 7.4 showing a definition for the accountancy term *deemed cost*, which begins

"an amount used instead of cost or depreciated costs at a specific date. Any following amortisation or depreciation is made on the assumption that ..."

I googled the term (in inverted commas, to get an exact match) and the second hit stated

'**Deemed cost**' is a surrogate for cost at a given date. For example if a building is purchased at \$100000 this is cost and also the **deemed cost** at that given date. ...

Accurate, extensive encyclopedic entries are very often already available, and very easily accessible via google, as here. A case has to be made for what value lexicographers are adding.

Lew, in his chapter, is good to his title, 'Online Dictionaries of English.' He introduces some useful criteria for thinking about online dictionaries – for example individual (standalone) dictionaries, *vs* clusters of dictionaries (eg, from the same publisher and on the same website) *vs* portals (websites that give links to lots of dictionaries) *vs* aggregators (which offer entries for a word from unrelated dictionaries, e.g., Dictionary.com). He then reviews a large number of online resources, mainly classifying and describing what is out there, sometimes evaluating.

In contrast Sanchez and Pascal review the case of online monolingual Spanish dictionaries and find just four, all closer to copycats than Rolls Royces. They then develop an account of what could be done, making use of the potential of the electronic medium to give a very rich account of a word within their Lexical Constellation Model.

Verlinde's chapter describes his *Base Lexicale de Français*: here, as also discussed at the e-lexicography conference in Bled, Slovenia, in November 2011, is already something we might call a Rolls Royce. As the chapter describes, and the Bled presentation demo'd, this is a dictionary – extensive, and widely-used – which applies many lessons of what can be done, online, with current technology, to customise according to the user's characteristics and information need.

eLex 2013 Electronic lexicography in the 21st century

The third international conference on electronic lexicography, eLex 2013, will take place in Tallinn, Estonia, on 17-19 October 2013. It is organized by the Institute of the Estonian Language, Estonia, and Trojina, Institute for Applied Slovene Studies, Slovenia.

The previous two eLex conferences (at Louvain-la-Neuve, 2009, and Bled, 2011) have shown that there is huge interest for such events from both academia and industry. This conference will be yet another great opportunity for presenting state-of-the-art projects and research in the fields concerned with e-lexicography, as well as discussing topical issues and networking.

The conference will be preceded by SKEW4, the 4th annual Sketch Engine workshop, which will take place on 16 October 2013 at the same venue, and is organized by Lexical Computing Ltd.

More information is available on the conference website, <http://eki.ee/elex2013>.

Iztok Kosem, Jelena Kallas
Co-chairs of the eLex 2013 organizing committee

Heid takes a well-established method from information science – the usability study – and applies it to dictionaries. He provides an overview of usability studies and shows how they can be applied in lexicography – and indeed how they can often provide an answer to the questions that motivate much of the book: what queries, forms, data and structure – give the user (with a particular information need) most help. The study in this chapter compares three ‘live’ online dictionaries, and suffers from the methodological problem that there are many differences between the three, so it is not clear which differences resulted in a more or less successful user experience. Heid notes the problem and discusses, as ‘further work,’ a model where the ‘dictionaries’ to be compared are closely controlled so there is just one parameter according to which they differ. He and his co-workers have since lived up to this promise, with a study presented at the Bled conference, in which they do just that.

The book concludes with ‘Ten Key Issues,’ a chapter which summarises the discussions from the Valladolid workshop, edited by Samaniego Fernández and Cabello de Alba.

The Aarhus School and linguistics

The Aarhus School denigrates linguistics with vigour:

“Linguists were the princes of meta-lexicographic discussions, and meta-lexicography and practical lexicography were subsections of the work done by these linguists. This era in the history of lexicography can rightfully be regarded as representing a form of *linguistic colonialism*.” (Gouws p22)

“... linguists who also masquerade as lexicographers...” (Bergenholtz and Bergenholtz p188)

Let me state my vantage point. My friends and colleagues Sue Atkins and Michael Rundell chose to include “practical” in the title of their *Oxford Guide to Practical Lexicography* and state:

“This is not a book about ‘theoretical lexicography’ – for the good reason that we do not believe that such a thing exists.” (p4)

Authors of most chapters in this book, however, state the need for theoretical lexicography. Unsurprising then, that my friends are its enemies, already identified as such (along with Henri Béjoint) in the introduction (p8), and explicitly denounced in Bergenholtz and Bergenholtz’s chapter (p189). I won’t pursue the question “is there such a thing as theoretical lexicography” as I fear it would be an arid discussion on

the meaning of *theory*.¹ I hope it is not contentious to say, theory or no theory, the Aarhus School is concerned to place the user’s information needs centre stage, and the ways and means and implications of doing that are the central theme of the book. This is somewhat in spite of its title, ‘e-Lexicography’. In the 21st century, pretty much all lexicography is e-lexicography, in the senses that the writing is based on digital evidence, takes place on a computer and employs dictionary writing systems, and most users will be accessing the data through a computer or other electronic device, and there is little more needing saying. The subtitle, ‘The Internet, Digital Initiatives and Lexicography’ does not add much. A more informative title would have been ‘Putting user needs at the centre of lexicography.’ The relation between user needs, and having the dictionary data in electronic form, is that we can show different users different things, according to their information needs. This is the link between the ‘e’ and the real topic of the book.

Another perspective on the role of linguistics in lexicography is this. The chapters of this book are mostly concerned with delivering information to the user. This is of course fully legitimate, and the questions “how much information” and “which information, when” are good ones – but none of the chapters discuss the risk of delivering false or misleading information. They proceed as if the truth were known and the database contained all and only correct material. Would that it were so! A careful review of any dictionary – see for example Hanks on *Merriam-Webster’s Advanced Learners English Dictionary* (*International Journal of Lexicography* 22.3, 2009) – will uncover points at which it is likely to mislead and confuse. Even in this book, where presumably the authors have chosen examples with care, I noticed a lexicographical bloomer. On pp 211-213 we have an analysis of the English phrasal verb *call back*. It is given six meanings of which the sixth is given the example “I cannot call his face back.” As an English native speaker, I go *eeeeeeugh*. This is blazingly wrong. (We might say “I cannot recall his face.”) A little research revealed that this ‘example sentence’ exists in a number of dictionaries and translation tools: a dictionary error that has been copied and recopied from dictionary to dictionary.

A simple and central case, in both general-language and special-language, concerns variability of set phrases and idioms. Somehow, if the user is trying to decode “quaking in one’s boots” – or even “quivering on one’s Doc Marten’s” we would like to direct them to the idiom that

(debatably) has as its core form “shaking in one’s shoes” (example from Moon 1994). This is a hard problem: it is both hard to work out how to represent the facts in a usable way, and then it is hard to work out, for each individual expression, what the facts are: different phrases allow variability to different degrees, in different ways. In Bergenholtz’s article on his dictionary of fixed expressions, he considers the Danish *på vulkaner*, *være på vulkaner*, *danse med vulkaner* and several other variants but does not discuss the challenge of how the lexicographer might discover the range of variation of the phrase, or of how this might be represented in the database or shown to the user.

To come back to the Venn diagram with which I opened the review, the linguistics part is for analysis. If we had a database containing all the facts and generalisations about the behaviour of all the words and phrases of the language, optimally structured, then we wouldn’t need linguistics. But we don’t. That is what linguistics aims to do, and what the lexicographer, when working on a particular word, aims to do for that word.

