

# DICTIONARY News

## Integrating phonetic transcription in a Brazilian Portuguese dictionary

Luiz Carlos Cagliari

### 1. The dictionary

K Dictionaries has developed a series of dictionaries for learners of various languages, including Portuguese/French. This dictionary was based on European Portuguese, whose words had a phonetic transcription with European pronunciation. A group of lexicographers from the University of São Paulo, under the supervision of Ieda Maria Alves, adapted the Portuguese entries to Brazilian vocabulary, and another team from UNESP (São Paulo State University) at São José do Rio Preto, under the supervision of Claudia Xatara, adapted to Brazilian the Portuguese translations of the French entries.

In line with the change from European to Brazilian Portuguese, I gathered a group of students to be involved in modifying the phonetic transcription. The collaborators were Patrícia Ormastroni Iagallo, Bruno Sérgio Sedenho, Priscila Maria Mendonça Machado, Cristiane Jussara Romanatto, Caroline Talge Arantes, Eliane de Oliveira Galastri and Lara Vieira Balieiro. The first three were working on their MA degrees and the other four were undergraduate students at UNESP / Campus of Araraquara.

I started learning how to operate the XML program that was used for this work and discussing all kinds of problems related to the phonetic transcription.

### 2. Problems with phonetic transcription

I have been working as a phonetician for the past three decades and during this period I have carried out many projects, exploring different areas of articulatory, acoustic and experimental phonetics. The task of doing phonetic transcription was present

in everyday work. It is always common for a phonetician to transcribe his or her own language. When doing that, all kinds of sound variation are registered, according to the speakers' pronunciation. On the other hand, when doing phonology, the sound patterns are interpreted following theories in order to get an abstract sound system of the language. In this case, the phonetic variation is accommodated into their phonemes. The phonological transcription does not represent a pronunciation of the language, in the same sense as phonetic transcription does. Moreover, phonological transcription does not relate to language in the same way as orthography does. According to the rules of writing systems, orthography has the function of neutralizing the phonetic variation of a language on the word pronunciation level, allowing all speakers to read. Consequently people read in their own dialectal variety or, if they wish, using another variety, for instance, a more culturally or educated pronunciation.

Of course, a dictionary presents the orthographic form for words. If it is decided to implement phonetic transcription in order to help users with the pronunciation, chiefly when the dictionary is bilingual, then nothing is more appropriate than asking a phonetician to do the job. However, not all phoneticians know exactly what is demanded and expected in such a case. Certainly, a phonological transcription is inadequate. For the reader, spelling works better than the phonological transcription in a language like Portuguese. On the other hand, which variety should represent the language in the dictionary? The option for educated pronunciation is a good choice, but it is not clearly defined.

1 Integrating phonetic transcription in a Brazilian Portuguese dictionary | **Luiz Carlos Cagliari**

6 Kernerman French-Portuguese dictionary: Adapting the translation from European Portuguese to Brazilian Portuguese | **Helena Sakano Fernandes and Claudia Xatara**

10 Kernerman Dutch Dictionary | **Truus Kruyt**

12 Anthropological and linguistic fundamentals of lexicographic work | **Miguel Eduardo Montoro**

19 Pedro A. Fuertes-Olivera (ed.). *Specialised Dictionaries for Learners* | **Amélie Josselin-Leray**

26 Gilles-Maurice de Schryver (ed.). *A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks* | **Elizabeth Knowles**

28 Henri Béjoint. *The Lexicography of English. From Origins to Present* | **Anatoly Liberman**

31 Invitation to ASIALEX 2011 in Kyoto | **Ai Inoue**

32 In Memorium: Shigeru Takebayashi, 1926-2011 | **Hiroko Saito**

Editor | **Ilan Kernerman**



K DICTIONARIES

© 2011 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



**Luiz Carlos Cagliari** received his MA in Linguistics from the State University of Campinas (UNICAMP) in 1974 and his PhD from Edinburgh University in Scotland in 1978. He started the Phonetic Laboratory in the Department of Linguistics at UNICAMP, where he taught phonetics and phonology until his retirement in 1997. In 1991 he got his professorship in Phonetics and Phonology (UNICAMP). He carried out research on writing systems at the School of Oriental and African Studies, London, in 1988 and investigated prosody as Senior Researcher at Oxford University in 2004. He has been teaching and supervising MA and PhD students at São Paulo State University (UNESP) since 2005. His main research interests include phonetics, phonology, writing systems, orthography and literacy, and he has authored books and numerous publications on these topics. In recent years his research has focused on writing systems and prosody with particular reference to intonation and rhythm. Professor Cagliari is responsible for the phonetic transcription of Brazilian Portuguese in the global series of K Dictionaries. [lccagliari@gmail.com](mailto:lccagliari@gmail.com)

Since the definition of the dialect was crucial, I started to look at dictionaries to find out what they have done in such regard. The dictionaries intended for native speakers of Portuguese do not usually include phonetic transcription. This is a pity. Only some orthoepic information would appear close to the headwords when it is absolutely necessary. With the recent spelling reform of the Portuguese language<sup>1</sup>, words like *econômico* indicate Brazilian pronunciation and *económico* European Portuguese. But, this is not enough if we take into account the difficulties users face, searching for the correct pronunciation. The problem intensifies when the user is not a native speaker of Portuguese.

In the *Aurélio Século XXI* dictionary (1999), in the section *Como Usar o Aurélio XXI (How to use the Dictionary)*, we find:

“5. A ortoépia entre parênteses e em tipo claro, após a cabeça, esclarece a pronúncia, quando passível de dúvida (vogal aberta ou fechada, pronúncia de consoante ou de encontro consonantal, hiato, etc.” (p. xv).<sup>2</sup>

“6. A transcrição fonética dos estrangeirismos vem logo após a cabeça, entre colchetes, e representa, no Alfabeto Fonético Internacional (adaptado) a pronúncia aproximada da palavra na língua de origem. A tabela de transcrições fonéticas encontra-se na página XVI.” (p. xv).<sup>3</sup>

That dictionary is useless with regard to phonetic transcription. It annotates irrelevant cases and does not indicate the correct pronunciation when necessary.

The *Dicionário Houaiss da língua portuguesa* (2004) presents some details about orthoepy and phonetic transcription (p. XIX). Like Aurélio, Houaiss uses the orthoepic information for Portuguese words and presents phonetic transcription only for foreign items. Houaiss employs the International Phonetic Alphabet (IPA) notation carefully in the phonetic transcription, but in the orthoepic indication, the spelling notation is an option (using diacritics). The transcription comes within // bars, which is misleading, since this symbol is properly used only for phonological transcription. Although this dictionary is more comprehensive in global terms, it does not really help readers with the pronunciation of the language.

Other dictionaries list the phonetic symbols in use with examples. In general, they are confused, incomplete and not rarely mistaken. (cf. *Dicionário Unesp do Português Contemporâneo*, 2004). This approach has the shortcoming of being nobody’s pronunciation. It pre-supposes the reader has a proficiency in the omitted cases.

We come across more detailed phonetic transcriptions in bilingual dictionaries. Because they transcribe all items, they must set up the variety of the language that is being used. Even in these cases, it is not always clear which is the chosen variety, because it is not indicated, and the transcriptions we see are not strictly close to one dialect. An additional problem is the use of phonetic symbols. In spite of statements indicating the use of the IPA, in practice, what we find is something adapted in a way that makes it painful to consult the dictionary.

In the *Novo Dicionário Folha Webster’s Inglês/Português, Português / Inglês* compiled by Antônio Houaiss and Ismael Cardim (1996), all English and Portuguese items have a phonetic transcription within | | bars (which are traditionally used for morphological transcription, not for phonetics). The dictionary begins with a “phonemic transcription” that is, in fact, a phonetic one. After presenting all vowels, consonants and diphthongs, there are some notes relating to spelling forms for pronunciation in the case of Portuguese. Nothing of the kind is said for English. It says:

“|ʒ| is a transcription for written *s*, *z* or *x* ending syllables, e.g. *as*, *estas*, *postura*, *atriz*, *exportar*, *expansão*. The pronunciation of |ʒ| has two patterns, one we should call old and Paulista pattern, and the other, Rio-Lisboa pattern.” (p. vii). (Examples: Old-Paulista pattern: *velhas santas* |vɛlas sãtas|; *fortes garras* |fɔrtɪz |gʌras|. Rio-Lisboa pattern: *suas tias*: |suax |tiax|; *fortes garras* |fɔrtɪj |gʌraxl.)

These examples show how confusing and inaccurate those transcriptions are.

In this dictionary, the phonetic transcription follows the Rio-Lisboa pattern, but the model has unacceptable transcriptions like the use of the same [ʀ] for Paulista |fɔrtɪzl and Rio |fɔrtɪj|. Obviously they do not have the same pronunciation. In Paulista we have a voiced tap (or a retroflexed “r”), and in Rio a voiceless velar fricative: *mar* |mʌx|, *porta* |pɔxta|.

The *Michaelis Moderno Dicionário Inglês/Português Português/Inglês* (2000) presents the words, their syllable boundaries for writing and the phonetic transcription between brackets. As usual, only the primary stress is indicated, but in this case, the stress mark is the sign [ˈ] placed just before the stressed vowel (*casa* [kˈazə]), which is very unusual.

“Foreign words were transcribed by using the phonetic scheme presented above, adapting the pronunciation as closely as possible to the foreign model (*hovercraft* [rˈɔvɛkrˈaft])” (p. xxi). There

is an advertisement about variation on different dialects.

“A word is rarely pronounced in the same way everywhere in Brazil. There are many variations, influenced mainly by the region where the language is spoken and by factors such as the speaker’s educational level, social group and age bracket. For this reason, a choice had to be made regarding pronunciation. To date, there is no official standard for spoken Portuguese, as is the case for the vernacular French, for example. Therefore, the formal pronunciation of speakers with higher education in the city of São Paulo, and that used by radio and television announcers on national network, was adopted. Their pronunciation tends to undergo minimal influence from regional variations for this dictionary.” (p. xx).

The dictionary uses the symbol [ə] and notes:

“Similar to the **a** in **about**. It corresponds to the final, unstressed open [a] pronounced (in fact, like all final unstressed vowels) faintly, almost inaudibly; in practice, it can be replaced by [a] without prejudice to communication.” (cf. example: *janela*, p. xx).

There is also the symbol [ʌ] with the following observation:

“Similar to the posterior, semi-open **a**, as in **bud**, and slightly nasalized; it is always stressed and followed by *m*, *n*, *nh*. (example: *cama*, *cana*, *cânhamo*)” (p. xx).

Words like *canto* and *órfão* are transcribed as [k'ântu], [ˈorfãw]. Therefore, there are two ways to transcribe the same sound [a] = [a] and [ə]; and two ways to transcribe the nasalized “a”: [ã] and [ʌ]. This kind of alternation is confusing for the reader and unnecessary. In spite of being published in 2000 (revised edition), the sound of “R”, transcribed as [r̄] (r with micron) (cf. *rato*, *carro*) is defined as a trill [r], a pronunciation that is very rare in the Paulista dialect today.

In the *Dicionário Larousse / Ática Básico Inglês-Português, Português-Inglês*, we see a systematic transcription based on the Carioca dialect (from Rio de Janeiro). It also marks the secondary stress. In practice, the dictionary avoids some vocalic reductions to approximate the phonetic transcriptions to spelling forms: *escola* [is'kɔla] but *embaixo* [ẽm'bajʃu]. The word *bem* is noted [ˈbẽj] but *bom* is transcribed [ˈbõ], which is nobody’s pronunciation (cf. [ˈbõõ]). Occasionally there are dialectal pronunciations like *boné* [bõ'ne] (cf. [bo'ne]). It transcribes *capaz* [ka'paz], not following a rule.

In general, the transcription characterizes the Carioca dialect but it does not conform strictly to the sounds, sometimes providing a more phonological representation of the sound pattern.

This dictionary tour has a purpose: which model should I use to make my phonetic transcription for the new Kernerman dictionary? I finished my phonetic journey with lots of doubts and no solution. Thus, I decided to apply my own model (Cagliari, 2007).

### 3. The model

- 3.1 The option for the International Phonetic Alphabet was unquestionable. However it should be used as settled by its conventions without adaptation. It means that any linguist or person with the knowledge of the IPA should get the right pronunciation.
- 3.2 For computer reasons, the IPA font chosen was Arial Unicode MS.
- 3.3 Since there was no real speaker to be transcribed, but a variety of the language, and taking into account the fact that the educated variety of the Brazilian pronunciation is less sensitive in the whole country than the Paulista dialect, the choice to transcribe that variety was considered to be the most convenient. The most important television news announcers are trained to use what is referred to as less socially marked pronunciation. In practice, they use a modern variety of the Paulista dialect. Other varieties appear on different TV programs and in regional news bulletins. However, there are adaptations, for instance, avoiding the retroflexed sounds of any kind, introducing the modern affricates for coronal stops, now disseminated among the new generations of Paulista speakers. That strategy seemed interesting and I decided to follow the TV news announcers’ pronunciation.
- 3.4 A group of students, Paulista dialect speakers, helped me with the phonetic transcription. But there still remained the most important and crucial problem to be solved: as it does in all varieties of a language, the Paulista dialect presents different pronunciations in particular contexts or even for a single word. That required the statement of some conventions that should not betray a good pronunciation, but on the other hand should get some generalization, convenient for a dictionary. This particular problem was related to the transcription of nasality being consonantal or vocalic. So, the next step demanded setting up some transcription rules.

K Dictionaries has been involved in creating a new global dictionary series since 2006. The series consists of lexicographic cores that serve as bases for developing monolingual, bilingual and multilingual dictionaries, for different target audiences, in print and digital media. The following language cores are available as of mid-2011:

- Arabic
- Chinese Simplified
- Chinese Traditional
- Czech
- Dutch
- English
- French
- German
- Greek
- Italian
- Japanese
- Korean
- Latin
- Norwegian
- Polish
- Portuguese Brazil
- Portuguese Portugal
- Russian
- Spanish
- Thai
- Turkish

The following language cores are in preparation and due in 2012:

- Hebrew
- Hindi
- Swedish



#### Kernerman French-

Portuguese Dictionary was developed as part of the series Dictionnaires Assimil Kernerman, launched by Assimil in Paris in 2009 (cf. KDN 17). It is fully bilingual, and bi-directional, and is targeted at French-speaking learners of European Portuguese (EP).

The French dictionary core was created by Nathalie Lanckriet (editor-in-chief), Valérie Lecœur, Marine Lesprit, Frédérique Melanie-Becquet, Emilienne Morruzzi and Caroline Moulart-Fremaux (reviewed by Colette Alouch).

The EP dictionary core was created by Lúcia Liba Mucnik (reviewed by Ana Margarida de Almeida Borges), who also did the translations (Portuguese-French translation revised by Marie-Pierre Gachet). The EP translation of the French-Portuguese dictionary was adapted to Brazilian Portuguese (BP) by Claudia Maria Xatara and Maria Cristina Parreira da Silva, along with Helena Yuriko Sakano Fernandes, Noelle Perez, Catharine Kamie Ramos Furukawa, Cristiane Jussara Romanatto and Bruno Sérgio Sedenho.



3.5 One more problem came up: the phonetic transcription needed to be as close to the spelling form of the word as possible. Portuguese orthography is friendly in this respect. In fact, this decision worked as a filter and a target when deciding between competitive pronunciations.

#### 4. Some transcription rules

4.1 To avoid retroflex sounds, the “R” was transcribed with the symbols [x]: *rato* [ˈxatu], *carro* [ˈkaxu], *honra* [ˈõxa] and [r] *caro* [ˈkaru], *porta* [ˈpõta], *mar* [ˈmar], *braço* [ˈbrasu]. The voiced [ʁ] was discarded since it varies with [x]: *barriga* [baˈʁiga] ~ [baˈxiga].

4.2 As pointed out above, the coronal stops were transcribed with the affricates [tʃ] and [dʒ] when appearing before an anterior closed vowel [i]: *tia* [ˈtʃia], *dia* [ˈdʒia]; or [ɾ]: *pote* [ˈpõtʃɾ], *sede* [ˈsedʒɾ], *desrespeito* [dʒɪsɪksɐˈpɛɾtu].

4.3 In word final position, the unstressed vowels were transcribed as [a, ɪ, u]. The option for [ə, e] instead of [a] runs only for European Portuguese.

4.4 The vowels [ɪ, u] next to another vowel mean the weak part of a diphthong except when a stress mark indicates as hiatus or in unstressed endings when [ɪ] or [u] precede another vowel. In this case, there is a variation between the hiatus and a diphthongized pronunciation: *funcionar* [fũõsɪoˈnar], *mágoa* [ˈmaɣua], *ciência* [sɪˈẽĩsɪa].

4.5 In word initial position, if the syllable is unstressed, there occurs the vowel [ɪ] when followed by a coronal fricative in coda position: *escola* [ɪsˈkõla], *externo* [ɪsˈtɛɾnu]. This represents a clear tendency in the dialect.

4.6 There is a tendency to have *des-* pronounced as [dʒɪs/z...]: *desbloquear* [dʒɪzblokɐˈar], *descobrir* [dʒɪskoˈbrɪr].