One surprising and disappointing aspect of the book is the poor standard of production. Given the topic, one would have expected a book where figures are readable, and where thought had been given to the best typography for presenting complex dictionary entries. But it took a magnifying glass to read the text on screenshots in Chapter 10, and throughout the book complex lexical entries are presented in plain text, with no use of font, font size or weight, indentation or other formatting to make them digestible. Running headers relate to chapter names rather than author names.

Conclusion

This is largely an Aarhus School book, with a number of interesting and useful chapters exploring and developing the model of the pluri-monofunctional dictionary.

While I find the Aarhus School’s attention to the information-science side of lexicography often useful and enlightening, I find its attacks on the linguistic side puzzling. The Aarhus School doubts the relevance of corpora for lexicography (explicitly, in the concluding chapter, p309). But you need corpora to get the facts right.

Notes

¹ We Anglo-Saxons are often dubious about grand statements of theory. When Wittgenstein pronounced “Whereof one cannot speak, thereof one must be silent” his English friend and colleague Frank Ramsey (no intellectual slouch, a founding figure in mathematical economics and decision theory) responded “What we can’t say we can’t say, and we can’t whistle it either.”

References

- Atkins, S. and Rundell, M. 2008. *Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Janssen, M. 2004. Multilingual Lexical Databases, Lexical Gaps, and SIMuLLDA. *International Journal of Lexicography*, 17.2: 137-154
- Moon, R. 1994. *Fixed Expressions and Text*. PhD thesis, School of English, University of Birmingham.

Adam Kilgariff

Lexical Computing Ltd
adam@lexmasterclass.com

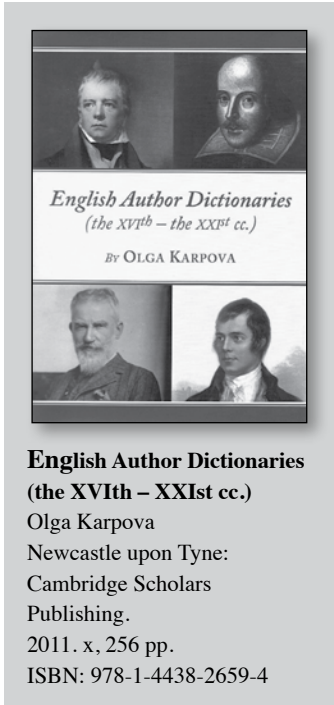
Olga Karpova. *English Author Dictionaries (the XVIth — the XXIst cc.)*

Olga Karpova’s *English Author Dictionaries* begins with her wonderment about why the English author dictionary has been neglected in dictionary research in spite of the fact that it “has at its disposal about 300 titles of linguistic and encyclopedic reference works to single and complete works of more than eighty writers” (p. ix).

I am much in agreement with this sentiment. We are well aware of the fact that, over the centuries, the “author’s dictionary,” or the reference work “which provides information on the vocabulary of a specific author” (Hartmann and James

2001, 10), and the “dictionary of authors,” or the reference work “providing literary, biographical and sometimes encyclopedic information about a selection of authors” (Hartmann and James 2001, 43) concerning English writers have been published in great number. We also know that a copious number of reviews have been written for such works.

However, when it comes to the point of how often they have been researched from a holistic perspective, it is quite another story. For instance, even in the voluminous *The Oxford History of English Lexicography*



(Cowie (ed.) 2009), which encompasses various types of dictionaries, the author dictionary is scarcely treated – although a chapter for the dictionary of quotations is provided, and this is one type of author dictionary. When observing this situation, I am inclined to wonder, like Karpova, why research on the English author dictionary has hardly been performed, despite its great value from a philological and linguistic viewpoint.

In this sense, Karpova’s monograph, which deals with the historical development of the author dictionary from a macroscopic perspective, may be regarded as highly innovative, having the potential to open up a new and significant area in the research of English lexicography. For this reason, I, a researcher of the history of English lexicography and of the history of the language, would like to celebrate its publication. However, at the same time, it is often the case that an innovative work is a prototype for the posterity. Therefore, I will analyze the volume in the hope for improvement of the research in the English author dictionary, aiming to clarify what Karpova has achieved and what is left for the future researcher. To be concrete, I will, at first, briefly refer to Karpova’s use of works by authorities related to her research, and, then, analyze her selection of the author dictionary based on which she performed her research for the volume.

Firstly, in reference to Karpova’s use of works by authorities, her volume can be divided into four parts, except for the preface and introduction: (1) the body text (pp. 10-173), (2) the list of books and papers of her reference (pp. 175-208), (3) the list of author dictionaries for British writers (pp. 209-253) and (4) the list of the names of British writers (pp. 254-256), out of which the second part can be subdivided into two parts: the list of author dictionaries for writers outside of Britain, and that of research books and papers by authorities which Karpova referred to in writing the volume.

On this premise, if we are to see how Karpova used works by authorities, we can know from the bibliography entitled ‘Other References’ in the second part that she referred to 128 books and papers by 90 authorities in all, with the exclusion of 5 items on the Internet whose authors are not clear. When we collate such books and papers with the contents in the body text, we notice Karpova has finely used the works to support her research, hardly making critical remarks about them. This attitude of hers seems to clearly show the fact that Karpova has had the intention to open up a new horizon in the historical study of lexicography, trying to make full use of the fruits of related research in the past.

Actually, there are quite a few cases in the body text where we can effectively know the books and papers to be referred to in regard to the terms and concepts of lexicography, as the following two instances indicate:

“[...] the problem of choosing Shakespeare’s [...] edition is considered to be of primary importance and difficulty in author lexicography (Andrews 1987, 277-279; Benko 1968, 649; Greg 1942; Parker 1945; Schaaber 1947; Culpeper 2004, 17-73; [...]).” (p. 11)

“Glossary is one of the oldest lexicographic forms in English national and author lexicography (Considine, Iamartino 2007; Dill 1959, 340-361, 369-375; Hüllen 1999; López 1977, 151-159; Müller 2001).” (p. 35)

Such will be judged to be a sound method and should be applied in the future research of the English author dictionary.

Then, on what and how many English author dictionaries has Karpova based herself in performing her research? With regard to this point, the very core of Karpova’s volume seems to lie in the list of author dictionaries for British writers which she entitled the ‘Index of Dictionaries to the English Writers (the XVIth – the XXIst cc.)’ (pp. 209-253). (It should be noted that the word ‘Index’ here only means the “(systematic) list,” having no relevance to the page number; this volume does not provide an index with the page number of the body text, concerning which I will comment later.) This is because Karpova remarks that she has (studiously) “been adding various types of author dictionaries to” the list since 1973, “working in different libraries in Russia and abroad” (p. 209). If this is the case, and if Karpova’s volume is essentially based on the dictionaries cited in the list, as it seems to be, we will not be able to make a correct assessment of the volume without analyzing it, thus revealing how much she has achieved and what challenges are left

Table 1. Karpova’s selection of dictionaries on English authors 1

Type	16th c.	17th c.	18th c.	19th c.	20th c.	2001-2010	Total
Lin.	2	–	4	28	155	17	206
En.	–	–	–	3	11	42	56
C+P	–	–	–	4	35	9	48
Q+P	–	–	–	6	17	5	28
Bible	3	2	1	–	1	12	19
Beowulf	–	–	–	–	5	–	5
E+I	–	–	–	–	53	–	53
Total	5	2	5	41	224-277	85-138	415

for the future in the research field of the English author dictionary.