4.7 A more complicated case involving unstressed syllables in word initial position occurs with the presence of nasality. If the word starts with the vowel, the best choice was the diphthong [ɛɪ + (nasal)]: *então* [ẽĩnˈtẽũ], *enraivecido* [ẽĩxarveˈsidu]. If it starts with a consonant, the pronunciation with a diphthong used to sound emphatic, mainly when followed by a stop consonant: *tentativa* [tentaˈtʃiva], *sentença* [senˈtẽĩsa]. The transcription marks the nasal consonant before a stop; otherwise the diphthong is the rule.

4.8 As a general rule, the nasalized vowels and diphthongs receive the tilde only when followed by a nonstop consonant. Before stop consonants, there is the transcription of a homorganic

nasal consonant: *cansado* [kẽˈsadu], *pensar* [pẽĩˈsar], *enlatado* [ẽĩlaˈtadu]; *campo* [ˈkẽmpu], *tinta* [ˈtʃinta], *trancar* [trẽŋˈkar]. The pronunciation without the nasal diphthong before fricatives and laterals is not encountered in the Paulista dialect.

4.9 The nasalization of the vowel “A” is always marked as [ɐ] before stop and nasal consonants and as [ẽ] before nonstop consonants: *bando* [ˈbẽndu], *frango* [ˈfrẽŋgu], *cansar* [kẽˈsar], *lanchar* [lẽˈʃar], *cama* [ˈkẽma], *banha* [ˈbẽna]. In Portuguese, there does not occur an homorganic nasal before nonstop consonants, but only the nasalized vowel or diphthong. One exception is the transcription with [a] instead of [ɐ] in unstressed syllables when the vowel precedes a nasal consonant at the onset of the following syllable: *banana* [baˈnẽna], *camada* [kaˈmada]. Phonetically, the two pronunciations occur (cf. [bɛˈnẽna], [kẽˈmada]).

4.10 At word endings, there occurs the nasalized vowel [ẽ]: *maçã* [maˈsẽ], *irmã* [iˈrẽmẽ] and nasalized diphthongs [ẽĩ], [ẽũ], [õĩ], [õũ]: *mães* [ˈmẽĩs], *pão* [ˈpẽũ], *corações* [koraˈsõĩs], *marrom* [maˈxõũ]. The spelling “-IM” and “-EM” were transcribed as [ɪŋ] and [ẽŋ]: *fim* [ˈfɪŋ], *assim* [aˈsɪŋ], *vem* [ˈvẽŋ], *armazém* [armaˈzẽŋ]. In these cases, the transcription with the nasalized diphthong [ĩ] and [ũ] would be a good one, but transcriptions like *fim* [fɪ], *vem* [vẽ] and *bom* [bõ] are misleading.

4.11 Close to the spelling form, inside the words, the phonetic transcription maintains the unstressed vocalic quality [i, u] for “I” and “U”. For spelling “E” and “O”, the transcription varies according to the pronunciation: [e] ~ [ɪ], [o] ~ [u].

4.12 The correct pronunciation of the diphthong is [ɔu] for the spelling “UL”: *sul* [ˈsɔu], *último* [ˈuʉtɪmu], *culpa* [ˈkɔupa]. The transcription [uɔ] is misleading. The right pronunciation is [au] for “AL” and “AU”: *alta* [ˈaʉta], *auto* [ˈaʉtu], *almoço* [aʉˈmosu].

4.13 Although we find in Paulista dialect diphthong reductions or diphthongization in some contexts, as used to be in Brazilian Portuguese, I have decided not to implement these rules, because without them, the pronunciation sounds more educated. Examples: *caisa* [ˈkaɪsa], *cadeira* [kaˈdeɪra], *pouco* [ˈpɔuɔ], *rapaz* [xaˈpas], *frequês* [freˈges], *nós* [ˈnɔs].

4.14 On the other hand, an epenthetic vowel [ɪ] is inserted to keep the syllable pattern correct: *objeto*

[obr'zɛtu], *advogado* [aðʒivo'gadʊ],  
*optar* [opr'tar].

- 4.15 Foreign words were transcribed as the Paulista speaker pronounces them: *shopping* [ʃɔpɪŋ], *outdoor* [autʃr'dɔr].
- 4.17 Only the stressed syllables received a stress mark: *destruição* [dʒɪstrui'sɛũ], *psicologia* [psikolo'zia].
- 4.18 In a dictionary there sometimes happens to be two orthographic words together, constituting a phonological word. One of them is unstressed and may occur before or after the principal word. In such cases, the option was to fuse them in a single phonetic utterance: *em comum* [ɛɪŋko'mũ], *espreguiçar-se* [ɪspreʒi'sarsɪ]. If the words were individually stressed, they are presented separately: *dar para* ['dar 'para], *ir para os ares* ['ir 'paraʊ'zarsɪ]. The reason is to improve the knowledge of the occurrence of stress in the language and to facilitate the phonetic reading in the dictionary.
- 4.19 The reader may come across phonetic transcriptions of special cases for which a general rule could be applied, but find variations. When discussing these cases with the group of transcribers, it appeared that one pronunciation is preferable for certain words rather than another. Typically this happens at the beginning of words, involving the alternation between [e] and [ɪ]. For instance, the word *encaixar-se* could be transcribed as [ɪŋkaɪ'farsɪ] or [ɛɪŋkaɪ'farsɪ]; *desaparecer* [dezapare'ser] or [dʒɪzapare'ser]. The choice was the first occurrence. Obviously, we are facing a sound change in the language: not all /e/ have already transformed into [ɪ], and for this reason some rules do not apply.
- 4.20 These rules govern most cases that were optimized in order to keep a right and educated pronunciation, and at the same time this shows a compromise with the spelling form. These procedures seemed to be very useful taking into account the use of a dictionary for foreign and native speakers of Brazilian Portuguese. All phonetic transcriptions reveal an exact and accepted pronunciation as noted by IPA symbols. They transcribe the exact pronunciation of one educated variety of the Paulista dialect. It does not mean that in particular cases and contexts some speakers might not use an alternative pronunciation.

## 5. Conclusion

The most important thing when transcribing the phonetic transcription for the dictionary

was to avoid incongruence. The main aim of any writing system is to allow easy reading for all users. To achieve this goal all systems must have an orthography, whose most important function is to neutralize the phonetic variation in the writing system. Although these objectives must be present, because of the mobility of the sound pattern of languages, the system is always incomplete and restructuring itself. This is a problem for linguists who describe the system and for the orthography that is always out of synchrony. The phonetician feels comfortable in transcribing single speakers. The phonologist has some problems to accommodate the phonetic data into the phonological system. The orthographers are always delayed in order to save a tradition before the process of language changing. This makes the linguists' life somewhat complicated. But with a good understanding of the phenomenon of linguistic variation it is possible to describe the language in all its peculiarities without losing its systematic character. The experience of writing the phonetic transcription for the dictionary was a challenge to the group of people involved in it, and a good opportunity to learn more about the profession of lexicography.

## Notes

1. The Portuguese language has undergone many orthographic reforms over the past hundred years. The latest reform was officially implemented in Brazil in 2009. That was a minor reform affecting a few words in specific contexts (*heróico* = *heroico*; *idéia* = *ideia*) and introducing new rules for the use of hyphen (*mini-saia* = *minissaia*; *contra-senha* = *contrassenha*), etc. Portugal has not officially implemented this reform.
2. "5. The orthoepy between brackets and in a clear type, after the headword, interprets the pronunciation, when a doubt may occur (open or closed vowel, the right pronunciation of a consonant or of a consonant cluster, a hiatus, etc.)" (p. xv). Example: *resinoso* (ô). This kind of notation is useless, because the language morphology has clear and specific rules for such occurrences.
3. "6. The phonetic transcription of foreign words comes immediately after the headword, between brackets, and is represented by the International Phonetic Alphabet (adapted) in an approximate pronunciation of that word in its original language." (p. xv). Example: *res nullius* [ɾɛs 'nullɛos]; *outdoor* [ʔæwtɔɔ]. With the modifications, the dictionary presents a non-conventional use of the IPA. It



The phonetic transcription of the BP translations was adapted by Luiz Carlos Cagliariari along with Patrícia Ormastroni Iagallo, Bruno Sérgio Sedenho, Priscila Maria Mendonça Machado, Cristiane Jussara Romanatto, Caroline Talge Arantes, Eliane de Oliveira Galastri and Lara Vieira Balieiro.

The EP core was adapted to BP by Ieda Maria Alves, along with Ana Maria Ribeiro de Jesus and Luciana Pissolato de Oliveira, from the University of São Paulo.

The new BP dictionary core now serves as a base for developing a series of bilingual learners' dictionaries targeted at Brazilian users. As above, each title is bi-directional, i.e. including an L2-BP part as well as the BP-L2 part.

The BP dictionary was augmented together with Instituto Antônio Houaiss and is entitled *Dicionário Houaiss Kernerman da língua portuguesa para estudantes* (cf. p. 9).

means that the eccentricity in the phonetic transcription demands extra knowledge to grasp the right pronunciation.

#### References

- Cagliari, L.C. *Elementos de Fonética do Português Brasileiro*. São Paulo: Editora Paulistana. 2007.
- Borba, F.S. (ed.). *Dicionário UNESP do português contemporâneo*. São Paulo: Editora UNESP. 2004.
- Ferreira, A.B.H. *Novo Aurélio Século XXI: o dicionário da língua portuguesa*. Rio de Janeiro: Editora Nova Fronteira, 1999.
- Hunter, S.J. e McNeillie, J. (ed.). *Dicionário Larousse / Ática Básico Inglês-Português, Português-Inglês*. São Paulo: Editora Ática/Larousse. 2001.
- Houaiss, A. (ed.), e Cardin, I. (co-ed.). *Novo Dicionário Folha Webster's Inglês/Português Português-Inglês*. São Paulo: Editora Folha de São Paulo. 1996.
- Houaiss, A. e Villar, M. de Salles. *Dicionário Houaiss da Língua Portuguesa*. Rio de Janeiro: Editora Objetiva, 2001.
- Weiszflog, W. (ed.). *Michaelis: Moderno Dicionário Inglês-Português, Português-Inglês*. São Paulo: Companhia Melhoramentos, 2000.

## Kernerman French-Portuguese dictionary: Adapting the translation from European Portuguese to Brazilian Portuguese

Helena Sakano Fernandes and Claudia Xatara

This article describes the process of integrating translation equivalents for Brazilian Portuguese (BP) in French dictionary entries that were previously translated to European Portuguese (EP). The resulting dictionary is intended for adults and students who learn or use Portuguese on the intermediate level as well as for Brazilians learning French.

The dictionary is fully bilingual and bi-directional. Each section has about 12,000 entries in the source language with translations for every sense, example and phrase in the target language.

The French entries are part of the vernacular: names of animals, flowers, common trees, fairly common dishes and drinks, frequent interjections (*vive!* – *viva!*), foreign words (camping – *camping*) and colloquial words (*chouette* – *gracinha*; *legal*, *bacana*), not forgetting the most complicated terms found in the press or in administrative documents (*directeur adjoint* – *diretor adjunto*; *vice-diretor*), as well as acronyms and abbreviations.

Common expressions usually appear as part of the entry, either within a given sense or forming a sense of their own. The meanings that denote a change of grammatical category are treated as sub-entries, but homographs come in their own entries.

The dictionary includes a bilingual geographical list with names of countries (or major regions), their demononyms and corresponding languages.

The question could arise, why adapt a dictionary from Portuguese to Portuguese? When we think of Brazilian learners the answer is quite obvious. Regarding

the direction French-Portuguese, the Brazilian audience wishing to decode texts produced in French might find it difficult to understand certain equivalents in EP or might use unsuitable equivalents of the Brazilian variant.

However, compared to the French audience, the benefits of having two French-Portuguese dictionaries may not be very clear. At first, it might seem commercially more viable to add the equivalents unique to BP, simply by indicating “Br”, for example. Nevertheless, separating the two Portuguese variants into two dictionaries can considerably help learners, by avoiding a lexical mix and incorrect text production that is inappropriate for any of the variants. Wittmann et al. (1995) give good examples of this type of ungrammatical phrases, which they call “aberrations”:

(1) \**Quando mo deu, ele não tinha se apercebido...*

According to the authors, this sentence “is not correct Portuguese because it mixes two syntactic phenomena from different variants: *mo* [contraction of the pronouns *me + o*] is not used in Brazilian Portuguese whereas the order of the clitic pronoun, *não tinha se apercebido*, is not acceptable in European Portuguese.”<sup>1</sup> We could add to this that in Brazil we prefer the verb *perceber* rather than the pronominal verb *se aperceber*.

(2) \**Encontrei o banheiro no bonde.*

This sentence, on the other hand, cannot be accepted because of the words that are used. *Banheiro*, meaning *lifeguard*, is only used in EP (in Brazil *banheiro* means *bathroom*) and *bonde*, meaning *tramway*, is a word used only in Brazil (in Portugal, *eléctrico* is used).



The two variants differ on all linguistic levels: phonological, lexical, morphological and syntactic. But it should be noted that the comments made in this article relate only to standard registers of these two main variants of the Portuguese language. If we compared the regional dialects, we could perhaps find other situations, but this will not be considered at this time.

It is also worth mentioning that not all words or structures of EP replaced in the Brazilian version of the dictionary are necessarily absolute differences, that is to say, they can exist as a possible alternative for the Brazilian variant but they are not usually the preferred option. For example, in the entry *bus*, we had *autocarro* as EP equivalent, which is not used at all in BP, being then considered an absolute difference. In contrast, *tomar o ônibus*, the expression we obtained by changing *autocarro* for *ônibus* in the example, can be used in Brazil, but the expression *pegar o ônibus* is much more common, so in this case we refer to a preferential contrast, to use the terminology of Wittmann et al. (1995).

**bus** [bys] *nm* abreviação de autocarro  
{ep/bp} autocarro [awto'karu] *m* ônibus *m*  
◇ *prendre le bus*  
{ep/bp} tomar o autocarro *pegar o ônibus*

The issue of **phonetic transcription** of these equivalents is an important modification to the French audience interested in BP, but it is not discussed in this article since it was performed by Professor Luiz Carlos Cagliariari, not by our team. In the examples given here the transcription thus remains in EP, such as [ɛbɛdu'nar] for *abandonar*, [ɛbdumi'nal] for *abdominal* or [ɛbur'daʒɛj] for *abordagem*.

The **orthography** adopted in Brazil follows the standards of the Orthographic Agreement recently signed by the Portuguese-speaking countries<sup>2</sup>. In this regard, there are several instances of changes, especially when EP has associations of consonants, such as *ct* (*atividade* > *atividade*), *DC* (*inspeccionar* > *inspecionar*), *CC* (*injecção* > *injeção*), *nn* (*connosco* > *conosco*), *pt* (*óptico* > *ótico*), etc.

For the equivalents, the **translation** work was based mainly on the electronic dictionary versions of Houaiss (2009) and Aurélio (2004), the online *Sensagent*, the *Grande dicionário francês-português* (1998) and the *Dictionnaire électronique d'expressions idiomatiques français-portugais* (online).

In order to find out the usage frequency of the words and phrases, we tried to look up the *Corpus Brasileiro*, a BP corpus available on the Internet (<http://corpusbrasileiro.pucsp.br/>) and thus accessible to the public. However, according to the introductory text of the project, this corpus is not ready, its

initiators have not achieved its billion-word target and, in addition, searching it still presents double results. We therefore decided to consult the Internet for information on frequency and collocations, despite its drawbacks (idiosyncrasies, “errors”, etc).

In many cases, it is quite difficult to distinguish between the changes due to morphological phenomena and those due to syntactic phenomena. Still, there are some modifications that can be considered as changes on the **morphological level**, for example the option in each of the variants for different suffixes attached to the same base, as can be seen in the entries *bégaiment* and *dessinateur*. In the first case, the suffix *-ez* is attached to the base form *gagu-* in EP, and the suffix *-eira* to the BP variant. In the second case, we have the same base form, *desenh-*, to which we add *-ador*, in Portugal, and *-ista*, in Brazil.

**bégaiment** [begemõ] *nm* fait de parler en répétant les syllabes  
{ep/bp} gaguez [gɛ'gɛ] *f* gagueira *f*  
**dessinateur** [desinatœr], **trice** [-tris] *nmf*  
personne qui dessine  
{ep/bp} desenhador/-ra [dɛzɛnɛ'dɔr/-rɛ] *m*  
desenhista

In the last example, there is a difference between French and Portuguese, as in the Portuguese language the gender is invariable while in French there is one form for masculine and another for feminine.

We also found some instances where the change is not of the suffix but in the thematic vowel. In the first example of the entry *constitution*, there is the word *équipe*, which in EP is *equipa* and in BP is *equipe*.

**constitution** [kõstitysjõ] *nf* 1 formation  
{ep/bp} constituição [kõstitwi'sɛw] *f*  
◇ *la constitution d'une équipe*  
{ep/bp} a constituição de uma equipa / equipe

In the dictionary there are also cases of differences in the prefix, as in the entry *informe*, which has the EP equivalent spelled the same way as in French, but the BP one has the prefix *in-* replaced with the prefix *dis-*.

**informe** [ɛfɔrm] *adj* 1 qui n'a pas de forme précise  
{ep/bp} informe [fɔrmɛ] disforme  
◇ *une silhouette informe*  
{ep/bp} uma silhueta informe / disforme

At the morphological level, there are still instances where the word's gender “changes” depending on the variant. Looking at the example from the entry *pichet*, we could also say it in other words: there are two forms in this language that differ only with regard to the gender. For example, whether in Brazil or in Portugal, the words *jarro* (*nm*) and *jarra* (*nf*) are possible, but in the first, the feminine form is preferred, whereas in the second it is vice-versa.



**Helena Yuriko Sakano Fernandes** obtained a superior degree in Translation in 2009 and is currently undertaking her master in Linguistics at UNESP (Universidade Estadual Paulista – the São Paulo State University) in São José do Rio Preto, Brazil.  
helenayuriko@hotmail.com



**Claudia Maria Xatara** teaches French, lexicography and translation at UNESP (Universidade Estadual Paulista – the São Paulo State University) in Araraquara, Brazil. She completed her translation studies at UNESP in 1983, has been teaching there since 1984, and obtained there her master (1994) and PhD in Linguistics (1998). Professor Xatara has a post-doctorate in lexicography from Université de Nancy II in France (2005) and has carried out research at the Lexique, Dictionnaire, Informatique laboratory at Université de Paris 13 (2011). She is co-chief translator of *Kernerman French-Portuguese (Brazilian) Dictionary* (with Cristina Parreira) and project editor of K Dictionaries' Brazilian bilingual dictionary series.  
xatara@sjrp.unesp.br

**pichet** [piʃɛ] *nm* récipient pour liquides  
{ep/bp} **bidão** [bi'dew] *m* jarra *f*  
◇ *un pichet de vin*  
{ep/bp} *um jarro / uma jarra de vinho*

Divergences between the number (singular/plural) and the precision (defined/undefined) of the word were observed in several cases, but the contexts in which it happens must be researched further before making other comments or generalizations on the subject. Nevertheless, we present here two examples of the singular/plural difference in order to illustrate this phenomenon.

**change** [ʃɑ̃ʒ] *nm* action de changer une monnaie contre une autre  
{ep/bp} **câmbio** [kãbju] *m*  
◇ *un bureau de change*  
{ep/bp} *uma agência de câmbios/câmbio*  
**chasse-neige** [ʃasneʒ] *nm inv* véhicule qui enlève la neige sur les routes  
{ep/bp} **limpa-neves** [fipe'nevoj] *m* / *limpa-neve m*

In the examples of the entries, we could often observe syntactic changes. Among the analyzed parts of the dictionary, the most common cases concern the position of the pronouns. The BP variant prefers in most situations the proclise, while the EP variant makes more use of the enclise; therefore, in many examples where there were reflexive verbs or complements in a pronoun form, we had to change their order in the sentence, as can be seen in the EP sentence *Eles arranjaram-se* which became *Eles se arranjaram*.

■ **s'arranger** [sarãʒɛ] *vpr 1* se mettre d'accord  
{ep/bp} *arranjar-se* [arẽ'zarsɐ]  
◇ *Ils se sont arrangés.*  
{bp/ep} *Eles se arranjaram-se.*

Other very common cases are those names or verbs that are the same in both variants but which preferentially control different prepositions, as in the case of *arrancar*, which is used with the preposition *a* in EP, but in BP it is used more often with the preposition *de*.

**arracher** [araʃɛ] *vt 1* enlever en tirant fort (...)  
◇ *arracher des aveux à un criminel*  
{ep/bp} *arrancar a confissão a/de um criminoso*

There are also many cases of *preposition a + infinitive* structures in EP which had to be replaced with a gerund/present participle form in BP, as in the example below.

**arriver** [arive] *vi 1* parvenir à destination (...)  
◇ *Le printemps arrive.*  
{ep/bp} *A primavera está a chegar/chegando.*

Another difference between the two variants is the usage of personal pronouns referring to the second-person speech. In such case, in EP the pronoun *tu* is used when it concerns an informal relation. This pronoun is also used in certain dialects in Brazil, but the preferred standard register is the pronoun *você*. As can be noted in the following example, extracted

from the entry *content*, this change of pronoun implies also a modification in the verb conjugation, since the pronoun *tu* follows the conjugation of the second-person singular speech while *você* obeys the third-person singular speech paradigm.

**content** [kõtã], **ente** [-ãt] *adj* heureux (...)

◇ *Je suis contente que tu sois venu.*

{ep/bp} *Estou contente que ~~tenhas~~ você tenha vindo.*

Finally, there are also modifications in the order of words in the phrase. In the following example, the structure *ter suficiente dinheiro* is unusual in Brazil, but if the position of the adverb is changed we obtain a structure that is much more common (in this case, we changed the adverb too, but if *suficiente* had remained the result would be about the same).

**assez** [ase] *adv 1* en quantité suffisante (...)

◇ *avoir assez d'argent*

{ep/bp} *ter suficiente dinheiro / o bastante*

At the **lexical level**, there are several examples of absolute and preferential contrasts. Regarding these, there are cases where the equivalent proposed in the entry is adequate for BP, but when used in the examples, it must be changed. The following example may well demonstrate the role of the context in the choice of words.

The first sense of the entry *article* has *artigo* as its equivalent, which is undoubtedly used in Brazil with the meaning of "short text that is part of a book". However when it comes to the context of dictionaries, the short text within it is not commonly known as *artigo*, but as *verbete*.

**article** [artil] *nm 1* texte court qui fait partie d'un ouvrage  
{ep/bp} **artigo** [er'tigu] *m*  
◇ *un article de journal*  
{ep/bp} *um artigo de jornal*  
◇ *les articles d'un dictionnaire*  
{ep/bp} *os artigos/verbetes de um dicionário*

Certain difficulties presented at the lexical level are due to **cultural differences**. Based on our data and on the reading of Wittmann et al. (1995), these difficulties can be divided into two types: words without equivalents and institutional contrasts. The first are lexical units not used in the other variant, often due to the absence of a referent, and which, therefore, have no equivalent. "In general, these are common names (not scientific) of certain plants, fruits or animals outside the current speech of the other variant (...)"<sup>3</sup> As an example, we take the word *bruyère*, or *urze* in Portuguese. This word seems to be much more common in EP (51,500 occurrences on Portuguese webpages<sup>4</sup>) than in BP (9,440 occurrences on Brazilian webpages), but as the referent is not very common in Brazil either, there is no other more usual equivalent that could



be used as a substitute word for *urze*, and it should thus be kept in the entry.

**bruyère** [bryjeʁ] *nf* plante à fleurs violettes ou roses

{ep/bp} urze ['urzə] *f*

Similarly, *azevinho* is much more frequent on Portuguese webpages (42,700 occurrences against 8,500 on Brazilian webpages) but it had to remain as the equivalent, since there was no other word to replace it.

**houx** [u] *nm* petit arbre à boules rouges

{ep/bp} azevinho [əzə'viɲu] *m*

In the following case, we have a slightly different situation. The proposed equivalent for *bidonville* in EP, *bairro de lata*, does not score many occurrences on Brazilian webpages (7,580), but that is not due to the absence or weak presence of the referent in our culture. On the other hand, in Brazil the marked presence of this genre of housing groups has played a significant role in our demographic development and in our history, whereas this presence in Portugal is due to various historical reasons and it plays a different and particular role in their culture, so it seems natural that each country has its own name for these types of dwellings. Thus, for the Brazilian variant, we put as an equivalent the word *favela*, which is much more common (923,000 occurrences) in our culture.

**bidonville** [bidɔvil] *nm* groupe d'habitations où vivent des gens très pauvres

{ep/bp} bairro de lata ['bajru de'late] *m* / favela *f*

Citing Wittmann et al. (1995), institutional contrasts “cover words and phrases related to differences on the organizational level” of countries.<sup>5</sup> In the case of institutional contrasts, we have chosen as an example the entry *lycée*, which was translated in EP as *liceu*. This word exists in the Brazilian variant and is also quite frequent (211,000 occurrences), but in another context: it is widely used in proper names of this type of schools that are generally designated by *colégio* or, more specifically, *escola de ensino médio*.

**lycée** [lise] *nm* établissement scolaire du second degré

{ep/bp} liceu [li'sew] *m*, secundária escola *f* [ɛf'kɔlɛsɔkũ'darje] / colégio *m*, escola *f* de ensino médio

With the help of examples extracted from the adaptation of a French-European Portuguese dictionary to the Brazilian variant, we have seen that the differences between the two variants of Portuguese are numerous and exist on all linguistic levels. So, in the hope that other dictionary publishers take the same initiative, we can only praise this initiative to avail one same macrostructure in French into two main variants of the Portuguese language, as this will increase the availability of bilingual dictionaries of

French for the Brazilian public and will also help the French public to produce texts in Portuguese without mixing variants.

#### Notes

1. “Assim, sustentamos que a seguinte frase, por exemplo, *Quando mo deu, ele não tinha se apercebido... não é português correcto, porque mistura dois fenómenos sintácticos de variantes diferentes: mo não é usado em português do Brasil, enquanto que a ordem dos clíticos não tinha se apercebido não é aceitável em português europeu.*”
2. Cf. the pull-out of the 5th edition of the *Dicionário da Academia Brasileira de Letras*, <http://www.academia.org.br/>, and the *Guia Prático da Nova Ortografia* (TUFANO, 2008).
3. “Em geral, são nomes vulgares (não científicos) de certas plantas, frutas ou animais não pertencentes à linguagem corrente da outra variante, embora não exista nenhum equivalente, como por exemplo *azinha* (PE) ou *sapoti* (PB).”
4. All frequency searches presented in this article were performed on March 7th, 2011.
5. “**Contrastes institucionais.** Cobrem palavras e expressões relacionadas com diferenças a nível organizacional entre Portugal e Brasil, como é o caso, por exemplo, do sistema educacional (*liceu, primeiro grau*), das regiões administrativas (*distrito, estado*) de instituições oficiais, etc. Este tipo de contrastes compõe um conjunto à parte por representar uma realidade equivalente mas não igual no âmbito cultural dos dois países.”

#### References

- Aurélio = Ferreira, A.B.H. *Aurélio século XXI* (versão 5.0). Rio de Janeiro: Positivo. 2004.
- Azvedo, D. *Grande dicionário francês-português*. 13th ed. Venda Nova: Bertrand. 1998.
- Dictionnaire électronique d'expressions idiomatiques français-portugais*. [http://cntrl.fr/dictionnaires/expressions\\_idiomatiques/](http://cntrl.fr/dictionnaires/expressions_idiomatiques/).
- Houaiss = Houaiss, A. *Dicionário Houaiss da língua portuguesa* (versão eletrônica 3.0). Rio de Janeiro: Objetiva. 2009.
- Sensagent = <http://dictionary.sensagent.com/>.
- Wittmann, L.H. et al. 1995. “Português Brasileiro e Português de Portugal: algumas observações”. In *Anais do XI ENCONTRO NACIONAL DA APL Lisboa, 2-4 de Outubro de 1995*. <http://linguateca.pt/Diana/.../WittmannPegoSantosAPL95.pdf/>.

Encyclopædia Britannica launched an educational website for primary schools (K-6) in Brazil in April 2011 (<http://escola.britannica.com.br/>), which features an online application of *Dicionário Houaiss Kernerman da língua portuguesa para estudantes*. This monolingual learner's dictionary for native speakers of Brazilian Portuguese consists of 30,000 entries and forms the first title in the new framework of cooperation between K Dictionaries and Instituto Antônio Houaiss from Rio de Janeiro. The editor-in-chief is Ieda Maria Alves and the editors include Ana Maria Ribeiro de Jesus, Luciana Pissolato de Oliveira, Liana Koiler and José Monteiro.

# Kernerman Dutch Dictionary

Truus Kruyt



**Truus Kruyt** worked on several lexicographical projects, among which are the historical dictionary *Woordenboek der Nederlandsche Taal WNT* (comparable with *Oxford English Dictionary* and *Grimm's Wörterbuch*), the Dutch *PAROLE/SIMPLE* computational lexicon and *Kernerman Dutch Dictionary*. She was head of the Language Database Department at the Institute for Dutch Lexicology in Leiden, and in this position has participated in many European projects in the field of language technology. She has lectured on Computer and Lexicon (Leiden University), Lexicography and Computer Technology (Ivanovo State University) and Dutch Linguistics (Westfälische Wilhelms-Universität Münster and Universität zu Köln), and is a member of national and international boards and scholarly committees. Since 2007 she has been working as a freelance linguist. [truuss.kruyt@hotmail.com](mailto:truuss.kruyt@hotmail.com)

## Introduction

K Dictionaries (KD) has developed a series of monolingual learners' dictionaries for over twenty languages, which should serve as a basis for bilingual dictionaries. A monolingual Dutch dictionary (12,000 entries) was compiled in 2007, by the lexicographers Rik Schutz, Truus Kruyt, Monique Woltring and Geert Bilsen. The dictionary has been published on *Woorden.org*, which has a collection of 170,000 entries derived from various online lexicographic resources. It is accessible for free. The entries from our dictionary can be recognized by the sentence "*Bovenstaande informatie is afkomstig van Kernerman Dictionaries*" (the information above is originating from Kernerman Dictionaries). Another distinctive feature is the pronunciation in IPA<sup>1</sup> representation. The entries from the other sources are less detailed and sometimes concern outdated words and spellings. The aim for the coming years is to improve and to update these entries, to start with words that are most frequently retrieved by the users of *Woorden.org*. For our dictionary, this ambition will be achieved in 2011 by an extension with another 12,000 entries (see the final section).

The dictionary was translated into French and translation to English, German and Spanish are on-going. They are all available on [Mijnwoordenboek.nl](http://Mijnwoordenboek.nl).

## Target users and headword selection

The monolingual Dutch dictionary was aimed to constitute the basis for bilingual dictionaries with a twofold target group: adults with a different mother tongue who learn Dutch in their own country, and Dutch native speakers who need a concise bilingual dictionary. The 12,000 entries of the dictionary are considered to represent a Dutch basic vocabulary which enables the user to largely understand an elementary text.

Sources for the selection of headwords included a wordlist provided by KD and several Dutch language resources for learners, among which *Van Dale Pocketwoordenboek Nederlands als tweede taal (NT2)* (*Van Dale's concise dictionary of Dutch as a second language*)<sup>2</sup>. Given the target group and the limited size of our dictionary, a major selection criterion was that the words to be included should frequently occur in Dutch daily life. For example, terminology,

less prevalent occupations, exotic animals and plants, rather unknown phenomena outside the Dutch-speaking regions, and non-naturalized international words and abbreviations were deliberately left out. Criteria for the distinction of homonyms were different parts of speech and different pronunciation, rather than etymology. The first two criteria are clear to learners, etymology is a too difficult matter.

The compilers were requested to critically consider the initial list of headwords, and to suggest entries to remove and to add in accordance with the target group. In this respect, texts on the Internet were often helpful, for example to check the current usage of words (printed dictionaries are inherently traditional) or as a source of inspiration for colloquial words to be incorporated in our dictionary.

## Structure elements of a special nature

The entries contain usual structure elements, such as headword variants, pronunciation, grammatical information (part of speech, gender, plural, singular past, past participle etc), meaning descriptions, synonyms, antonyms, examples, collocations and fixed expressions (idioms). Here we will pay attention to some special features.

## Pronunciation

As mentioned above, the pronunciation is represented by use of the International Phonetic Alphabet (IPA). Some sounds are not pronounced by all Dutch speakers, for example the *n* at the end of verbs (e.g. *eten* / to eat) and the *t* in a word like 'informatie' (information). We represented cases like these as follows: 'etə(n) and ɪnfər'mɑ(t)si (stressed syllable marked by a preceding ')'. This representation is in accordance with other Dutch reference works and with some bilingual dictionaries produced under the authority of the *Committee for International Lexicographical Resources*, subsidized by the *Nederlandse Taalunie (Dutch Language Union)*.<sup>3</sup>

## Meaning description / definition

For the determination of word meanings, the Internet corpus proved to be very useful. Of course, the output of searches is to be critically judged by an experienced lexicographer. But apart from meanings already recorded in dictionaries, the Internet corpus revealed new meanings

and new meaning extensions. For example<sup>4</sup>, dictionaries describe *geheugencapaciteit* (a compound of *memory* and *capacity*) as a computer term only, whereas the Internet also reveals the meaning ‘how much or how long you can remember things’, in contexts like ‘de geheugencapaciteit van je hersenen verbeteren’ (to improve the *memory capacity* of your brain) and ‘Een goudvis heeft een geheugencapaciteit van drie seconden.’ (A goldfish has a *memory capacity* of three seconds.).