In the list of English author dictionaries, Karpova classifies the relevant dictionaries into 6 types, for which I will use the

following abbreviations, respectively, in my analysis below:

Lin. for Linguistic Dictionaries
En. for Author Encyclopedias, Guides and Companions

Table 2. Karpova's selection of dictionaries on English authors 2

Author	Lin.	En.	C+P	Q+P	E+I	Author	Lin.	En.	C+P	Q+P	E+I
Arnold, M.	1	1	–	–	–	Austen, J.	1	2	1 (1)	–	1
Bacon, F.	1	–	–	–	–	Blake, W.	1	1	1	–	1
Bronte, C.	1	–	– (1)	–	–	Bronte, E.	–	–	– (1)	–	–
Browning, R.	1	2	–	–	–	Burns, R.	2	1	–	–	1
Byron, G.	1	1	–	–	–	Chaucer, G.	13	14 (1)	4	–	5 (1)
Christie, A.	–	–	1	–	–	Coleridge, S.	1	1	–	–	–
Congreve, W.	1	–	–	–	–	Conrad, J.	5	2	–	–	–
Cowper, W.	1	–	–	–	–	Daniel, S.	– (1)	–	–	–	–
Darwin, C.	1	–	–	–	–	Dickens, C.	2	2	5	–	2
Donne, J.	1	–	–	–	–	Drayton, M.	– (1)	–	–	–	–
Dryden, J.	2	–	–	–	–	Eliot, G.	–	–	2	–	–
Franklin, B.	1	–	–	–	–	Frost, R.	–	1	–	–	–
Goldsmith, O.	1	–	–	–	–	Gray, T.	1	–	–	–	–
Hardy, T.	–	1	4	–	–	Herbert, G.	1	–	–	–	–
Herrick, R.	1	–	–	–	–	Hopkins, G.	2	–	–	–	–
Housman, A.	1	–	–	–	–	James, H.	–	–	1	–	–
Johnson, S.	1	–	–	–	–	Jones, D.	1	–	–	–	–
Joyce, J.	9	2	2	–	1	Keats, J.	2	–	–	–	1
Kipling, R.	–	–	1	–	–	Lamer, S.	1	–	–	–	–
Larkin, P.	1	–	–	–	–	Lawrence, D. H.	1	–	–	–	–
Lyly, J.	1	–	–	–	–	Malory, T.	1	–	–	–	–
Marlowe, C.	2	–	–	–	–	Marvell, A.	1	–	–	–	–
Milton, J.	11	–	1	–	3	Pope, A.	2	–	–	–	–
Rawling, J. K.	–	–	–	–	–	Scott, W.	–	1	2	–	–
Shakespeare, W.	104	33	18	16	25 (1)	Shaw, B.	3	2	1	1	–
Shelley, P. B.	1	1	–	–	1	Sidney, P.	1 (1)	–	–	–	–
Spenser, E.	1 (1)	–	–	–	–	Swift, J.	2	–	–	–	1
Taylor, E.	1	–	–	–	–	Tennyson, A.	4	–	1	–	– (1)
Thackeray, W. M.	–	–	1	–	–	Thomas, D.	1	–	–	–	–
Thomas Kyd	1	–	–	–	–	Tolkien, J.R.R.	–	–	–	–	2
Traherne, T.	1	–	–	–	–	Webster, J.	1	–	–	–	–
Wells, H. G.	–	–	1	–	–	Wordsworth, W.	1	–	–	–	– (1)
Wyatt, T.	1	–	–	–	–	Yeats, W. B.	2	–	–	–	–

C+P for Dictionaries of Characters and Place Names [concerning the works of English authors]

Q+P for Dictionaries of Quotations and Proverbs

Bible and *Beowulf* for Dictionaries to the Bible and *Beowulf*

E+I for Electronic and Internet Dictionaries to English Writers

Under this premise, the dictionaries cited in the list can be analyzed from two perspectives: the one is an overall perspective of Karpova's general tendencies in the selection of relevant dictionaries, and the other is a specific perspective of authors treated in the dictionaries.

In the first place, from an overall perspective, my analysis of the list has resulted in the information as shown in the following table, in which each figure indicates the number of relevant dictionaries.

This table shows such facts as the following:

1. Karpova, in the list, has cited, in all, 415 dictionaries that are judged to form the basis of her volume.
2. The 'Linguistic Dictionaries' (Lin.) account for 206, or 49.6%, of the 415. As compared to this, the 'Dictionaries of Quotations and Proverbs' (Q+P) account for 28, or 13.6%, of all.
3. From a chronological viewpoint, the dictionaries published in the twentieth century account for 224 to 277, or 54.0% to 66.7%, of the 415. As compared to this, the dictionaries published between the sixteenth century and the nineteenth century account for 53 (5+2+5+41), or 12.78%, of all.
4. Karpova seems to have greatly succeeded in searching for the 'Author Encyclopedias, Guides and Companions' (En.) published in the twenty-first century; they comprise 42 in number, or 30.4%, out of all 138 relevant dictionaries published during this period.

I am not certain to what extent such facts

reflect the trend and transition in the publishing world, but it may safely be said that Karpova's research is particularly meaningful in knowing the situations of the 'Linguistic Dictionaries' and the 'Dictionaries of Quotations and Proverbs,' as well as the 'Electronic and Internet Dictionaries,' which were published in the twentieth and twenty-first centuries. And, at the same time, it may be said that the research in the English author dictionary is expected to develop in the future by further investigating its situations until the nineteenth century.

In the second place, from a specific perspective, my analysis has revealed that Karpova, in making the list, has selected 382 dictionaries in which a total of 66 specific authors are dealt with; in this analysis, I have excluded dictionaries whose types are 'Bible' and '*Beowulf*,' as well as those with titles such as *Forgotten English* and the *Collins Gem Dictionary of Quotations* in which the name of a specific author is not indicated, having taken up, instead, the dictionaries with titles such as the *Cambridge Companion to Chaucer* and *Who's Who in Dickens*. Out of the 382 dictionaries, 206 are 'Linguistic Dictionaries,' 68 are 'Author Encyclopedias, Guides and Companions,' 47 are 'Dictionaries of Characters and Place Names,' 17 are 'Dictionaries of Quotations and Proverbs,' and 44 are 'Electronic and Internet Dictionaries' – here again, we are reminded of Karpova's close attention to the 'Linguistic Dictionaries'.

Then, my analysis has resulted in Table 2, each figure showing the number of relevant dictionaries, and the indication "–" meaning that there are no relevant dictionaries selected. In this table, there are cases where the figure is in parentheses, as seen in the space for En. of "Chaucer, G." and that for Lin. of "Daniel, S." This shows the fact that the titles of relevant dictionaries indicate more than one specific author, as the *Companion to Chaucer and His Contemporaries* and the *Concordance*

Table 3. Karpova's selection of dictionaries on English authors 3

Author	16 th c.	17 th c.	18 th c.	19 th c.	20 th c.	21 th c.	E+I
Austen, J.	–	–	–	–	3 (4)	1	1
Chaucer, G.	2	–	–	4 (1)	13	12	5 (1)
Conrad, J.	–	–	–	–	5	2	–
Dickens, C.	–	–	–	1	7	1	2
Joyce, J.	–	–	–	–	12	1	1
Milton, J.	–	–	–	4	8	–	3
Shakespeare, W.	–	–	4	25	96	46	25 (1)
Shaw, B.	–	–	–	–	7	–	–
Tennison, A.	–	–	–	1	4	–	– (1)

Table 4. Karpova's selection of dictionaries on American authors

Author	~'70	~'80	~'90	~'00	'01~	Author	~'70	~'80	~'90	~'00	'01~
Crane, H.	–	2	–	–	–	Crane, S.	1	–	–	–	–
Dickinson, E.	1	–	–	–	1	Dreiser, T.	–	–	–	–	1
Faulkner, W.	4	6	1	2	1	Fitzgerald, S.	–	–	–	–	1
Frost, R.	–	1	–	–	1	Hemingway, E.	–	1	1	2	4
Hughes, L.	–	1	–	–	–	Moore, M.	–	1	–	–	–
Poe, E.A.	–	–	–	–	2	Pound, E.	–	1	–	–	–
Stevens, W.	1	–	–	–	–	Twain, M.	–	–	–	–	1

to the *Sonnet Sequence of Daniel, Drayton, Shakespeare, Sydney and Spenser*. The table reflects that there are 12 such cases in all.