For the meaning descriptions we adopted a concise and colloquial style with wordings as simple as possible. Some examples: *schuld* (*debt*) is defined as ‘geldbedrag dat je nog moet betalen’ (sum that you still have to pay), *sekse* (*gender*) as ‘feit dat je een man of een vrouw bent’ (fact that you are a man or a woman), *understatement* as ‘uitspraak waarmee je iets (opzettelijk) zwakker uitdrukt dan het in werkelijkheid is’ (statement by which you (deliberately) express something weaker than it actually is). In fact, the ambition of defining this way was one of the most challenging ones in writing the dictionary. We avoided idiomatic expressions in the definition and the use of a synonym as the sole definition was only allowed for words for which a common alternative is available. Sometimes a paraphrase rather than a definition was given. For example, the pronoun *jullie* (plural *you*) is described as ‘je zegt dit woord als je tegen twee of meer mensen praat’ (you say this word when you speak to two or more people), and *aha* has the description ‘uitroep bij positieve gevoelens’ (exclamation going with positive feelings).<sup>5</sup>

### Examples

There should be an example of usage (at least one) for each sense of polysemous entries. Monosemous entries have examples as well, as considered appropriate by the editors, and compositional phrases may also serve as examples in certain cases. The main function of the example is to demonstrate current lexical and grammatical collocations, frequent inflections or a characteristic grammatical construction. Furthermore, it serves to clarify the concise meaning description, in particular in the case of polysemous words for which the difference between meanings might be difficult to grasp for the target group. Examples can also give encyclopedic, cultural, social or other functional information to clarify the word’s meaning and usage. In general, the aim was to show the word in its natural context.

The example is preferably not a full sentence but rather a short phrase, or even a characteristic compound. For example,

‘een slagvaardige ondernemer’ (a decisive entrepreneur) (lemma underlined). We preferred examples that immediately come into mind hearing the word. In other cases, the Internet was helpful as a source of inspiration. We deliberately avoided the special features of many dictionary examples, such as overspecificity, use of the past tense, little relationship with the current social reality, etc.

### Working method

Preparatory work (headword list, author’s instructions, customized XML editor, etc) was done in consultation with the coordinating team and technical staff at KD. The lexicographers worked at home and sent their files by e-mail to each other for a quality check and feedback. Although useful in itself, this supervision was needed due to the high time pressure. The coordinator was Rik Schutz, who also did much correction work on the files after the project.

### Future perspectives

As mentioned in the introduction, our dictionary has so far been published online with other lexicographic resources, not primarily oriented towards learners. User statistics show that 10,000 out of our total of 12,000 entries were looked up in the first half year. This is a fairly high percentage given the basic character of the vocabulary.

In February 2011, a follow-up started, that is to be finished in autumn this year, extending the dictionary with another 12,000 entries. A main criterion for the selection of new headwords is frequency of retrieved words by the users of *Woorden.org*. These words often differ in nature from the headword list of the initial product. For example, the amount of intricate words and specialist terms appears to be much larger (e.g. *oormerken/earmark*, *synergie/synergy*, *concipiëren/conceive*, *verdisconteren/discount*, *agiol/premium*). It was also decided to extend the headword list with word categories that were left out in the first phase, for example abbreviations and acronyms, a broader range of international words in the Dutch language, such as *scouting*, *sealen* (*to seal*) and *schwung* (*dash*), words for phenomena outside the Dutch-speaking regions, such as *tsunami* and *gletsjer* (*glacier*), and names of exotic animals and plants. As a consequence, the target group has shifted towards a somewhat higher level than before. Still, the accessible style of the definitions and the example sentences is maintained. For example, *kubisme* (*cubism*) is defined as ‘kunststroming waarbij alle vormen met rechte lijnen worden weergegeven’ (trend in art that represents all shapes by straight lines).

The Dutch dictionary core is part of the new global series of K Dictionaries. It was compiled by Rik Schutz (chief editor), Truus Kruyt, Monique Woltring and Geert Bilsen, and is currently being revised and expanded by Rik Schutz, Truus Kruyt, Wilfried Dabekaussen and Hanne Bussels. It is used both in monolingual form and as a base for developing various bilingual versions.



New headwords have been inserted between the existing articles. The latter are being optimized during the compilation of the new headwords. In particular, many more synonyms and antonyms are added. This time of the compilation is assisted by data from Opentaal Foundation (including speech indicators, inflected forms of nouns and verbs, and word frequency information) and the Sketch Engine corpus of Dutch. The present lexicographers are Rik Schutz, Truus Kruyt, Wilfried Dabekaussen and Hanne Bussels.

Registered user needs may guide future extensions of the dictionary.

#### Notes

- 1 International Phonetic Alphabet.
- 2 Van Dale Lexicografie, Utrecht-Antwerpen. 2003, 2006.
- 3 <http://taalunieversum.org/taalunie/>.
- 4 Example from the current work.
- 5 Example from the current work.

## Anthropological and linguistic fundamentals of lexicographic work

Miguel Eduardo Montoro

### 1. Introduction

I got the opportunity to work in dictionary compilation in quite an unusual way. Although I had studied lexicography, the reason it was included in my studies and research differed radically from the way I used it over the last six years, in which it has become my livelihood.

My biggest passion is philosophical research, and before I started to work on dictionaries I had the privilege of being guided by one of the great masters in philosophy, Andrea Di Maio (Professor at the Pontificia Università Gregoriana, Rome), who steered me towards lexicography.

The reason that philosophy led me to lexicography is that every thought uses words as a vehicle for expression, and to understand the thinking of a philosopher in a profound way requires understanding precisely the corpus of words he or she uses, their polysemy, and the total structure of his/her language understood as a semantic constellation of hierarchically inter-related words—where some emerge in importance over others and determine the total structure of his/her thought. Ultimately, to understand an author thoroughly implies, at least mentally, ‘to make a dictionary’ of the author.

To do this, we face the textual corpus of an author as if we were facing an unknown language for which we need a dictionary in order to understand what he or she is actually saying. We study the words used by the author statistically, make a word frequency list, select (by contrast with the common language) those terms considered as *keywords*, and try to understand the author’s ‘own language’, constructing a

semantic map with all word meanings that are used and their close inter-relationship. This is carried out by using particularly arduous techniques, an explanation of which is beside the point here, but which requires a microscopic study of words.

Such a study is so deep that, in fact, by studying a single keyword from a particular author we could write a full thesis, which leads to the paradox that the dictionary of an author can become a very rich series of such theses.

For example, in my case, I started to study the word *resolution* in St. Thomas Aquinas, and after completing a synchronic study about this word within the Thomistic corpus, my supervisor asked me for a brief diachronic introduction to the word as found in Aquinas. That is to say, I had to write the history of how various influences had contributed in the course of time to Aquinas’s particular semantic nuances. So it was not enough to analyse the term as a lexicographer does, but rather as a historian of word semantics. I started this arduous research following the historical course and the clues all led me to about 17 centuries before Aquinas, to Plato. Since Plato wrote in Greek, I obviously could not find the very same word in his work, but an equivalent, *διαλεκτική* (dialectic). As a result, what should have been a minor aspect of my thesis turned into a thesis in its own right and I ended up doing a lexicographic study of *διαλεκτική* from Plato.

This is my background relating to lexicography, which eventually I started calling *high lexicography*, in contrast to *hard lexicography*, that is, dictionary making,

This article was translated from Spanish by Sabrina da Silva Pascoal (shimina\_pascoal@hotmail.com) and Gabriel Giglio (gmgiglio@gmail.com) from UNESP, Brazil.

The English translation was revised by Ilan Kernerman, and further updated by Charles Levine.

which later became my livelihood.

After deciding to change my lifestyle I started looking for work that would direct me towards linguistic areas such as translation, where I had reasonable success, to help me keep a good standard of living. Working as a translator I then had the opportunity to work as a lexicographer, which reminded me about my old studies, so I gladly applied for such work.

I pursued my lexicographic initiation in ‘doing dictionaries’, and not researching, but focused on general language dictionaries. That is when I hit the *hard lexicography* road.

My tutor in this journey, and boss at the same time, was Ilan Kernerman and his colleagues from K Dictionaries. With enormous patience, they transmitted to me their know-how. I must admit that I always followed them very obediently, suggesting at times what I thought was better suited, but trying to interpret what they asked for. After all, it seemed to be *their dictionary* rather than mine.

I must confess that sometimes I followed the instructions and explanations somewhat blindly, without a deep understanding explicitly formulated in my mind that could explain the way I do things. Over time, increasing my *hard lexicography* knowledge, and reflecting over it, I achieved my own way to formulate all the deep reasons that would endorse the guidelines received from Ilan and his team.

This is precisely the point of this article – the deep anthropological and linguistic fundamentals, as well as the advantages and unique characteristics, of the dictionary that our team has developed under the guidance and supervision of K Dictionaries. It is therefore possible that a lexicographer, who is engaged in *hard lexicography* research, is likely to find here some statements that seem naive when seen from a merely lexicographic point of view, which is however not our focus here.

Let’s discuss, then, the anthropological and linguistic foundations of the features of the Spanish dictionary core that we developed.

## 2. Succinct definitions

You may imagine how difficult it is for someone used to do a thesis based a single word to follow this simple instruction: ‘definitions should be succinct’.

When you do not assimilate in a practical way a particular concept you never know whether you are coming or going in this field. It takes a long time to realize it and to adapt to such an instruction, which may sound simple but is not simple to implement. Very slowly, and after several

attempts of trial and error, like all really strong and long-lasting learning, you start to assimilate what this means and all its practical implications.

Our instruction was that definitions should be as short and meaningful as possible, accomplishing the task of describing a given sense and disambiguating it in respect of any other sense or entry; that is, we should define the word itself and its semantic referent as distinct from any other word and also have the definition disambiguate a particular meaning as distinct from other meanings of polysemous words. All of this must be done in a ‘succinct’ way, that is to say using the smallest number of words and choosing the most meaningful ones with most expandable meanings in the task of disambiguating the entry.

Why, then, is having a *succinct* entry an advantage? Because—to begin with—the language function, understood as a systematic amount of orderly related words, is primarily apophatic.

What is the meaning of ‘apophatic’? It says what something is *not*, rather than what it really *is*. That is to say that in the proper construction of a definition we look rather at the word in relation to other words than at its existential semantic referent. The purpose is to ‘disambiguate’ it, that is, to distinguish it from other words. It is thus clear that when we define, we look more at the relationship between words than at the word as a semantic vehicle of a reality.

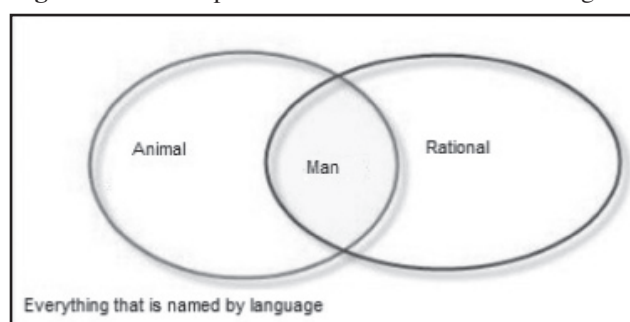
Mathematical logic can represent such relationships between the words employed in the act of defining, by means of mathematical sets (cf. [http://en.wikipedia.org/wiki/Mathematical\\_logic/](http://en.wikipedia.org/wiki/Mathematical_logic/)). For example, if we want to define *man* and take the Aristotelian definition, ‘a rational animal’ (just for the sake of illustration, putting aside all actual criticism and controversy about this definition), *animal* will be our hypernym (genre or super-category), and *rational* its particular characteristic (or specific difference)<sup>1</sup>. In mathematical logic, this could be represented as appears in Figure 1.



**Miguel Eduardo Montoro**

obtained his PhD in Philosophy with lexicographical orientation in 2003 at Pontificia Università Gregoriana, Rome, and has since taught various philosophical subjects in academic institutes in Argentina, Brazil and Italy. Dr Montoro is chief editor of the Kernerman Spanish Dictionary, and translator to Spanish of the French and Brazilian Portuguese dictionaries in the global series of K Dictionaries. emontoro@gmail.com

**Figure 1:** The complete definition in mathematical logic



First of all, what we see is a big picture, a frame of reference in which the activity of defining will be performed, this being the universe of all things named by language. So defining means to separate within the universe of nameable or utterable things that which may interest us among all other nameables. The first major division is the hypernym, which we represent in Figure 2 with a circle inside the square of the universe.

This circle already sets a limit, a distinction, a difference between what is inside the circle and what is outside it. And it tells us, therefore, that what I will define is *inside the circle*, and in a negative way that *it is not outside*. But this circle is not enough to give us precise coordinates of what we want to define, because inside it there is more than the very thing we define, which is *man*. In this circle are all sentient beings or animals together with *man*. Although this provides us with a coordinate by means of which we have separated *man* from the rest of the universe, i.e. the non-sentient beings, that is not enough. It is still indefinite inside that set, and this indefiniteness is what we call ambiguity.

Therefore, another coordinate is necessary, a particular characteristic or specific difference that distinguishes inside the circle of animals this sentient called *man* from the rest. This characteristic is, according to Aristotle, rationality. Thus, inside the box we draw another circle of rational beings, as shown in Figure 3.

We see, then, that the 'animal' circle intersects the 'rational' circle and the intersection of both forms the unique and particular space of *man*. So what we just did was to give the coordinates by means of which we could distinguish *man* from:

- all non-sentient and non-rational beings, such as plants, minerals, other physical or chemical entities, etc. This is the striped area in Figure 4;
- all sentient beings: animals (in the Aristotelian sense, including *man*), as shown in Figure 5;
- all rational beings: men, angels, God (apart, obviously, from their real existence; here

I refer only to their linguistic existence), as appears in Figure 6.

In this way, the intersection of these spaces locates the word *man*'s exclusive place of linguistic existence, as shown in Figure 7.

If we focus on nothing but this exclusive place, we will realize that, as a word, that is to say, as significant (according to Saussure, 1916), what we have is actually a set of negative coordinates of what is not *man*:

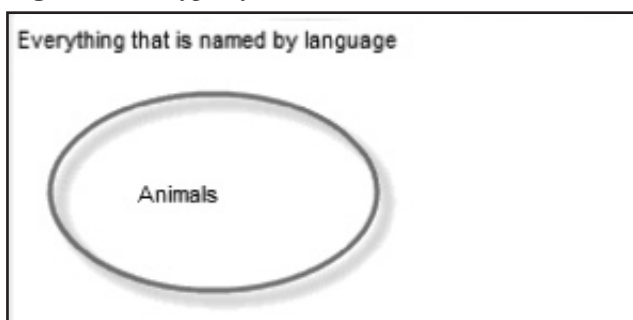
- it is not the set of all non-sentient beings: plants, minerals or other physical or chemical entities, etc;
- it is not the set of all sentient beings with no rationality, called in our modern semantics *animals* (not including *man*);
- it is not the set of all rational beings that lack sensibility, such as angels and God.

This implies that the word, as appropriate, is absolutely relative to where it is located in the semantic map or semantic constellation of a total language. So a definition is merely functional, serving only to give us the coordinates that are unique to what we want to define. But it could happen that for any reason these coordinates cease to be unique to that word. For example, if we discovered sentient rational life outside Earth, the Aristotelian definition would not be functional any more, given that it would not denote *man*'s new reality the way we would know it. We could then call *rational animals* both *man* and such aliens.

Definition is functional, knowledge is not. What *man* is will always be present in our cognitive capacity, because knowledge is contact, it is to touch the thing. Defining, however, demands a further activity to produce a representative '*verbum*' that could be transmittable to others and to me (significant), which entails the act of naming something. And it is because naming does not mean knowing that in our hypothetical case, in which the Aristotelian definition would cease to be functional, we would resort to the same knowledge to name new characteristics that distinguish *man* from aliens.

This analysis makes it clear that defining means giving coordinates, so a dictionary becomes a set of coordinates of the most significant words in a language. A dictionary is a GPS that tells us where we are when we look for a word. It demarcates an exclusive space where the word is and where the rest of the words are not. And the best way to give coordinates is using the most clear and determinant references to locate a place. Anyone travelling to an unfamiliar city knows that a long explanation of how to arrive somewhere is nothing but confusing and often ends up making us go astray or having trouble finding the place. It is usually enough

Figure 2: The hypernym





to explain directions with a minimum of the best and clearest references used as location coordinates. The same happens with words and language.

Some might argue that it is better to define with many words because that provides more information. And this is somewhat true. But let us think of the function of a dictionary. I insist on the image that a dictionary is not used to learn deep content, but to place a word somewhere in my total space of knowledge – to put it in a certain semantic place, uniquely. With this piece of information I can continue my research on the word, not in a dictionary, but inside this semantic place that we have negatively defined using our linguistic GPS. Thus, to move inside this linguistic space we need a specific GPS, more powerful, but somewhat more limited, as it only serves to explore this space and this is precisely the specific literature on a topic, which on the other hand is not exactly the function of a dictionary.

Because of all these arguments, I find this simple instruction I was given very beneficial: ‘definitions should be succinct’.

### 3. Dual disambiguation of meanings in polysemous entries

The second main guideline I was given was to provide another element of disambiguation for each sense of polysemous entries.

And I say ‘another element’ because the disambiguation by itself is provided by the definition. It means that to every polysemous entry we should add a meta-element that by itself segregates semantically the semantic place where the meaning must be comprised. For example, let’s look at the word *desfloración*:

**desfloración** [desfloraˈθjon] *nf*

1 =marchitamiento; envejecimiento del aspecto de alguien o de algo ◊ *la desfloración de sus energías vitales*

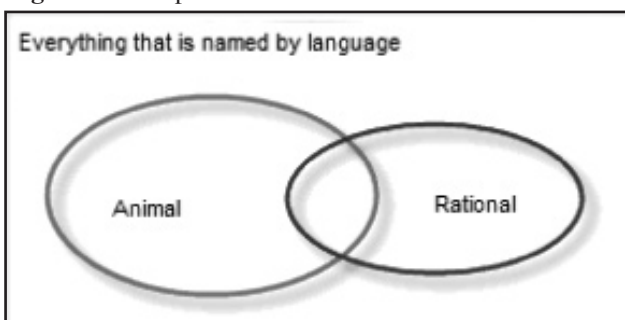
2 *bot* acción de sacar la flor a una planta ◊ *la desfloración de los rosales*

3 *coll* =desvirgamiento; acción de hacer perder la virginidad a una persona ◊ *desfloración de mujeres*

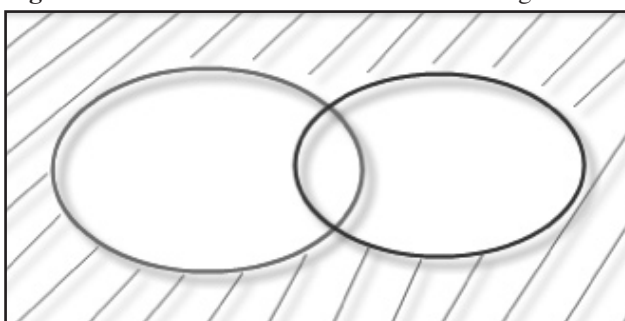
In addition to the definitions, which in themselves have the ability to disambiguate, different meta-elements are added to reinforce the semantic place in which the definitions must be located. To the first sense, a synonym is added, which by itself segregates semantically the other concepts. The same happens to the second sense by means of a subject field (*botany*) and to the third sense by the language register (*colloquial*) and a synonym.

So, if we compare the dictionary to linguistic GPS coordinates, each of these meta-elements would be another coordinate

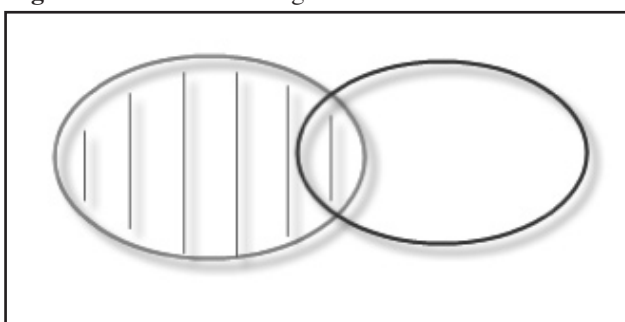
**Figure 3:** The specific difference



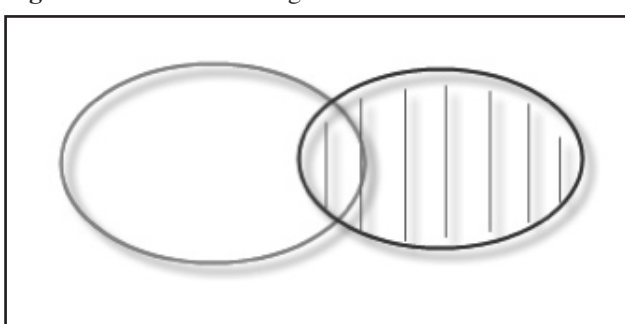
**Figure 4:** All non-sentient and non rational beings



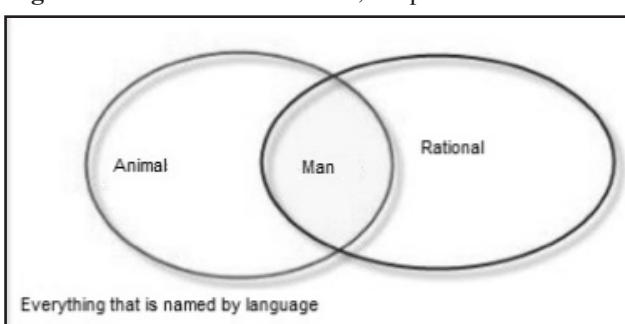
**Figure 5:** All sentient beings



**Figure 6:** All rational beings



**Figure 7:** Intersection of the sets, the place of the definition



**cañón** [ka'non] *nm*

1 *mil* tubo por el cual sale impulsado el proyectil de un arma

◇ *cañón del rifle*

2 *geog* estrecho profundo entre dos montañas

◇ *cañón del Colca*

**cañonazo** [kaɲo'naθo] *nm*

1 =descarga; bombazo disparado por un cañón

◇ *La ceremonia comenzaba con un cañonazo.*

2 =ruido; estruendo y daño producido por esta descarga

◇ *La ciudad quedó destrozada por los cañonazos.*

3 *sport coll* lanzamiento fuerte de la pelota al arco

◇ *Tiró un cañonazo al arco.*

**cañonear** [kaɲone'ar] *vt mil* disp

arar proyectiles con un cañón

◇ *El capitán dio la orden de cañonear el puerto.*

**cañoneo** [kaɲo'neo] *nm mil*=

bombardeo; lanzamiento de proyectiles

◇ *Comenzó el cañoneo al puerto.*

**cañonera** [kaɲo'nera] *nf*

1 *naut* hueco al costado de una embarcación o pared para disparar artillería

◇ *El fuego artillero salía de las cañoneras del buque.*

2 *mil* carpas usadas por los soldados en sus campañas

◇ *Desembarcaron e instalaron las cañoneras cerca de la costa.*

Sample entries from  
*Kernerman Spanish Dictionary*

that indicates the meaning in polysemous entries.

Why is it particularly advantageous to provide these elements in polysemous entries?

First, because more coordinates provide more possibilities to place the meaning of a particular concept linguistically.

But is this not the same as putting more information into a definition, making it longer?

The answer is no, because this information is provided, precisely in the manner of a meta-element, by presenting at the same time the information category. Many dictionaries also carry such information, that the term *desfloración* is the same as *marchitamiento*, without using the equal sign [=] that indicates the category of such information and shows that both terms are synonyms. This offers an additional advantage because more coordinates are given, not only by the amount of information itself, which is larger, but by the presentation structure of the information. If all the information were instead presented at the same level in a long definition without meta-elements, this would be more demanding on the reader, who, with the meta-element style of presentation, can have his/her questions solved in a clear and precise way, like well-demarcated signals, each one contributing to understanding a meaning.

The anthropological fundamentals of this feature lie in the fact that our mind organizes ideas and concepts as a three-dimensional map, where on the horizontal plane we can draw together the inter-relationship coordinates of words in what is termed a 'conceptual map'. For example, see the conceptual map of the word *plant* in Figure 8.

But such a conceptual map is used not only in the horizontal dimension of ideas, which involves the meaning relations among the words, but it also includes the

hierarchic dimension of the concepts' interaction among themselves, placing some concepts above others, showing the hierarchical supremacy of certain ideas that form the base of the concepts that they dominate. In this sense, expressing the contents by means of meta-elements is not only giving more content, but it also helps to create 'mental boxes' for the dictionary user, constructing his/her own conceptual word map in a more efficient way than the linear and non-specific account of a long definition that is full of data but which lacks the enhancement and the coordinates that are explicit in a conceptual map, and are also explicit in a definition done by means of meta-elements.

**4. All the entries have an example**

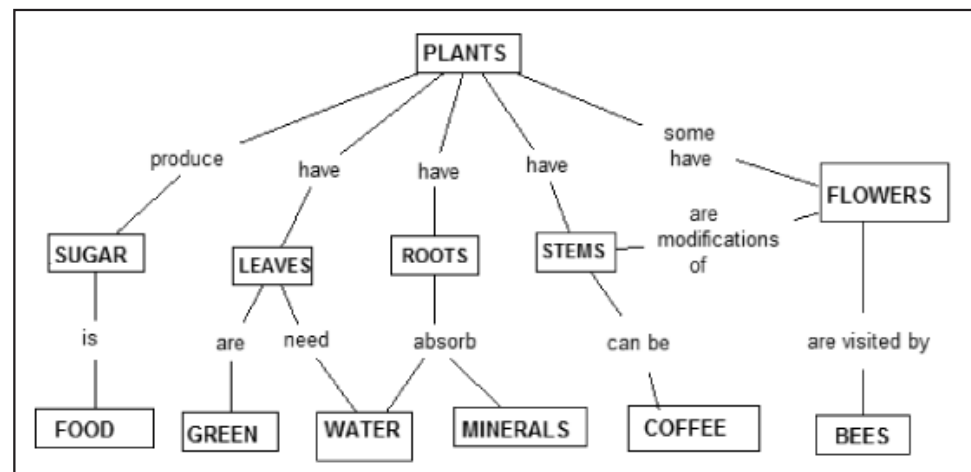
Perhaps this feature would sound more naive to lexicographers. Whom may it concern to discuss the importance of examples of usage in a lexicographic publication?

On the one hand, it is necessary to remember the viewpoint of this article, which is not quite lexicographical, but is rather anthropological and linguistic.

On the other hand, at least as regards the Spanish language, I have never encountered any monolingual dictionary that applies a standard rule of having examples of usage following each definition. But in this case, a lexicography professional may argue, and even be right, that such a principle, in which every definition must be followed by an example, is not necessary.

Nevertheless, in order to show the lack of examples in common Spanish monolingual dictionaries, I carried out a small statistical research—not scientific, though, but indicative (on account of the sample's size)—about the proportion existing between definitions and examples in polysemous entries (the average number of definitions per entries in the entire

**Figure 8:** Conceptual map



sample was 33). For this purpose, I used the three dictionaries most known and used for the Spanish language: The *Diccionario de la Real Academia*, *El Gran Diccionario de la Lengua Española Larousse* and the *Diccionario General de la Lengua Española*. The outcome was that in such context of highly polysemous entries, in these three dictionaries only 8.5% of the defined senses came with an example.

Therefore, on the concrete practice of lexicography, the obvious importance of using examples does not seem to be translated into real facts. Undoubtedly, there is the need to make other statistical studies on a scientific level, not merely indicative, to confirm the proportions and the hypothesis formulated above. And there is also the need to know the criteria that determines whether an example should or should not be used in a given definition, so that the person looking for the information may have a better judgment on the matter.

That being said, stepping back from the lexicographic view and into the anthropological view, we attempt to explain something tremendously obvious such as the importance of a definition being accompanied by an example.

To address this issue, I should state one of my theses with a personal opinion that may not be shared by others:

‘One truly understands a definition when the definition takes concrete shape, in an internal way, in which what is being defined is materialized by an example.’

I started to be truly aware of such thesis, even though I already knew the anthropological fundamentals previously, while making the corrections of the 50,000 entries of the dictionary. Then I realized that most of the time my mind could only distinguish the meaning of an entry or the meaning of a definition of a polysemous entry when I read its example(s). Moreover, when compiling entries, what first came to my mind was an example, and based on the example I would start to search for the genre (hypernym) to which the word belonged as its differential characteristic and, only then, I could shape the definition. Immediately after that, my mind started to check whether this definition would fit all the possible examples in relation to the defined meaning. What I discovered was that the mind makes a double recognition: from the example to the abstract definition and from the abstract definition to the example and to all its possible examples.

I mentioned above that this phenomenon happened ‘most of the times’, because every once in a while, in cases in which the meanings are more popular, the presence of an example

was not necessary in order to understand what was being defined. Nevertheless, I realized that in such cases the mind elaborates its own ‘examples’, its own materialization of the abstract definition known in classical anthropology as *phantasma*<sup>2</sup> and in modern evolutionary psychology, such as with Jean Piaget (cf. 1936, 1950, 1924), as *knowledge structures*. One may object that the *phantasma* of the scholastics is more static and that Piaget’s *knowledge structure* is, on the other hand, more dynamic, but in reality neither *phantasma* nor *knowledge structure* is, respectively, static or purely dynamic. Both describe the same episode, but from different viewpoints, leading to different conclusions.

The crucial aspect for the scholastics was to describe the structure of knowledge and, therefore, the *phantasma* can give the impression of being something static, when it is not, because it is not simply an evident representation, as a photograph would be, but it is a schematic representation that precedes the concept, in which, and by means of which, the intellect retains the concept. Besides, the intellect always needs this representation when we think and use the concepts by means of the *conversio ad phantasmata* (*conversion to sense experience*).

As to Piaget, the most important point is to describe, not only the structure, but the dynamics of the assimilation process. He does not consider knowledge as a pre-supposition, as scholastics do, but as the result of a process and of a movement. From this perspective, there are two types of scheme: the first, of action, and the second, of knowledge. The action schemes are principles of learning economy: once an action considered successful is executed, and then repeated several times, it creates an action scheme that turns such action into a permanent habit of the individual. On the other hand, the knowledge scheme is more structuring of the perception and it functions exactly as what organizes this perception. The scheme converts itself into a mental structure that is determined by an object. For example, this structure allows the possibility of executing actions towards absent objects and, therefore, it configures the perception towards a present or absent object. After this massive generalization, of structuring the perception of the object, a series of subtler differentiations according to its similarities and differences starts, and this is exactly the knowing process. Finally, it is possible to state that such schemes can be transferred and generalized.

However, beyond the difference between classical anthropology and its *phantasma* and modern anthropology and its *schemes*, what is most important

**cañonero, -ra** [kaño'nero, -ra] *nm/f sport* jugador que tiene un tiro potente y efectivo

◇ *el cañonero del equipo*

■ – [kaño'nero, -ra] *adj* nave, embarcación: que está armada con uno o varios cañones

◇ *un barco cañonero*

**canónica** [ka'nonika] *nf rel* vida monástica de los teólogos según las reglas antiguas

◇ *la canónica de la Catedral de León*

**canonical** [kanoni'kal] *adj* que se relaciona con el reglar

◇ *vestimenta canonical*

**canónico, -ca** [ka'noniko, -ka] *adj*

1 *rel*=eclesiástico; relativo a los cánones y disposiciones de la iglesia

◇ *Está estudiando derecho canónico.*

2 *rel* texto, libro: que es establecido o admitido por una tradición o religión

◇ *evangelio canónico*

3 =adecuada; que se ajusta a las características de un canon con exactitud

◇ *norma canónica*

**canóniga** [ka'noniya] *nf coll*

siesta que se toma antes del almuerzo

◇ *Se echó una canóniga en la última hora de clase.*

**canónigo** [ka'noniyo] *nm rel ecl* esiástico de una catedral

◇ *canónigo asesor*

Sample entries from  
*Kernerman Spanish Dictionary*



*Kernerman Spanish Dictionary* was developed from 2006 to 2010 under the editorship of Miguel Eduardo Montoro. The first part of the dictionary was compiled by Sebastián Cerón, Sergio López and Verónica Elizondo, and revised by Irene Renau Araque. The second part was compiled by Sergio López, Patricia Teijeiro, Ivana Fasano and Federica Urquiza, and revised by Elisenda Bernal Gallén. The third part was compiled by Sergio López, Ivana Fasano and Patricia Teijeiro. The first part corresponds to the 12,000 most common words, the second to the 24,000 most common words, and the third to the 50,000 most common words.

is that both movements teach and support that there is no pure thought (in the human being, obviously) and that the process of actualization with concepts is accompanied by a *phantasma* or a *scheme* that allows thinking in action.

This is the important point, when someone gives us a definition of something we can probably understand each word separately, but we may not understand what it truly refers to until the moment we create this *scheme* or *phantasma*, which comes from a sensitive order (not only thoughts) and implies, in a sense, a sensitive universal schematic representation of every materialization possible of the semantic space limited and demarcated by the definition.

In this respect it is important to stress the importance of examples in the areas of pedagogy and education. The example contributes directly to the elaboration of the *scheme* that allows us to understand the definition. Sometimes, when the definition refers to something that is extremely clear, there is no need for examples, because our mind elaborates the *scheme* as a sort of sensitive generalization of every possible example that might work for the definition.

But this does not happen all the time, nor most of the time and, especially, this is not how it works when we learn something new.

If there is something in which the majority of thinkers are in agreement, and that most thinkers in history had the tendency to use, it is the ‘definition–example’ pair, when something new is being taught. First, it is necessary to state the definition as an abstract mold that needs to be filled by the materialization of the example that always follows such a definition. This way of passing on education and knowledge has not changed since the early days of history until today, and it indicates something structural within human beings.

‘Something structural’ refers to the convenience and/or necessity for those who learn, that examples are essential to truly understand what is being said.

However, if this is how it works for any kind of teaching/education, that is evermore so for a dictionary, which is the first giver of the semantic coordinates of a word. It does not matter whether it is a native speaker who is consulting a monolingual dictionary or someone learning a foreign language consulting a bilingual dictionary.

To sum up, I consider an advantage and a very important characteristic of the dictionary we elaborated the fact that each and every one of the definitions of the entries is accompanied by (an) example(s) of usage.

## 5. Conclusion

As expressed by the title, this article does not merely present a lexicographic point of view, nor a particularly lexicographic perspective, but mostly explains the anthropological and linguistic insights of a person who has been meticulously following the procedures of the lexicography professional. Surely, on a lexicographic level, there are strong reasons for the guidelines described above for creating a dictionary—to write succinct definitions; use meta-elements to disambiguate polysemous terms; and add illustrative examples for each sense. By writing this article, I wanted to share the amazement witnessed by a philosopher who is discovering, from his own perspectives and also from his science, that everything that he once executed blindly in creating a dictionary in fact has a deep foundation in the human being, in the way we learn, in the way we communicate and in the way we assimilate contents. I hope this interdisciplinary essay is, somehow, useful for lexicography professionals.