In Table 2, the fact is immediately noticeable that Karpova has selected an exceedingly large number of dictionaries on William Shakespeare, which actually account for more than 51% of all dictionaries whose titles indicate the names of specific authors. This can be regarded as due to the fact that Karpova is a talented authority on Shakespeare, though she says in the body text that “[i]t should be noted that Shakespeare scholarship has a profound set of theoretical works where lexicography plays an important role” (p. 10). Karpova provides the bibliography entitled the “Publications by Olga Karpova” where she listed 92 books and papers, of which 25 concern Shakespeare. She also provides the sub-sections ‘Shakespeare Concordances’ (pp. 11-17), ‘Russian Monolingual and Bilingualized Shakespeare Dictionaries’ (pp. 131-132), ‘Shakespeare Quotable Calendars’ (pp. 151-154), and ‘Bilingual Russian Shakespeare Dictionaries’ (pp. 161-162). She does not do this for any other author. For this reason, Karpova's volume is judged to be especially valuable concerning the historical development of dictionaries on Shakespeare.

In contrast, the dictionaries on other authors seem to be relatively small in number. I do admire the herculean effort of Karpova to have read an extensive range of the English author dictionary and to have tried to judge the quality of each by herself. At the same time, we may see here a problem in the research of the author dictionary that will hardly be solved by an individual: it will be far from possible for any one single person to be a specialist of dozens of English authors and evaluate the dictionaries on them. Instead, it might not be such a difficult task for the interested authorities to point out more than one specialized dictionary worthy of being selected, concerning, for instance, William Congreve, Charlotte and Emily Bronte, John Donne and D.H. Lawrence,

or William Wordsworth, bringing out abundant potential in the research of the English author dictionary. In this sense, the result of my analysis that is revealed in Table 2 may also suggest the necessity of forming a team for the research in the author dictionary.

Saying this, I, of course, never mean that such a limit of the individual capacity mars the significance of Karpova's volume. Especially, she has selected more than five dictionaries for each of nine authors (Jane Austen, Geoffrey Chaucer, Charles Dickens, James Joyce, John Milton, William Shakespeare, Bernard Shaw, Alfred Tennyson) as follows, providing a clear example of how the survey of the English author dictionary should be.

By the way, the volume's title *English Author Dictionaries* has dual meanings; the one is the “dictionaries on British authors” and the other the “dictionaries on authors who wrote their work in the English language.” In this regard, I formed Table 4, concerning dictionaries on American authors, all of which were published in the 20th and 21st centuries, based on Karpova's other list on ‘Dictionaries, Encyclopedias, Guides and Companions to American Writers’ (pp. 176-178), to which, however, we can only scarcely see her reference to such dictionaries in the body-text of the volume.

A regrettable aspect is the lack of an index which, according to Hartmann and James (1998, 72), should allow “the user access to each [relevant description in the body-text] by means of page numbers.” Providing this type of index for the names of authorities, the English authors and the dictionary titles would make this volume much more usable and valuable for its readers.

Lastly, Karpova remarks, concerning the origin of the English author dictionary: “Historical roots of English author lexicography go back to the XVIth c., when concordances of the English translation of the Bible and glossaries to the complete works of G. Chaucer were published” (p. 171). However, if we leave aside the classification of monolingual and bilingual dictionaries, the origin of the dictionary in Britain may

be traced further back to the eighth century, closely connected to the very beginning of its history of lexicography. This is because, in a sense, British lexicography began with Latin-English dictionaries, the *Epinal Glossary* (early 8c.?), which treated 1,186 words, and the *Corpus Glossary* (early 8c.?), which treated 2,175 words, both compiled anonymously, for the commentaries on St. Augustine's homiliary. (Besides, there are facts that the *Leiden Glossary* (9c.?) and the *Erfurt Glossary* (late 9c.?), which also were anonymously compiled, followed the two glossaries with the same purpose, and that, in this context, Thomas Elyot compiled the *Dictionary of Syr Thomas Eliot Knight* (1538), a Latin-English dictionary, quoting from the works of Marcus Tullius Cicero, Publius Vergilius Maro, Gaius Julius Caesar, and so on.) Such a perspective may also be necessary for the future development of the promising and creative research field of the author dictionary.

In line with Samuel Johnson's maxim, "Lay the foundation, and leave the superstructure to posterity," with this volume Karpova has opened up and laid the foundation of new research

in lexicography – that of the English author dictionary, which seems to be highly significant from the viewpoint of philology and linguistics. There are high expectations that research in the field will significantly develop in the future, and I believe Karova's volume will provide a good starting point for this.

References

Cowie, A.P. (ed.). 2009. *The Oxford History of English Lexicography*. New York: Oxford University Press.

Hartmann, R.R.K. and James, G. 1998 (revised, 2001). *Dictionary of Lexicography*. London & New York: Routledge / Taylor and Francis.

Hayashi, T. 1968. *Eigo Jisho Hattatsu-shi [A History of English Lexicography]*. Tokyo: Kaibunsha.

Stein, G. 1985. *The English Dictionary before Cawdrey*. Lexicographica Series Maior, 9. Tübingen: Max Niemeyer Verlag.

Kusujiro Miyoshi

Soka Women's College, Tokyo
kusujiromiyoshi@gmail.com

Olga Timofeeva and Tanja Säily (eds.). *Words in Dictionaries and History.* *Essays in honour of R.W. McConchie*

As print journals are replaced by journals online, readers increasingly click on articles they want to read rather than flip through pages of scholarship in which a fact or a figure, a graph or a quotation, might arrest them for a moment. Scholarly reading today is all about efficiency, but efficiency has its costs — we rarely know all of what we might know, or even what, given our interests, we need to know. The festschrift is inevitably miscellaneous, and so it begs to be read in leisure, with an open mind. Alas, its inefficiency has all but killed it. Though there are exceptions, of course, too many festschriften have gathered too many decades of dust on too many library shelves. Librarians are reluctant to buy them, and most publishers have turned their backs on them. Thankfully, however, some have not, John Benjamins prominent among them, a recent volume of whose series Terminology and Lexicography Research and Practice, titled *Words in Dictionaries and History. Essays in honour of R.W. McConchie*, edited by Olga Timofeeva and Tanja Säily, is an

outstanding specimen of the genre.