## Notes

1. This is referred to by modern lexicographers as ‘formal definition’ (cf. Trimbley, 1985: 75-76), which is sufficient for our purpose.
2. Cf. Aquinas, *Summa Theologiae* I, 85, 1, in *Opera Omnia*, 1992.

## References

- Diccionario de la Lengua Española de la Real Academia Española*. 22nd ed. Madrid: Spanish Royal Academy. 2001.
- Diccionario General de la Lengua Española*. Barcelona: Vox. 2009.
- Gran Diccionario de la Lengua Española*. Barcelona: Larousse. 2008.
- Opera Omnia Thomae Aquinatis cum hypertextibus in CD-ROM*. Busa, R.S.J. (ed.). Milano: Editel. 1992.
- Piaget, J. 1956 [1924]. *Le jugement et le raisonnement chez l'enfant*. 4th ed. Paris: Delachaux & Niestlé.
- Piaget, J. 1957 [1950]. *La construction du réel chez l'enfant*. 2nd ed. Paris: Delachaux & Niestlé.
- Piaget, J. 1998 [1936]. *La naissance de l'intelligence chez l'enfant*. 9th ed. Paris: Delachaux & Niestlé.
- Saussure, F. de, 1949 [1916]. *Cours de linguistique générale*. 4th ed. Paris: Payot.
- Trimbley, L. 1985. *English for science and technology: A discourse approach*. Cambridge & New York: Cambridge University Press.

## Pedro A. Fuertes-Olivera (ed.). *Specialised Dictionaries for Learners*

### 1. General presentation and structure

The work under review, edited by Pedro Fuertes-Olivera (University of Valladolid, Spain), is a Festschrift in honour of Professor Enrique Alcaraz Varó (University of Alicante, Spain). Professor Alcaraz (who died in 2008) was an academic, a sworn translator, and a dictionary compiler, who dedicated most of his career to LSP and translation (in particular in the field of English Law), and thus inspired a change in direction in English studies in Spanish universities. It is worth noting that the contributors to this collection were chosen not because of their acquaintance with Alcaraz, but because they were scholars who “could make a real contribution to the development of pedagogical specialized lexicography” (p.2), which is the focus of this book. This might account for the fact that Alcaraz’s works are only cited in two articles, which seems surprising at first sight.

After a preface by the editor, a short tribute to Alcaraz Varó and a selected bibliography of his works by two of his former students, the core of the volume is divided into three (unequal) parts. The first, which represents almost half of the book, comprises five papers by advocates of the Danish ‘Function Theory’ of lexicography – either founders of the theory or disciples – that show how it can contribute to the development of specialized learners’ dictionaries (SLDs). The second part groups four papers dealing with the contribution of linguistics as a whole to the development of SLDs, with a particular emphasis on culture-bound items, figurative meaning, lexical semantics and actants, and corpus linguistics. The third part contains only two papers, intended to pave the way for new developments in SLDs in two promising areas: Chinese lexicography and Internet terminological dictionaries. The book ends with a two-fold bibliography of the dictionaries quoted within the book (approximately 60) and other literature. There are also useful notes on the contributors, and a name and subject index. There is no general conclusion to the work.

This volume is the first, to our knowledge, to explicitly bring together the dual issues of specialized dictionaries and learners’ dictionaries, i.e. SLDs. We thus expect to find here a discussion of the notion of ‘learner’, a rather complex concept since it can refer to the learner of a language or of a specific field of interest; as well as of

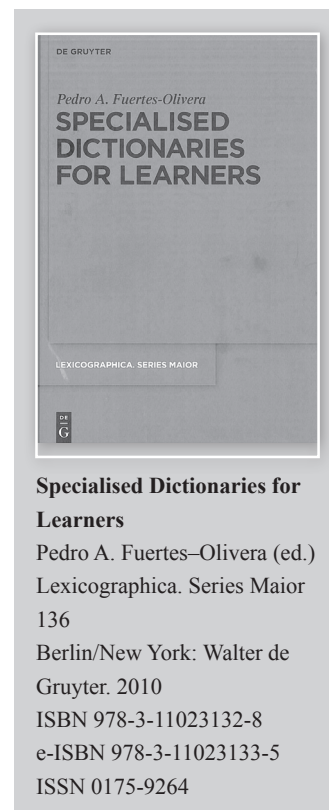
both theoretical and practical considerations related to the purposes and the designing of dictionaries aimed at this particular category of user.

### 2. Description of the chapters

#### • Part I

The very first chapter, Henning Bergenholtz and Sven Tarp’s paper entitled “LSP Lexicography or Terminography? The Lexicographer’s Point of View”, sets the tone for the first part of the work, which is deeply rooted within the framework of the Function Theory (FT) developed at the Aarhus School of Business in Denmark. Along with the second chapter, also by Tarp, it outlines some of the tenets of this theory relating to user needs and the functions of dictionaries. The last three chapters focus on particular types of dictionaries: Chapter 3 deals with monolingual SLDs while Chapters 4 and 5 both focus on bilingual ones.

Chapter 1 opens with a discussion about the difference between lexicography and terminology. The authors, who insist on saying the title of their paper was not their own choice but the editor’s, claim that there is no clear dividing line between specialized lexicography and terminography, which they actually envisage as synonyms. Although most of the arguments put forward are valid, the last one, about ‘research funding’ and ‘influences and positions at universities’, seems debatable. In what appears to be a response to one of Humbley’s papers (2002), the authors demonstrate that the FT bridges the gap between specialized lexicography and terminography and does not widen it. To do so, they focus on the concept of user needs, which, they say, is very closely linked to the specific situations the users are in. Quoting Tarp (2008), they explain that “a lexicographical function is defined as the satisfaction of the specific types of lexicographically relevant needs that may arise in a specific type of potential user in a specific type of extralexicographical situation”. They list nine criteria to be used when drawing the profile of dictionary users, then remind the readers that the FT distinguishes between three types of situations: cognitive, communicative and operational. The rest of the article focuses on two main types of communicative situations: translation and reception. The section on translation tries to show how two translation dictionaries in which the authors were involved (the English Gene Technology



### Specialised Dictionaries for Learners

Pedro A. Fuertes-Olivera (ed.)

Lexicographica. Series Maior

136

Berlin/New York: Walter de Gruyter, 2010

ISBN 978-3-11023132-8

e-ISBN 978-3-11023133-5

ISSN 0175-9264

**eLEX2011**  
**Electronic lexicography**  
**in the 21<sup>st</sup> century:**  
**New applications for**  
**new users**

The second conference on electronic lexicography, eLEX2011, will take place this year in Bled, Slovenia, November 10-12. It is organised by Trojina, Institute for Applied Slovene Studies, under the aegis of the European Association for Lexicography (EURALEX).

The conference aims to build on the first eLEX conference, held in Louvain-la-Neuve, Belgium (<http://uclouvain.be/en-cecl-elexicography.html>), which was very successful, and proved that electronic lexicography needs an event where current projects are presented and topical issues are discussed.

The programme will consist of presentations of papers, software, and posters. A wide range of topics will be covered, from usage practices of electronic dictionaries, new dictionaries and dictionary databases, and the latest developments in dictionary writing systems, to the exploitation of language resources, such as corpora and Wordnet, and natural language processing tools.

The keynote speakers come from both academia and industry, and include Sylviane Granger (University of Louvain), Simon Krek (Amebis software company and Jozef Stefan Institute), Erin McKean (Wordnik), Michael Rundell (Lexicography MasterClass), and Serge Verlinde (University of Leuven). ▶

Dictionary and Spanish Gene Technology Dictionary, 1998) meet the needs of their intended users (translators of specialized texts) because those had been anticipated by the compilers. The dictionaries thus provide knowledge on the subject field by means of systematic introductions to the field of molecular biology, and cater for the needs of the translators through every phase of the translation process: translators can find explanations in the mother tongue for the understanding phase, equivalents with relevant meaning discrimination for the transfer phase, and grammatical and syntactic information for the production phase. In the section dealing with reception, the authors introduce the difference between lay people, semi-experts and experts, and show how one single headword, *RNA*, has (or could have) two different definitions in the *English Gene Technology Dictionary* depending on the user's level of expertise. In the concluding remarks, while the authors seem to express doubts about whether terminographers are as interested in user needs as specialized lexicographers, they maintain that a "superior theory for a tool for specific needs" is much wanted, no matter whether it originates from lexicographers or terminographers.

As mentioned above, Sven Tarp is also the author of Chapter 2, entitled "Functions of Specialized Learner's Dictionaries". He postulates right from the start that "the very concept of a specialised learners' dictionary can only be defined by determining the possible lexicographic functions which these dictionaries may have". He claims that, so far, studies on SLDs have limited the concept of 'learner' to that of the learner of a language, neglecting those acquiring knowledge about a specific field. Before outlining what he means by lexicographic functions, Tarp explains that the user's needs are punctual and situation-dependent information, which leads him to analyze the various types of situations related to the learning process. A learner must acquire knowledge, but also skills, which he/she finds necessary to divide into two main categories: communicative/linguistic skills and practical skills. The former consist mainly of skills for text reception, text production and translation – which can be further sub-divided into 24 sub-categories, while the latter may be interpretive or operative. Knowledge and skills, he argues, are distinct categories, even as far as language is concerned. While knowledge can be rather easily acquired by learners, two mediating elements are required to transform the information retrieved from the dictionary into linguistic skills and practical skills: communication

and practical exercises/training. During the learning process, the learner also finds him/herself in various cognitive situations where the SLD can be of great help. For instance, when the learner systematically studies the subject field, he/she may find it useful to turn to a dictionary (instead of a handbook) to have a systematic overview of the field, such as the ones found in the above-mentioned Gene Technology dictionaries, or a systematic introduction to its specific LSP. In addition to cognitive situations, which are the "easiest to deal with" according to Tarp, communicative and practical situations are those in which an SLD may also satisfy users' needs in a way that should be further investigated. By combining the user profile with the various situations described, the lexicographer may be able to identify the essential functions of an SLD. Although the chapter is well-written and most of the explanations are clear, more examples would have been welcome.

Chapter 3, by Rufus H. Gouws from the University of Stellenbosch in South Africa, focuses on a particular type of SLD: the monolingual one. It starts with a much needed discussion of the concept of the learner, which is twofold in the case of pedagogical specialized lexicography: it can refer to the learner of a language or to the learner of a subject field. Each of these two types of learners can be further sub-divided according to their degree of competence or knowledge, which has to be taken into consideration by the lexicographer. The language learners can thus be sub-divided into beginners, intermediate and advanced, while those of a subject field can be lay persons, semi-experts or experts. For the author, the average user of an SLD is a semi-expert at the intermediate language level. After analyzing dictionary users in accordance with FT criteria, the author examines dictionary functions. He believes that, contrary to what is advised for *general* learners' dictionaries, in SLDs text production should be given more prominence than text reception. This means that fewer lemmata can be included, but included items have to be given a more comprehensive treatment. Gouws makes the point that the dictionary structure should logically result from the analysis of the dictionary functions and deals with the following aspects: (i) data distribution: he insists on the importance of including texts dealing with issues related to the subject field outside the macrostructure (called "outer texts"); (ii) access structure: very interesting suggestions are made regarding ways of improving information retrieval in the dictionary depending on the needs of the user – the example of a possible treatment of the term *nitric acid* is



particular convincing; (iii) microstructure: the discussion about the treatment of meaning and the phrasing of the definitions depending on the profile of the user is very relevant and the author also suggests adding a bilingual dimension to the microstructure, while specifying that this would limit the target user group; (iv) grammar: since SLDs are more geared towards production, grammatical data should be included, but outside the entry proper so as not to impede access to subject-specific data. The next section discusses the problems related to the type of monolingual specialized dictionary to be produced, (basically single field versus multifield), in particular the placement and ordering of outer texts to which the author seems rightly attached. The last section related to culture-dependent or culture-independent items is rather disappointing as it only alludes to the issue, which would have warranted a more in-depth analysis, but this gap is partly filled in Chapter 6.

With Chapter 4, by Sandro Nielsen from the Aarhus School of Business, we move from the monolingual dictionary to what is specifically called *specialized translation dictionaries for learners*, within the framework of the FT. The author makes the case for dictionaries that view translation as a whole, not limited to term-to-term equivalence (it is texts, not words, that are translated). He is very critical of what he calls the traditional specialized dictionary (though giving no precise reference) which, according to him, only helps to translate a small portion of LSP texts. He feels that too little attention has been paid to research in the field of translation studies, which is more useful to specialized lexicographers than surveys dealing with LSP translation that provide them with only vague ‘hints’ of what to do. The findings of translation studies that seem most useful to him are the following: translation is concerned with units larger than words; the clause and sentence levels are of primary importance; and, the receiver’s perception of the target text is increasingly important. Just like Gouws in the previous chapter, Nielsen identifies several categories of users depending on their degree of competence in the language or the subject field, but admits that the categorisation of users is very hard to do. For him, the most common user of an SLD is a lay person or a semi-expert “at the best”, which is a slightly different view from Gouws’s. He also recommends that lexicographers identify the factual, linguistic, textual and cultural competences of the intended users. The rest of the article is devoted to two focal points of LSP translation: linguistic structures and

genre conventions. The third focal point, terms (which includes collocations), has already received much, if not too much, attention according to him. Since syntactic structures in the source language and the target language can differ considerably (he gives the example of structures typical of the Danish and German legal language, which are not found as such in French and English translations), he insists that contrastive information should be given in the dictionary. As far as genre conventions are concerned, he states that learners must produce translations that conform to the conventions of the target language culture: they must therefore be systematically made aware of differences in cases such as the use of capital initial letters in some specific genres or sub-genres. He proposes that all the information relating to syntactic structures, genre conventions and what he calls “translation strategies” should ideally be placed in separate sections in the back matter, and these sections should contain illustrative examples that show translation strategies. For him, the best dictionary is a single-field dictionary, possibly with an electronic dimension, that focuses primarily on usage, and can be considered as an augmented reference tool.

Specialized translation dictionaries are also the focus of chapter 5, by Ildikó Fata, a Hungarian lexicographer from the University of Pílicscsaba. At first sight, two chapters on the same topic may seem redundant, but the two are rather different. Although the article is divided into several sections (the numbering of which is not really clear when references are made within), it actually consists of two main parts: a theoretical part based on the FT, describing the specialized translation dictionary as a particular type of dictionary, and a more practical part consisting of a report on a bilingual German-Hungarian dictionary in the field of pension insurance, explicitly aimed at translators and interpreters, in which the author was involved. In the first part, Fata shows that bilingual specialized translation dictionaries are at the crossroads of several scientific disciplines, drawing on the findings of translation studies, metalexigraphy, corpus linguistics, LSP research and terminology. Since the author adopts Tarp’s definition of a translation dictionary as one “designed to assist the user in solving problems related to the translation process” (2004), she finds it useful to rely on a model of the translation process also adopted by Tarp (2007), to identify the various functions of the dictionary in relation to each of the steps of the process. In accordance with Nord (2002), she includes bilingual specialized

► The conference organisers – Trojina, Institute for Applied Slovene Studies – are currently working on projects that involve developing dictionary and corpus resources for Slovene, with particular focus on user-friendliness of dictionary tools and corpus interfaces. Other research conducted at the institute includes exploitation of lexicographic and corpus resources for pedagogic purposes. The town of Bled, where eLEX2011 will take place, is a small alpine resort in Slovenia, known for its lake and beautiful surroundings, which attract guests throughout the year. We warmly invite you to the conference, which we hope will prompt interesting debates on current trends in electronic lexicography and give the participants the opportunity to network and meet partners for future collaborations.

<http://trojina.si/elex2011/index.html>

**Iztok Kosem**

Trojina, Institute for Applied Slovene Studies  
elex2011@trojina.si

## ABBYY LINGVO for iOS apps

ABBYY is integrating K DICTIONARIES content in the LINGVO new application for iOS platforms, which enables iPhone, iPad and iPod Touch users to access a wide range of dictionaries for different languages from leading publishers worldwide.

The ABBYY LINGVO application has reached Top-10 Grossing Apps position in Russia for 4 months, is #1 Reference App in Russia and other countries, and was chosen as a Staff Favourite App in the iTunes section in 61 countries in April 2011. Features including fast and easy look-up and smart search for misspelled and inflected words allow users to make the most of the dictionary content on their iOS devices.