The contributions to *Words in Dictionaries and History* are loosely connected insofar as the volume "aims to represent and advance studies in historical lexis," as the editors put it. They all also represent areas of particular concern to R. W. McConchie, who has long been a leading scholar of Early Modern English lexis and lexicography, especially medical vocabulary and medical glossaries. He began his career focused on Old English language and literature and is now, among many other things, writing about the language of Jane Austen's novels. Thus, it should be no surprise that the contributions to his festschrift cover a challenging array of discrete subjects. Patient readers will have gaps in their knowledge filled, for instance, by Anatoly Liberman's etymology of *yeoman*, or Samuli Kaislaniemi's discovery of a rare word for sex, as well as unexpected possibilities raised, for instance, by John Considine's recovery of a lost (or at least very well hidden) dictionary project, or Joshua Pendragon and Maggie

Scott's skirmish with the *Oxford English Dictionary* over the lexicon of swordplay. It should certainly warm Rod McConchie's heart that he has inspired work of such breadth and interest, not just lexical but cultural, and of such excellence. In both respects, contributors are simply following his example.

To explain the book's overall structure, one cannot do better than the editors: "The articles fall into two parts. The first part focuses on the history of dictionaries, analysing them in diachrony from the first professional dictionaries of the Baroque period via Enlightenment and Romanticism to exploring the possibilities of the new online lexicographical publications. The second part looks at the interfaces between etymology, semantic development and word-formation on the one hand, and changes in society and culture on the other." I know what the editors mean, but the value of the book, I think, is in having the historical, lexicographical, and linguistic material, as well as the social and cultural, all on one and the same hand, or, if separate hands are necessary, with the fingers of those hands intertwined.

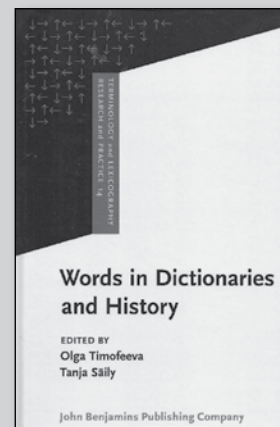
The contributions focused on dictionaries are, of course, no less culturally interested than those focused on words. In "*The Flores of Ovide* (1513): An early Tudor Latin-English textbook," Ian Lancashire introduces us to a mostly overlooked glossarist, Walter, and the unique copy of the early printed book, in which his "complementary English—Latin and Latin—English glossaries" are preserved. Yet the article is not merely bibliographical or lexicographical, but also about the role of glossaries in sixteenth-century English pedagogy. Jukka Tyrkko, in "'Halles Lanfranke' and its most excellent and learned expositive table," hopes "to provide a description of an early English glossary, as well as shed some light on its compiler John Halle." Along the way to doing so, he "reinforce[s] the notion that the medical profession was in many ways in the vanguard of English dictionary-making," a core subject of McConchie's scholarship. But again, the argument exceeds bibliography and lexicography, for Halle (or as some might know him better, "Hall") was a poet, a composer, a biblical translator, and a reminder that none of us is just one thing and the variety of our experience informs our cultural productions, so, Tyrkko suggests, "Halle's work on religious texts must have informed his medical and perhaps especially his lexicographical work, particularly when it came to appreciating the importance of lexical precision." John Considine considers the origin and fate of "John Lane's *Verball*:

A lost Elizabethan dictionary," which was proposed as an aid to the writing of quantitative verse in English. Though, like Lancashire's and Tyrkko's, focused on a particular book, Considine's contribution also extends to an intellectual tradition, that of Latin and English guides to prosody.

Each of these articles is excellent, and Considine's is a perfect specimen of its kind. First, Considine is a master of the note, and each section of his article accomplishes more, this jealous reader observes, than it has any right to do. Second, in trying to discover who the author of the anonymous *Verball* was, Considine explores an array of sixteenth-century genealogical connections with an almost savage zeal. If any of John Lane's family and other connections had hoped to bury their relationships to the *Verball*'s author, Considine has unearthed them beyond burying again.

This is not to suggest that the articles are without any weaknesses. Lancashire at times seems rather breezy. Can he really know that *The Flores of Ovide* was "the first and only intrusion of Ovid's *Ars Amatoria* into sixteenth-century grammar school education"? Sometimes, I would have appreciated a citation or note: "Colet and William Lily ... in a small committee that also included Thomas Linacre devised a grammar textbook, the so-called *Short Introduction of Grammar* (STC 15610.10; Allen 1954; Flynn 1943)," Lancashire informs us, helpful references in place; but then he declares, "Colet and Lily taught English school children for several centuries," and we are left to rely on his authority, without further explanation. Lancashire has long been a leading scholar, and of course we can rely on his very deep, precise knowledge of Early Modern affairs, yet he might not presume this so easily.

Tyrkko, on the other hand, does not always write with a sure grasp of his subject: Halle's *The Courte of Vertue* (1565) was published as a pious alternative to the poetic miscellany called *The Courte of Venus*, first published in 1537 or so, and first attacked by Halle in *Certayne Chapters taken out of the Proverbes of Solomon* (1549/1550) — like Lancashire's Colet, Halle had no time for *ars amatoria*. We are told in a note that "*The Courte of Venus* is a coterie compilation of poems critical of the church. It has been attributed, in its entirety or in part, to Chaucer (Fraser 1952)." While some poetry of the Chaucer apocrypha does appear in *The Courte of Venus*, the article by Russell Fraser that Tyrkko cites does not mention those attributions. Importantly, *The Courte of Venus* is certainly NOT attributed entirely to Chaucer, and Fraser's article discusses instead Hall's parodies of poems



Words in Dictionaries and History. Essays in honour of R.W. McConchie

Olga Timofeeva and Tanja Säily (eds.)

Terminology and

Lexicography Research and Practice, 14

Amsterdam/Philadelphia: John Benjamins.

2011. xiv, 292 pp.

ISBN 978 90 272 2338 8

e-ISBN 978 90 272 8690 1

by Thomas Wyatt the Elder. *The Courte of Venus* is notable, not because it raised Hall's ire, but because it is the first printed poetic miscellany in English, for proof of which one can consult Fraser's definitive edition (1955) of the book's three extant fragmentary copies.

There are a few other lapses. When reporting the entry for *chirurgery* in Halles' glossary to Lanfranc, Tyrkko remarks that the definition there is "rather meaningless," but I don't quite see how it is: "**Χειρουργία** is sayd of euery arte, whose function consisteth in manuelle action or handye operation," as opposed, say, to the arts of the apothecary, a quite meaningful distinction. Finally, it is hard for me to understand why one contributing to a festschrift for McConchie would refer to Richard Howlet as *Huloet*, though many others have done so, since McConchie himself has written decisively on the lexicographer and his name (see, for instance, McConchie 2007).

Readers of this review will think I am nit-picking. Really, are these the only criticisms I have of the book as described so far? It says something about the book's general excellence that I can do no better. None of my niggling concerns really diminishes Tyrkko's thorough historical and partly forensic analysis (following Julie Coleman and Sarah Ogilvie in the *International Journal of Lexicography*, 2009) of the glossary appended to Halle's translation of Lanfranc. His focus, after all, is not on Halle's poetry but on his treatment of medical lexis. Very often, too, Lancashire's magisterial breeziness makes for good reading — Lancashire doesn't get bogged down in learned citations and historiographical controversy, but instead tells a good story and often delights his readers with a clever phrase.

While their scope may seem narrow at first glance, the several contributions actually reach to significant cultural issues. Considine's central figure, John Lane, aspired to be a lexicographer before lexicography was a plausible target of aspiration. We should look for others who did the same; we should wonder, in historical context, just what sort of aspiration it was, and what it indicates about Early Modern English society. Similarly, Gabriele Stein, in "The linking of lemma to gloss in Elyot's *Dictionary* (1538)," focuses on a slight feature of entry structure, but this leads inevitably to interest in the logic and developing rhetoric of dictionary entries in what would prove a rhetorical age, no less relevant today than at the advent of the dictionary genre. Elizabeth Knowles, writing about "*Chaos and old night*: A case study in quotation usage," explores

quotation, transmission, and alteration of Milton's famous phrase, raising the question of when a quotation is allusive and when it detaches from its source and is used idiomatically, unallusively, effectively how a language is infused with quoted material once "owned" by authors but finally by speakers. The articles may seem like small hooks of scholarship, but thereby hang some big tales.

In addition to the articles already noted, the volume's first section includes Giles Goodland's "Music amidst the tumult," which considers the ways in which making *A Dictionary of the English Language* (1755) required that Johnson "repress" his "poetic side," yet another chapter in the developing distinction of lexicography among the genres of English letters. Julie Coleman's "Online dictionaries of English slang," proposes that "online slang dictionaries can be categorized along a spectrum from the static to the dynamic," the former authored in traditional ways and put online in a more or less finished state, the latter inviting and responding to user contributions, which, while "they do not generally fulfil the requirements of traditional dictionary users in terms of content, quality or reliability," nonetheless offer information that slang lexicographers can use as material to determine frequency, distribution, origins and semantic development." All of this signals a newly symbiotic relationship between dictionary makers and dictionary users, and the development of the web as a platform for amateur lexicography, both of whom shifts in our notion of the dictionary nearly as significant as establishment of the dictionary as a pedagogical tool and a target of literary aspiration 400 or so years ago.

From Walter the Almost Anonymous Glossarist to *Urban Dictionary* — are they really so far apart? After all, aren't there notable similarities between Walter and someone identified only as Nony, who entered *chester* 'pedophile' (< *Chester the Molester*) in *Urban Dictionary* on 13 March 2005, in spite of their obvious differences? The first half of *Words in Dictionaries and History* is not a systematic study of its subject, but it is nonetheless informative and challenging, especially to those already immersed in lexicography and its history, and fully repays an afternoon's reading. The second half of the book rises to the rather lofty standard set by the first.

First in the second half is Matti Kilpiö's "Old English etymologies in Christfrid Ganander's *Nytt Finskt Lexicon* (1787)," which evaluates the adequacy of Ganander's etymologies of Finnish lemmata when they include Old English elements. Ganander comes through this scrutiny well for a

lexicographer working without benefit of the New Philology, and Kilpiö makes the case that Ganander's work on English etymology should not be dismissed. He is exactly the sort of lexicographer overlooked by most but consulted by Anatoly Liberman, who hopes to balance, if not replace, the "dogmatic" tendency of most English etymology with an "analytic" one in his *Analytic Dictionary of English Etymology* (or *ADEE*; 2008-, one volume to date). Liberman's "The etymology of the word *yeoman*," which immediately follows Kilpiö's piece, is what W. W. Skeat called a "scorched-earth" etymology (see *ADEE*, p. xxv-xxvi): it accounts, not only for the preferred solution to the etymological problem at hand, but also for the relative inadequacy or outright impossibility of all the alternative explanations. Liberman's is a thoroughly satisfying display of etymological method and, not incidentally, a compelling solution to an unsolved etymological crux. Anyone who doubts the value of *festschriften* should consider that Liberman includes citations from 513 of them in his *A Bibliography of English Etymology* (2010) — they prove more useful than many scholars, librarians, and publishers imagine.

Samuli Kaislaniemi's "Early East India Company merchants and a rare word for sex" is at least as interesting as Liberman's account of *yeoman*, though its conclusions are not quite as sound. The rare word in question in *lapidable*, defined in Early Modern dictionaries (mono- and bilingual) as 'stonable [< L *lapid-* 'stone']' and 'marriageable'; the *OED* proposes the first definition and calls evidence of the second "a strange mistake ... copied in some later Dicts." With the benefit of newly available digital resources, Kaislaniemi provides contextual evidence for the 'marriageable' meaning and goes further to show, from both dictionary and contextual sources, that *marriageable* was code for 'sexually desirable, available.' All of this is a wonderful service to lexicography and cultural history, admirably executed. The stones in question, Kaislaniemi proposes with good reason, are those of the male — a *lapidable* woman is 'fuckable' or so desirable that one "gets one's rocks off." But the argument is not quite scorched-earth. Kaislaniemi quotes from *Jemmy Carson's Collections* (1744) as follows: "but if thee pursues it farther, to know whether she be *Lapidable*, or not, thee art certainly a Tyrant: For the *Hammer of thy Loins*, will at length beat down the *Fortress* of her *Porto Bello*; and the *Pillars of her Tabernace*, will be spread abroad, until thee has plundered the City, and taken the *Precious Stones* away." The alternative to the male stones is the

female *lap* 'pudendum' (*OED* sv **lap** *n*¹ in sense 2b), which Kaislaniemi dismisses, but without, tyrant-like, taking the etymological hammer to the pillars of this very evidence and hauling the stones away, as it were. Good as the argument is, the etymology remains unsettled.

Cynthia Lloyd's article, "From denominal to deverbal: Action nouns in the English suffix *-al*," by way of extending the metaphor, leaves no stone unturned. In this, however, it's no different from the others, but just as excellent. It provides a useful typology of *-al* suffixations and a splendid diachronic account of the suffix's semantics, fortified with persuasive contextual examples culled from the *OED* and the *Middle English Dictionary*. McConchie is a similarly scrupulous investigator of affixes (many of which are strongly associated with medical vocabulary), and it was doubtless written to reflect that shared interest. I was particularly grateful to read the beginning of Lloyd's article, at a point when I thought the volume was drifting among words without much purpose: "This book," she writes, "includes papers on both Old English and Latinate Renaissance lexis in English (Liberman and Diller respectively). It also contains studies of the transition from Latin to Old English (Hall), and of the subsequent revival of interest in Old English during the Latinised English Renaissance (Karlas-Tarkka). Between these two points, the OE vernacular became reunited with Latinate culture and lexis through the medium of the conquering language, French." This is exactly the point from which Lloyd's argument embarks, all of the contributions besides Liberman's are in front of the reader, and their relations to one another helpfully anticipated.

Alaric Hall's "*A gente Anglorum appellatur*: The evidence of Bede's *Historia ecclesiastica gentis Anglorum* for the replacement of Roman names by English ones during the early Anglo-Saxon period" immediately follows Lloyd's contribution. Hall argues that the inherent instability of place names accounts for the gradual shift from Roman to English place-names, leaving Celtic names behind in spite of considerable demographic continuity, and that Bede's *Historia* provides indirect evidence to support that model. It is a subtle and learned argument and will undoubtedly lead to future research on the issue. Leena Kahlas-Tarkka's "William Lambarde and Thomas Milles in search of the golden past," is one of the most elegant accounts I have ever read of Early Modern English antiquarian interest in Anglo-Saxon language and culture as a means of throwing off the

KERNERMAN NORSKE ORDBØKER

Editors:

Norwegian: Carina Nilstun and Sarasvati Jacobsen, led by Ruth Vatvedt Fjeld

Norwegian to French:

Clémentine Choubrac, Isabelle Ljøsne and Sabine Burlot

French: Nathalie Lanckriet (chief editor), Marine Lesprit, Valérie Lecœur, Frédérique Melanie-Becquet, Emilienne Morrucci and Caroline Moulart-Fremaux

French to Norwegian:

Torill Simonsen

Norwegian to German:

Nora Pröfrock

German:

Susanne Billes (chief editor), Monika Böhme, Sabine Corso, Susanne Dyka, Marion Hahn, Christine Moehrs and Emmanuelle Worbs

German to Norwegian:

Per Einar Fosser

Norwegian to Italian:

Berit N. Johnsen

Italian:

Palma Gallana

Italian to Norwegian:

Christian Skaug

Norwegian to Polish:

Dawid Gut

Polish: Iwona Sosinska, Aleksandra Marciniak, Alicja Huflejt-Hamon and Dawid Gut

Polish to Norwegian:

Dawid Gut, revised by Nanna Natalia Karpinska Dam Jorgensen

Norwegian to Spanish:

María Carmen Lagos Andersen

Spanish:

Miguel Eduardo Montoro and Irene Renau Ararque (chief editors), Sergio Mario López, María Verónica Elizondo and Carlos Sebastián Cerón

Spanish to Norwegian:

Alfhild Elisabeth Hirsch Kopperdal

Series of Norwegian bilingual learners' dictionaries published by Vega Forlag. See back cover.

so-called Norman Yoke and establishing ancient national identity. It, too, is subtle work, and it is an excellent model for young scholars.