A total of 38 titles from K DICTIONARIES, covering 20 European languages plus Chinese, are available, consisting of pocket-sized dictionaries for travel and daily communication. The following language pairs are on offer:

Chinese (Simplified) – English  
 Chinese (Simplified) – French  
 Chinese (Traditional) – English  
 Czech – English  
 Danish – English  
 English – Chinese (Simplified)  
 English – Chinese (Traditional)



dictionaries among translation auxiliary materials. This section would have merited clearer treatment, all the more so as some quotations or figures (e.g. Figure 6) have not been translated from German and are not understandable for people who do not master this language. She also mentions a survey without giving sufficient information about it. In short, the first part of the article could have benefited from the inclusion of more examples. In the second part, which seems slightly disconnected from what preceded, she explains the context of developing Hungarian metalexigraphic research which has led to the publication of the translation-oriented bilingual LSP dictionary of pension insurance. Although the main target users are native Hungarian translators and interpreters, five other categories of potential users with varying backgrounds were identified (for instance, pension experts who have excellent knowledge of the field but little knowledge of its LSP). Since the range of target users is very wide, the macrostructure had to take into account what was the minimum and maximum background knowledge required. The number of user situations envisaged was up to 18, the most interesting being those grouped under the heading ‘operational-orientated user situations’, whose treatment is however not discussed at all. The ‘introduction’ of the dictionary is mentioned several times but never quoted. Finally, the author claims that the novelty of the dictionary lies in the fact that each part is bilingual, something that is in fact not uncommon. The last section is devoted to the pedagogical dimension of the dictionary, which is interesting and could have been further developed. For instance, more concrete detail on how the corpus that was designed helped to compile the entries and the so-called ‘mini-contexts’ would have been welcome. All in all, this chapter would have been better served focusing on one aspect rather than two, the theoretical or the practical.

### • Part II

The second part deals with the treatment of specific linguistics topics in relation to SLDs. The notion of culture is the first to be dealt with, in Chapter 6, by Aquilino Sánchez, from the University of Murcia in Spain. The article discusses at great length the intricate relationship between language and culture, showing that language is a tool for the transmission of cultural knowledge, and that cultural features “are necessary ingredients for a complete understanding of a word”, and as such cannot be dissociated from words. Some examples are taken from everyday language, such as *breakfast*

in English and its so-called equivalent in Spanish, *desayuno*. The author insists on the fact that cultural features are best identified when two languages are compared, and quotes the well-known example of the various ways of referring to *snow* in ‘Eskimo languages’, as well as several examples from legalese in English and Spanish. He extends the problem to that of the varying ‘semantic space’ between several languages, relying on the in-depth contrastive analysis of *wood* in English and *bosque* in Spanish. The section devoted to ‘culture in traditional lexicography’ states that dictionaries have rarely paid direct attention to the cultural dimensions of words, but no traditional dictionary in particular is analyzed. Later on, though (in section 5), there is a critical presentation of LDELIC. Two main drawbacks are pointed out: the first has to do with the format of the dictionary, whose linguistic perspective is too traditional; the second concerns the inclusion of cultural notes, whose very existence seems to indicate that language and culture are two separate things, something unthinkable from the author’s point of view. It is only in section 7 that the reader finds some (heterogeneous) analyses of cultural items in existing dictionaries, some of which are not specialized. Surprisingly enough, the solutions proposed by Svensén (1993) are barely commented upon. In short, the author does not add much that is new. While he does outline interesting possibilities for future development such as the notion of a culture-dependency scale, or the idea of using Elementary Meaning Units to identify the cultural elements, these ideas are not developed. And while specialized dictionaries, either monolingual or bilingual, are discussed, SLDs are not, which puts this article outside the scope of this volume.

Chapter 7, by Geart van der Meer (University of Groningen, the Netherlands) is the shortest of all chapters. It tackles the very complex subject of the treatment of figurative meaning in monolingual SLDs. After stressing the importance for a learner of a specialized language to make the link between the metaphorical meanings and the literal meanings of a word or expression so as to grab their full meaning (something which is part and parcel of the native speaker’s competence), the author analyzes several examples taken from the business field, comparing definitions from various specialized dictionaries. He convincingly argues that it is more than desirable for a dictionary to make learners aware of what he calls the “double-sidedness” of the words, but then very honestly questions the feasibility of it in a monolingual SDL.

He comes to the conclusion that for some metaphors whose origin is rather obscure (like *benchmark*), there is no need for a tentative explanation in the dictionary, but that for others whose origin can be traced back (like *bubble* or *scorched earth policy*), some information should be presented to the user. The least space-consuming solution lies in the wording of the definition, which, thanks to collocations or words suggestive of the literal field, could help the learner make a link between the basic sense and the metaphorical one. He actually undertakes to adapt some of the definitions found in existing dictionaries, which proves a success.

With Chapter 8, we are back to purely linguistic matters: Marie-Claude L’Homme (University of Montréal, Canada) examines the treatment of arguments (‘actants’ in the theory of Explanatory Combinatorial Lexicology developed by Mel’čuk et al., cf. 1995) in several existing dictionaries, and reports on her team’s project to turn what was a rather formal terminology database, the *DiCoInfo*, into an SLD in the field of computing and the Internet. After a very useful review of the way five existing dictionaries (among which a learner’s dictionary such as COBUILD, and an SLD such as DAFA) deal with the problem of actants (through pronouns, a numbering system, general semantic labels, etc), she presents the original database of *DiCoInfo*. That database, unlike former specialized dictionaries, does not focus on conceptual information, but rather on linguistic information; moreover, it is clearly corpus-based, and does not limit itself to the noun grammatical category. Through the example of the entry *Internet*, the reader is shown the various information categories that appear in the entries, among whose striking features are the actantial structure, the linguistic realizations of actants and the lexical relation section. In the original version of *DiCoInfo*, actants are presented by means of actantial roles described by fifteen different labels, the most common being Agent, Destination, Instrument and Patient. These labels may sound familiar to linguists, but are definitely opaque for most users. This is the reason why the *DiCoInfo* team undertook to (automatically) convert the existing labels into more user-friendly ones. Relying on the above-mentioned review of the way actants are presented in existing dictionaries, the team decided, for various reasons that are very clearly outlined, that the most appropriate means of representing actants would be to resort to typical terms (as in DAFA). A thorny issue is the choice of the ideal typical term: several criteria are suggested, but the author shows

how difficult it is to combine all of them. Since learners always have difficulty with actants, and especially when dealing with LSP, this paper usefully feeds the discussion about SLDs both from a theoretical and a practical point of view.

The last chapter of the second part is dedicated to corpus linguistics. Lynne Bowker (University of Ottawa, Canada) first reviews the main lexicographic concepts at stake, i.e. specialized dictionaries, learners’ dictionaries, SLDs – concepts that have either already been defined or should have been defined earlier in the volume. Quoting Varantola (2003), she interestingly points out that a learner is rather different from a non-native speaker. The most novel part of this review lies in the presentation of what can be called *hybrid* learners’ dictionaries: dictionaries, such as MEDAL, which mainly deal with general language but tend to include a very large number of specialized terms. In the next two sections, she first explains the fundamentals of corpus linguistics, defining the term *corpus*, showing its advantages over introspection and giving general facts about the size of general-language corpora – a useful reminder for readers who would not be familiar with the subject. Then she traces the history of the contribution of corpus linguistics to general learners’ dictionaries, explaining that the use of corpora in *specialized* lexicography is not as systematic but is (slowly) taking up. Relying on a corpus helps general lexicographers in several areas: information about frequency can help to construct the headword list, compile the definitions and order the senses; it can also provide authentic examples – something of great help to learners – and help to identify grammatical and phraseological patterns. What seems more original in her explanation is the resort to learner corpora. The very last section of this article, about the potential of corpora for SLDs, is a key issue that, in our view, should have been the main focus of the chapter. Bowker demonstrates that even though the availability of specialized corpora is not so much of an issue any more, not many dictionaries seem to rely on them yet. According to her, all the above-mentioned advantages of resorting to corpora for general learners’ dictionaries can be applied to specialized ones, all the more so as the focus of the latter has recently shifted from text reception to text production by LSP learners. She seems to be in favor of a hybrid approach, but applied in reverse: new corpus-based specialized dictionaries could include both specialized terms and general words frequently used within the field. She recommends that SLDs rely both on corpora recording native speaker usage and learner

► English – Czech  
 English – Danish  
 English – Finnish  
 English – French  
 English – German  
 English – Greek  
 English – Hungarian  
 English – Icelandic  
 English – Italian  
 English – Latvian  
 English – Lithuanian  
 English – Polish  
 English – Portuguese (Portugal)  
 English – Romanian  
 English – Slovak  
 English – Slovenian  
 English – Spanish  
 English – Swedish  
 English – Turkish  
 French – English  
 German – English  
 Hungarian – English  
 Icelandic – English  
 Italian – English  
 Latvian – English  
 Lithuanian – English  
 Portuguese (Portugal) – French  
 Slovak – English  
 Slovenian – English  
 Spanish – English  
 Turkish – English  
 ABBYY LINGVO Dictionaries for iOS are available on the AppStore:  
<http://itunes.apple.com/us/app/lingvo-dictionaries/id391989146?mt=8#/>

Published in collaboration with K DICTIONARIES



The 9th International School on Lexicography  
*Multi-disciplinary Lexicography: Traditions and Challenges of the XXI century*  
 will be held at Ivanovo State University, Russia, September 8-10, 2011.  
 For more information please contact Professor Olga Karpova: olga.m.karpova@gmail.com or lexico2011@gmail.com

corpora. One last promising suggestion is to use spoken-language corpora, a road yet to be trodden.

### Part III

Finally, the last part deals with the challenges ahead for pedagogical specialized dictionaries. In Chapter 10, Zhang Yihua and Guo Qiping (from Guangdong University of Foreign Studies, China) give the reader a very interesting insight into Chinese pedagogical specialized lexicography. They start by demonstrating that although China has produced many *general* specialized dictionaries (GSDs), it still lacks SLDs. Most existing dictionaries are not appropriate because they are intended only for professionals (either field experts or translators), and do not record enough linguistic information. This is backed up with the results of a survey conducted among 128 undergraduate students who attend specialized courses taught in English: only a third use bilingual English-Chinese specialized dictionaries, because they find the existing dictionaries unhelpful. The authors infer from their findings that users are demanding a new type of dictionary. The growing bilingual context of education in China, from pre-school to higher education, makes this country a “huge potential market” for SLDs, with many textbooks written in English. Hence the need for a theoretical framework for a new type of dictionary. The authors first show how SLDs differ from GSDs in terms of users, purposes, linguistic, functional and structural features, then how they differ from general learners’ dictionaries. Next, they examine the functions of English-Chinese SLDs, which are both communicative and cognitive. They insist that both active and passive vocabulary be included in the dictionary, together with a sufficient amount of specialty knowledge. The structural features of the ideal SLD are also under scrutiny: dictionary design should take megastructure, microstructure, distribution and access structure into consideration. The sub-section dealing with the inclusion of lexical relationships is very complex, and Figure 4 in particular would have required further explanations. Section 5 discusses definitions; the authors argue that they should be ‘multidimensional’ (based on cognitive domains, and with a focus on the foreign tongue), ‘holistic’ (situated within a larger frame, with a mention of hypernyms, differentiae, co-hyponyms, etc) and ‘pertinent’ (in accordance with specific academic fields). They give a few much needed examples to illustrate their point, but the source of these examples is not clearly specified. Finally, the last section deals

with the tricky issue of the translation of culture-bound items. The authors argue that ‘calque’ – be it semantic, morphological, phonological or else – is a more effective approach than the usual free translation strategy. Although this section is very interesting from a contrastive point of view, its relevance regarding lexicography is not obvious, and it could have been shortened or more directly connected to lexicographical issues.

The very last chapter was written by the editor of the book, Fuertes-Olivera, and truly deals with “Lexicography for The Third Millennium” as it focuses on Internet dictionaries. After a reminder about the functional approach, which is the theoretical framework used in his paper, the author tries to refine De Schryver’s typologies of electronic dictionaries (2003) by adding three more criteria: (i) the identity of the compiler, (ii) the type of access to the dictionary – free vs. restricted, and (iii) the intended user of the dictionary. This leads him to highlight the category of institutional Internet reference works, which can be further sub-divided into restricted institutional Internet reference works and free ones. The former correspond to electronic versions of ‘traditional’ dictionaries, and are thus of lesser interest than the latter, which are the focus of the rest of the article. Those dictionaries, compiled by “amateur lexicographers” working in identifiable institutions, have rarely been studied, which makes the contribution really valuable. Based on the analysis of terms starting with the letter E in the dictionaries of finance retrieved by the Yourdictionary.com portal, the author has identified three sub-types: (i) glossaries compiled by consultancies or private companies, (ii) reference works by national or international organizations, which, more often than not, are electronic versions of previous paper reference works, and (iii) reference works compiled by amateur trained lexicographers working in the language industry. Unsurprisingly enough, it is the last category that fares the best as to users’ communicative needs. The lexicographic quality of the first two is indeed rather poor. The last part of this chapter consists of recommendations regarding the construction of future free institutional Internet reference works. The author lists six essential lexicographic requirements: the lemma selected should be relevant; the treatment of meaning should include contextualization and be aimed at semi-experts more than laymen; semantic relationships should be signalled; grammatical information should be separately included; the access structure should be user-friendly

and, information about the subject field should be provided. An additional set of recommendations is Internet-specific and reinforces those already made by Almind (2005). They include resorting to more user-friendly search engines and proper hypertextuality, and to regular updating. More importantly, the author advocates integrating the dictionaries in “teaching packages”. Although the validity and desirability of all those recommendations cannot be questioned, they seem hard to put into practice as the compilers of such dictionaries are said to be amateurs, and not professional lexicographers.

### 3. Conclusion

In the first section, we mentioned that one of the things we expected to find when reading the work under review was discussions about the concept of the learner. The book really was up to our expectations in that respect. As far as theoretical considerations are concerned, we can say that it exceeded all expectations, since approximately two thirds are dedicated to theoretical aspects. However, the practical applications of such bountiful theory are not sufficiently developed in the book in our opinion.

If we set aside our own expectations, is this volume true to what it claims to be in the preface and in the introduction? According to the preface (p.2), it has three objectives:

(i) “defend[ing] a function-based transformative approach centred on the dictionary and the users, investigating which lexicographic theories and principles are best suited for learners enrolled on LSP courses and/or Translation degrees”: the first sub-objective of the first objective is more than fulfilled, since the FT seems to be the only theory outlined here, which is bound to leave little room for a discussion about which lexicographic theory is best suited for learners. The book is slightly misleading in that respect as it might give the (false) impression that very few other scholars have studied the issue of user needs;

(ii) “honour[ing] Enrique Alcaraz’s pioneering visions and daily activities as a teacher, translator of specialized texts, and lexicographer”: although the book is of great relevance for teachers, translators and lexicographers, it is not directly linked to Enrique Alcaraz’s ideas, something which could have been the focus of a general conclusion;

(iii) “open[ing] up new lines of research in terms of the construction of pedagogically-oriented specialized dictionaries”: this is probably where the book proved to be more successful, and in that respect it can be said to honour the memory of Professor Alcaraz.

However, a more practical implementation of many of the theoretical lines of research would be welcome.

Finally, it is said in the introduction (p.24) that the book defends three main ideas – that “there is a need and a market for specialised dictionaries for learners”, “most of the existing specialised dictionaries are not adequate for learners”, and “we need a sound theoretical framework for coping with known and unknown challenges in the realm of pedagogical specialized lexicography”. From our point of view, the first two can be considered more as basic premises underlying the book than as topics actually discussed in the papers; it is only the chapter on Chinese lexicography that explicitly shows there is a market for this type of dictionaries, and the limitations of existing dictionaries are not systematically pointed out. The third idea does run throughout the volume, but is limited to the presentation of a single theoretical framework, which might be seen as too restrictive.

### References

#### 1. Dictionaries

- COBUILD = Sinclair, J. (ed.). (1987). *Collins COBUILD English Language Dictionary*. 1st ed. London/Glasgow: Collins.
- DAFA = Binon, J. et al. (2000). *Dictionnaire d'apprentissage du français des affaires. Dictionnaire de compréhension et de production de la langue des affaires*. Paris: Didier.
- DiCoInfo = L’Homme, M.-C. (ed.). *Le Dictionnaire fondamental de l’informatique et de l’Internet*. <http://olst.ling.umontreal.ca/cgi-bin/dicoinfo/search.cgi/>.
- English Gene Technology Dictionary = Kaufman, U. and Bergenholtz, H. in cooperation with B. Stumman, S. Tarp, L. de la Rosa Marabet, N. la Serna Torres and G. la Serna Miranda (1998). *Encyclopedic Dictionary of Gene Technology. English (with Spanish equivalents)*. Toronto: Lugus.
- LDELIC = Summers, D. (ed.). (2000). *Longman Dictionary of English Language and Culture*. 3rd ed. London: Pearson ESL.
- MEDAL = Rundell, M. (ed.). (2007). *MacMillan English Dictionary for Advanced Learners*. 2nd ed. Oxford: Macmillan Education.
- Spanish Gene Technology Dictionary = Kaufman, U. and Bergenholtz, H. in cooperation with B. Stumman, S. Tarp, L. de la Rosa Marabet, N. la Serna Torres and G. la Serna Miranda (1998). *Diccionario Enciclopédico*



#### KAMUS LENGKAP

Inggris-Indonesia  
Indonesia-Inggris  
A Dictionary for Learners of  
English  
Penerbit Erlangga  
Jakarta, Indonesia  
July 2010  
551 pages, 235 x 158 mm  
Hardcover  
ISBN 978-979-075-181-1  
<http://erlangga.co.id/>  
[http://kdictionaries.com/  
products/junior/ppid.html/](http://kdictionaries.com/products/junior/ppid.html/)

From the series  
KERNERMAN SEMI-  
BILINGUAL DICTIONARIES

*de Ingenieria Genética. Español (con equivalente in inglés)*. Toronto: Lugus.