The penultimate contribution is Hans-Jürgen Diller's "Contempt — The main growth area in the Elizabethan emotion lexicon." Diller acknowledges that "Contempt is *not* a nice topic for a *Festschrift*," but his article about the lexical field "Contempt" is a generous gift to McConchie and to other readers, as well. Diller takes material on "Contempt" and "Disrepute" from the database underlying the *Historical Thesaurus of the Oxford English Dictionary* (Kay, Roberts, Samuels, and Wotherspoon 2009) and examines it rigorously from literally every direction with vertical bar graphs and horizontal line graphs that contrast features, field size, growth of the field relative to the whole lexicon of Emotion terms, and much more — it is a tour de force, very demanding of readers, and, as such, I think the editors realized, probably not the best piece on which to end the volume. Cleverly, they end instead with Joshua Pendragon and Maggie Scott's "A lexical skirmish: *OED3* and the vocabulary of swordplay," which is itself not light fare, but is appealingly written and, given McConchie's published interest in the subject, a very palpable hit.

Just as we must admire the editors' astute arrangement of the various contributions, we must also praise the care authors and editors have taken preparing the text for publication. Of course, innocuous errors occasionally survive even the most diligent proofreading. Some errors, however, are potentially more confusing, even to specialists. So, when Liberman writes "The *ODDE* is entirely dogmatic," *ODDE* may represent a text accidentally omitted from his references, but probably is meant to be *ODEE*, the *Oxford Dictionary of English Etymology* (Onions 1966), which one does find among them (and is, indeed, entirely dogmatic). The unwary (or less phonologically minded) reader may have more trouble with "A few early forms of yeoman listed in the *OED* are spelled with *-mn-*, and they presuppose assimilation from **-n* (<*ng*) + *m-*. However, variants with *-mm-* are in the minority and can be explained in more ways than one (for example, by the analogical shortening of the root vowel in other words ending in *-man* or by the erratic habits of Middle English scribes)." The scribe would be erratic indeed who wrote *-mn-* rather than *-nm-*, for only the latter cluster would be evidence of the assimilation required to get from *yongman* to *yeoman*. It seems unfortunate, too, that the title of McConchie's great work, *Lexicography and Physicke* (1997) has been

truncated accidentally to merely *Physicke* in David E. Vancil's amiable and informative preface to the *festschrift*, a brief account of McConchie's career and interests that helps to justify the volume's range of subjects. These are all small matters, but the last, at least, is perhaps not the best sort of error to make in a *festschrift*.

The first responsibility of *Words in Dictionaries and History* is to honor R. W. McConchie, which it does by the uniform excellence of the articles included in it, the way those articles respond to McConchie's varied interests, and their frequent citation of McConchie's works, which merely underscores the significance of his work in the history of English and English lexicography. The same excellence appeals to its fortunate readers, and I hope the next *festschrift* I pick up is half as good as this one.

References

- Coleman, J. and Ogilvie, S. 2009.** Forensic Dictionary Analysis: Principles and Practice. *International Journal of Lexicography*, 22.1: 1-22.
- Fraser, R.A. 1952.** An Amateur Elizabethan Composer. *Music and Letters*, 33.4, 329-332.
- Fraser, R.A. 1955.** *The Court of Venus*. Durham: Duke University Press.
- Kay, C., Roberts, J., Samuels, M. and Wotherspoon, I. 2009.** *Historical Thesaurus of the Oxford English Dictionary*. 2 volumes. Oxford: Oxford University Press.
- Liberman, A. 2008.** *An Analytical Dictionary of English Etymology. An Introduction*. Minneapolis: University of Minnesota Press.
- Liberman, A. 2010.** *A Bibliography of English Etymology. Sources and Word-List*. Minneapolis: University of Minnesota Press.
- McConchie, R.W. 1997.** *Lexicography and Physicke: The Record of Sixteenth-Century English Medical Terminology*. Oxford: Clarendon Press.
- McConchie, R.W. 2007.** The Real Richard Howlet. In Considine, J. and Iamartino, G. (eds.), *Words and Dictionaries from the British Isles in Historical Perspective*. Newcastle upon Tyne: Cambridge Scholars Publishing, 39-49.
- Onions, C.T., with the assistance of Friedrichsen, G.W.S. and Burchfield, R.W. 1966.** *The Oxford Dictionary of English Etymology*. London: Oxford University Press.

Michael Adams

Indiana University, Bloomington
adamsmp@indiana.edu

Grande Dicionário Houaiss da Língua Portuguesa

Mauro de Salles Villar

In the 1980s, Brazilian diplomat and linguist Antônio Houaiss, and a small group that I was part of, started to put together a project for a general-purpose dictionary that would differ from all others available in Portuguese.

Four fundamentals were established for the task to be undertaken:

1. The screening, by using various gathering techniques (from text corpora, to literature, to reading scientific books, to dictionary research, etc.), of the most comprehensive nomenclature consistent with the scope of a general-purpose dictionary that would also incorporate Portuguese linguistic performances outside the Portugal/Brazil axis.
2. A thorough inventory and analysis of Portuguese morphemes as a base for establishing large families of words with common ancient etymons.
3. A full-fledged effort to date lexical units to act as illustrative quotations, and to convey reliability in terms of the way that senses appear arranged within the entries.
4. An in-depth etymological study of each unit to anchor the entries' definitions.

The first edition of *Dicionário Houaiss da Língua Portuguesa*, now called *Grande Dicionário Houaiss*, resulted from an endeavor that started in February 1986 and ended in December 2000, and which involved the work of 35 lexicographers and 43 specialized external collaborators, as well as contributions from Portugal, São Tomé and Príncipe, Guinea-Bissau, Cape Verde, Angola and Mozambique.

The outcome was a dictionary with approximately 228,500 lexical units. The work is not a synchronic view of any given chronological or geographical segment of the language. Rather, from a historical perspective, it focuses on phenomena that mostly concern modern-day usage of the language in Brazil and Portugal. But it also contains a selection of words from both old and archaic Portuguese whose inclusion is warranted, for instance, by their rate of occurrence in the literary history.

Dicionário Houaiss was published in 2001, fifteen years after it was begun in 1986. It was first issued in print, later becoming available both on CD-ROM and online. A European version complying with the orthography and the linguistic norm

then officially adopted in Portugal came out in 2002.

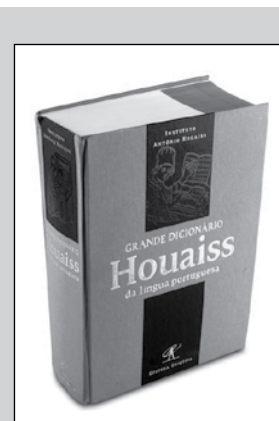
The second edition has just been completed. It is the outcome of another 10 years of revising and adding new materials to the first set. As compared to the first edition, hundreds of thousands of amendments were incorporated into every feature of the work — from definitions to phonetic notations, from dating references to the list of sources, from revised etymologies to the inclusion of morpheme-headwords. In addition, the lexicon was also revised, as were the specific descriptions concerning scientific and technical subjects such as biology, ecology, physics, astronomy, computer science, zoology, botany, and so forth.

This second edition features over 234,000 entries, of which 194,936 are words, word combinations, or morphemes; 29,393 are nestings; 4,861 are plurals carrying their own distinct meanings; 210 are phrases subsumed into the entries' microstructures; and 4,889 are words used on maps.

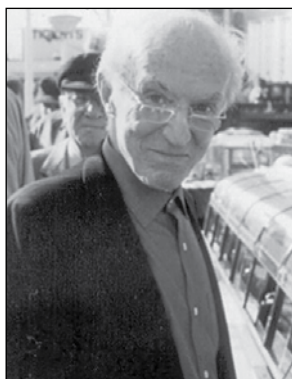
The first edition of our dictionary enjoyed an excellent response among common readers and specialized critics alike. This time, we are set on providing an even more rigorous, pondered, updated, and improved version of what had previously been accomplished.

Instituto Antônio Houaiss and K Dictionaries have launched the new joint brand name **HOUAISS KERNERMAN** for Portuguese bilingual dictionaries.

Mauro de Salles Villar was born in Rio de Janeiro in 1939 and is a lexicographer, encyclopedist and lawyer. He is the Chief Editor of *Grande Dicionário Houaiss da Língua Portuguesa*, co-director of Instituto Antônio Houaiss of lexicography, and member of the Brazilian Academy of Philology. maurovillar@iah.com.br



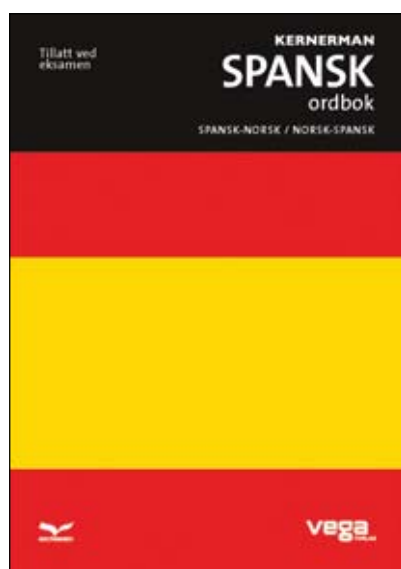
Grande Dicionário Houaiss da Língua Portuguesa
 Antônio Houaiss and Mauro de Salles Villar
 2001. Rio de Janeiro:
 Editora Objetiva
 Hardcover, lxxxiii + 2,925
 pages, 305 x 230 mm
 ISBN 85-7302-383-X
www.objetiva.com.br



Antônio Houaiss, 1915-1999, was a Brazilian lexicographer, writer and translator, who served as the Minister of Culture, and was a member of the Brazilian Academy of Letters. He was the son of Lebanese immigrants, and was born and died in Rio de Janeiro. He is best known for his translation of James Joyce's *Ulysses*, and for supervising the compilation of the *Dicionário Houaiss da Língua Portuguesa*.

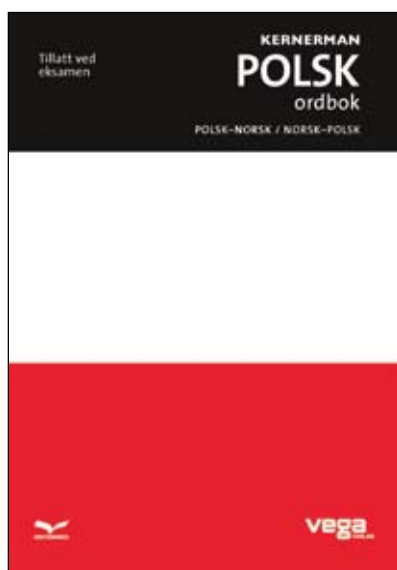
Kernerman Norske Ordbøker

Vega Forlag • Oslo • June 2012



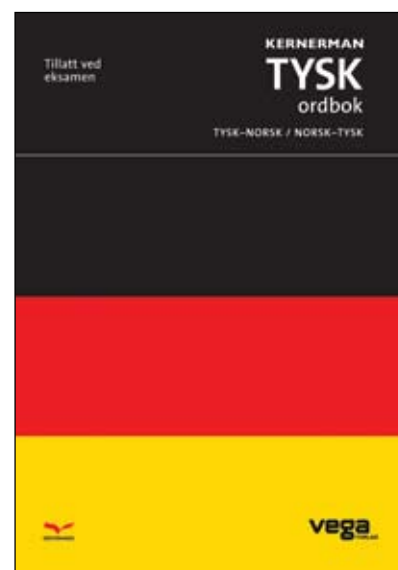
KERNERMAN SPANSK ORDBOK

Spansk-Norsk / Norsk-Spansk
936 pages, 210 x 145 mm
Paperback + CD-ROM
ISBN: 978-82-8211-022-8
Reprint



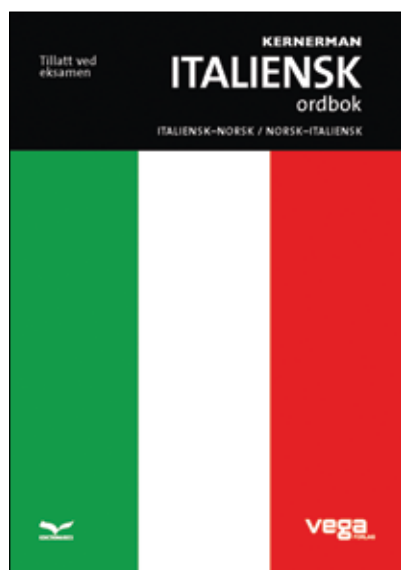
KERNERMAN POLSK ORDBOK

Polsk-Norsk / Norsk-Polsk
751 pages, 210 x 145 mm
Paperback + CD-ROM
ISBN: 978-82-8211-277-2



KERNERMAN TYSK ORDBOK

Tysk-Norsk / Norsk-Tysk
1103 pages, 210 x 145 mm
Paperback + CD-ROM
ISBN: 978-82-8211-135-5



KERNERMAN ITALIENSK ORDBOK

Italiensk-Norsk / Norsk-Italiensk
1253 pages, 210 x 145 mm
Paperback + CD-ROM
ISBN: 978-82-8211-133-1
Reprint



ENGELSK SKOLEORDBOK

Engelsk-Norsk / Norsk-Engelsk
704 pages, 210 x 145 mm
Paperback + CD-ROM
ISBN: 978-82-92489-78-9
Reprint



KERNERMAN FRANSK ORDBOK

Fransk-Norsk / Norsk-Fransk
TBA pages, 210 x 145 mm
Paperback + CD-ROM
ISBN: 978-82-8211-134-8

Vega Forlag AS

Gjøteborggata 6a, N-0566 Oslo

Telefon +47 21 09 04 10, Telefax +47 22 37 15 50, post@vegaforlag.no, [www.vegaforlag.no](http://vegaforlag.no)