## 2. Other

- Almind, R. 2005. "Designing Internet Dictionaries". *Hermes, Journal of Linguistics* 34: 37-54.
- De Schryver, G.-M. 2003. "Lexicographers' Dreams in the Electronic-Dictionary Age". *International Journal of Lexicography* 16:2, 143-199.
- Humbley, J. 2002. "Nouveaux dictionnaires, nouveaux rapports avec les utilisateurs". *Meta* 47:1, 95-104.
- Mel'čuk et al. 1995. *Introduction à la lexicologie explicative et combinatoire*. Louvain la Neuve: Duculot/Aupelf – UREF.
- Nord, B. 2002. *Hilfsmittel beim Übersetzen: Eine empirische Studie zum Rechercheverhalten professioneller Übersetzer*. Frankfurt am Main: Peter Lang.
- Svensén, B. 1993. *Practical Lexicography. Principles and Methods of Dictionary Making*. Oxford: Oxford University Press.
- Tarp, S. 2004. "How can Dictionaries assist Translators". In Chan S.-W. (ed.), *Translation and Bilingual Dictionaries*, 23-39. Tübingen: Max Niemeyer Verlag. (Lexicographica Series Maior 122).
- Tarp, S. 2007. "Lexicography in the Information Age". *Lexikos* 17: 170-179.
- Tarp, S. 2008. *Lexicography in the Borderland between Knowledge and Non-Knowledge. General Lexicographical Theory with particular focus on Learner's Lexicography*. Tübingen: Max Niemeyer Verlag.
- Varantola, K. 2003. "Linguistic Corpora (Databases) and the Compilation of Dictionaries". In van Sterkenburg, P. (ed.), *A Practical Guide to Lexicography*, 228-239. Amsterdam/Philadelphia: John Benjamins.

**Amélie Josselin-Leray**

CLLE-ERSS, Université Toulouse le Mirail,  
France  
josselin@univ-tlse2.fr

## Gilles-Maurice de Schryver (ed.). *A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks*

This *Festschrift for Patrick Hanks*, like Hanks's own career, covers a period of extraordinary interest and technological change for the practice of lexicography. As Yorick Wilks, looking back to the early 1980s, points out, "Computational search within large corpora ... was simply an aspiration." Dictionaries were compiled in hard copy (in 1978, only storage in fireproof cabinets saved fourteen years' worth work for the *Historical Thesaurus of the Oxford English Dictionary*, when the Glasgow premises were gutted by fire). Real-language evidence was similarly likely to exist only in the form of citation-based handwritten files. By 2010, dictionaries and reference had moved online, and the existence of significant corpora is now taken as a norm. *A Way with Words* charts some of the paths between the two extremes in the thirty years covered by the book. Understandably for a

festschrift, papers come from those (many of them long-term friends, colleagues, and associates) who have found Hanks's work particularly fruitful as representing the contribution of a key mover and shaker in the field.

### Introduction

In "Getting to the Bottom of How Language Works", Gilles-Maurice de Schryver sets the scene by outlining Hanks's career and significant publications (for example, 'Word Association Norms, Mutual Information, and Lexicography', co-authored with Ken Church in 1989). The three divisions of the book, theoretical, computational, and lexicographic, reflect the main areas of that career. This explicit connection provides a linking thread between the papers—as good a way as any of achieving a level of homogeneity for the book. A few of the links



are a little tenuous: for example, Jonathan Green's enjoyable article "ARGOT: the Flesh Made Word", on the development of Francophone slang, apparently had its genesis in a request by Hanks for a similar piece for the Elsevier *Encyclopedia of Language and Linguistics* (2005). However, the piece itself is so interesting that no-one could cavil at its presence. It is perhaps to be regretted, given Hanks's work in names lexicography, that there is no article which reflects this specific area.

As an addendum to the Introduction, sixteen pages are devoted to a bibliography of Hanks's publications, ordered chronologically. One of those listed as 'forthcoming' is also one of the most warmly referenced: *Lexical Analysis: Norms and Exploitations* (Cambridge, MA: MIT Press), which Wilks describes as "his forthcoming magnum opus". de Schryver writes with understandable admiration of his subject's continuing productivity: there is clearly no question but that the next few years are likely to be busy ones as 'forthcoming' is replaced by publication dates.

### Theoretical aspects and background

The first two papers in the theoretical section underscore the chronological reach of the book. The first of them, "Defining the Definiendum", is actually the last paper (unfinished at his death in 2007, and lightly edited here by Rosamund Moon) written by John Sinclair. Starting with the typical treatment in a dictionary of idiomatic phrases (appearing at the end of an entry, with citation forms that are likely to be ad hoc), Sinclair argues that corpus evidence offers a strong case for multi-word units of meaning to be given the status of headwords. Yorick Wilks's paper on "Very Large Lexical Structures", on the other hand, introduces a seminal paper from the past: it was first published in 1977, and opens a fascinating window on what is now a vanished world, since (as Wilks puts it) the text "refers to a thesaurus where one would now refer to WordNet".

### Computing lexical relations

Ken Church's paper "More is More", which opens the section on computational lexicography, takes a backward glance at the pioneering days when he "mocked up" something similar to the COBUILD corpus using the AP (Associated Press) newswire. ("I chose the AP wire merely because it was handy.") However, it swiftly comes up to the present day, engaging with Adam Kilgarriff's contention that "Googleology is bad science." The paper provides a stimulating overview of what (even allowing for provisos about dirty data) can be achieved, before coming to a

positive conclusion: "Everyone has more access than they ever had before. Life is good." Other papers in this section deal with computational approaches to the lexicon in English, German, and Czech—springing from Hanks's activities in the United States, Germany, and the Czech Republic.

### Lexical analysis and dictionary writing

Rosamund Moon opens the section on lexical analysis and its impact on dictionary-writing with "Words that Spring to Mind", a paper which presents a corpus study of the phraseology of *spring to mind*. Starting with an examination of contrasting dictionary treatments of the item, her contribution exemplifies her quotation from Hanks on the necessity for "patient studies at the word-face" as a prerequisite for demonstrating systems and formulating explanations. Sue Atkins provides a detailed account of the development of a recent and major database, the DANTE project, as developed for the *New English-Irish Dictionary* (2010), and considers its possible application to FrameNet. Other papers in the section include Kilgarriff and Richly's reflections on a possible route from corpus to dictionary: "Semi-Automatic Dictionary Drafting".

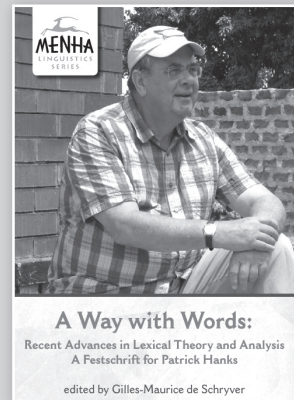
Human beings seek certainties. *A Way With Words* concludes with Michael Rundell's thought-provoking paper "Defining Elegance". In it, Rundell looks back to the early days of corpus lexicography, when real-language examples could replace lexicographers' constructs. He recalls the plaintive protest "But it was in the corpus" when he had to "confront members of the team with outlandish examples in text they had compiled". The point may seem an obvious one, but it is a useful corrective: Rundell has captured a moment at which what in any terms represented an exponential shift in resource was for that very reason being given a status beyond question. Every advance needs to be welcomed with an element of testing and scepticism.

### Conclusion

This is an enjoyable and thought-provoking volume, which brings together accounts of the early days of computational lexicography, with speculation as to where the future might take us. It will be of interest both to those who are intent today on exploring the latest developments, and those whose focus lies in tracing the changing history of lexicography over the past thirty years.

### Elizabeth Knowles

Editor, *Dictionaries* (Journal of the Dictionary Society of North America)  
elizabethmary.knowles@btinternet.com

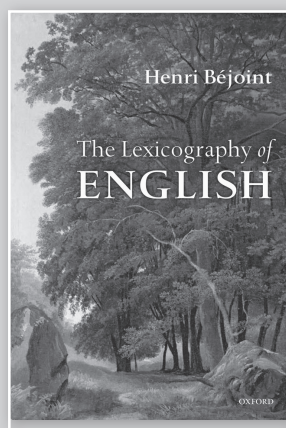


### A Way with Words: Recent Advances in Lexical Theory and Analysis. A Festschrift for Patrick Hanks

Gilles-Maurice de Schryver  
(ed.)

Kampala: Menha. 2010  
ISBN 978-9970-101-01-6

## Henri Béjoint. *The Lexicography of English. From Origins to Present*



**The Lexicography of  
English. From Origins to  
Present**

Henri Béjoint  
New York: Oxford University  
Press. 2010  
ISBN 978-0-19-829967-7

This well-written book is a treasure house of information. Since English prides itself on a greater number of various dictionaries than any other language, a major work on English lexicography inevitably becomes a survey and analysis of general lexicography. The balance between “theory” and “practice” in such works depends on the author’s tastes and predilections, because dictionary making hardly needs “theory” in the sense in which we understand *linguistic theory* or *theory of numbers*. Those attempt to explain the nature of certain phenomena, while “lexicographic theory” generalizes the experience of the profession. The difference between such questions as “What is a phoneme?” (or “What is a separate word?”) and “What is a dictionary?” requires no elaboration. Béjoint knows it and, most fortunately, never promotes *Theory* with capital *T*; Chapter 10 shows that in his case a small *t* may sometimes also be unnecessary. He has a realistic view of the uneasy union between lexicography and linguistics; especially revealing is the section on linguistics and structuralism (pp. 264-66). Lexicography cannot disregard the progress in semantics, but a dictionary purports to be a convenient reference book, and, inasmuch as no amount of theorizing will tell us where the line between several remote senses of a word and homonyms lies (to give just one example), every time we write an entry on a polysemous word, the age-old question presents itself anew. The same holds for the question about the descriptive versus the prescriptive mode (yes, dictionaries should provide us with a faithful transcript of the chosen language at any given moment rather than lay down the law, and yes, millions of people open dictionaries to find out how to spell, pronounce, and use words correctly, that is, according to the accepted standard) and about the breadth of inclusion (the more words between the covers, the better, but something must be left out, so where should one draw the line?). The recent literature on such matters is enormous, and lexicographers profit by knowing it. However, even the most erudite among them still depend on the requirements of the publisher, their own common sense, and intuition (a fancy synonym for experience that has become second nature). This book is certainly descriptive, not prescriptive, but Béjoint has given so much thought to lexicography that anyone who is interested in the subject will learn a good deal about all its angles from his exposition. His

additional strength lies in his expertise in French lexicography. The field being what it is, Béjoint had many serious predecessors. References to B.T. Sue Atkins, Michael Rundell, Patrick Hanks, and especially Sidney Landau have been strewn most generously in the text.

Despite its title, Béjoint’s book does not discuss English specialized dictionaries, dictionaries of slang, etymology, local words, usage (they are only mentioned in passing), let alone bilingual dictionaries of physics, engineering, medicine, and the like. An attempt to cover everything would have resulted in a multivolume version of *The Oxford History of English Lexicography*, a utopian project. As could be expected, not the same attention has been given to every major dictionary. For example, James Stormonth’s *A Dictionary of the English Language* (New York: Harper & Brothers, 1885) fell through the cracks. *The Century Dictionary* got minimal coverage (less than a page, p. 89). Béjoint says that it still has admirers. Since I am one of them, I am grieved to see how underestimated this magnificent work is. The same holds for Henry Cecil Wyld’s 1932 *The Universal Dictionary of the English Language* (“... good on pronunciation and etymology, and could have been successful if it had been published in other circumstances,” p. 110). This verdict and note 18 on p. 268 (“He [Wyld] wrote *A Short History of English* (1914), *A History of Modern Colloquial English* (1920), etc.”) show that Béjoint is not quite aware of Wyld’s stature and the excellence of *The Universal Dictionary*. On pp. 233-34 we read about the cases in which the outcome of court procedures depended on dictionary definitions. If Béjoint had followed the history of *Dictionary of American Regional English*, he could have added a few more curious examples to this section.

On the other hand, occasionally the book contains more than its structure and indexes suggest. For example, I looked up *etymology* in the index and found five references: etymology in Blount, p. 58; in Bailey, p. 64; in Richardson, p. 83; in *OED*, p. 102; in *W*, p. 134. Additionally, on p. 72, Béjoint speaks about etymology in Samuel Johnson; on p. 83 about Horne Tooke; on p. 140, about the *American College Dictionary*; on p. 85 about Webster’s derivations; on p. 45 about Walter von Wartburg’s *Französisches etymologisches Wörterbuch* (incidentally, headwords

in it are not Latin but reconstructed Proto-Romance or Proto-French forms, and what do the dates 1922-28 mean?). On pp. 59-60 *Gazophylacium Anglicanum* (1689), an etymological dictionary, is said to be “considered anonymous by some but attributed to Stephen Skinner, the famous author of the *Etymologicon Linguæ Anglicanæ* (1671) by others, with a second edition called *A New English Dictionary Showing the Etymological Derivation of the English Tongue* in 1691, which explained the etymology of ‘all common English words’, and it had a second part with proper names.” *Famous* is a relative concept, but who suggested that *Gazophylacium* was written by Skinner? The book seems to have been an abridged pirate translation of the *Etymologicon* by a well-hidden publisher. On p. 47, a footnote refers the readers to a bibliographical survey of English etymological dictionaries. Note 13 on p. 101 warns them that Bayle’s 1696 *Dictionnaire historique et critique* is a dictionary of history but adds that in 1992 a two-volume etymological dictionary (*Dictionnaire Historique de la Langue Française*) appeared. In connection with incongruent titles Béjoint cites Jacques Azaïs’s *Dieu, l’homme et la parole*, “a highly original etymological dictionary” (1853; p. 8). The usual trend is opposite, namely, to call a dictionary etymological only because it supplies the included words with etymologies. In England this tradition was perpetuated by Nathan Bailey. A more recent teaser using the adjective *etymological* as a marketing ploy is Chambers’ dictionary (for more than a century, beginning with 1867, *not* an etymological dictionary despite the promise on the title).

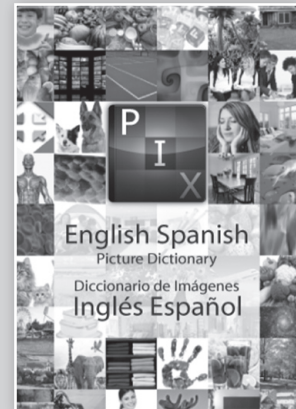
Collecting such crumbs for the index must have looked like a waste of time to Béjoint, the more so as, according to him, “[t]here were also ...dictionaries of etymology, although the field is more than adequately covered by the OED: the *Oxford Dictionary of English Etymology*, 1966, by G.W.S. Friedrichsen, R.W. Burchfield and C.T. Onions, no less, which had several reprints, the *Chambers Dictionary of Etymology*, 1999, etc.” (125). *No less*, as I understand, refers to the impressive team; the rest shows that it is better to stay away from the subject of which one has no firsthand knowledge. Needless to say, the *OED* is not an etymological dictionary, though Murray and Bradley were great etymologists and though their etymologies are splendid. Other great dictionaries also used to employ specialists familiar with the study of word origins. As pointed out above, polysemy is hard to keep apart from etymology, and we find remarks on this subject on pp. 263-64,

265, 275, 276, and 283. A passage (from an article by Rundell) on p. 282 is worth quoting:

“Most learners... would probably see some connection between *bay*, in its meaning of ‘an indentation in the coastline’, and *bay*, when it means ‘a recess’ (as in a *loading bay* or a *bay window*); conversely, few learners would see any connection whatever between the two main meanings of *club* (‘a society that people join’ and ‘a heavy stick used as a weapon’). Nevertheless, the historically motivated (but counter-intuitive) organization of the native-speaker tradition has in general been carried over into the MLDs [monolingual learner’s dictionaries], so that *bay* appears in L[ongman] D[ictionary of] C[ontemporary] E[nglish] and A[dvanced] L[earner’s] D[ictionary] as five separate noun homographs, while *club* appears as just one. Even more confusingly, *drill* (‘a tool for making holes’) and *drill* (‘a form of instruction based on repetition’) are grouped together in one homograph, while *drill* (‘an agricultural tool for planting seeds’) is shown as a separate entry.”

On the same page, Béjoint discusses *game* ‘activity, sport’ and *game* ‘wild animals’ and says (note 37): “From an Old Saxon word meaning ‘fellowship’, and the use of the word for an amusement, and a metonymy.” For *Old Saxon* read *Gothic*, and of course an English word, unless it is a borrowing, can be from neither Old Saxon nor Gothic. *Gaman* ‘amusement, diversion’ was attested in Old English. All this goes a long way toward showing that someone who wants to appreciate this book to the full should read it from cover to cover. I devoted so much space to etymology because it is my field, but I have no illusions about its place in lexicography and could offer many quotations like the following from Howard Jackson’s 1988 book *Words and their Meaning*: “[L]exicographers consistently, or perhaps persistently, put into dictionaries certain kinds of information for which the vast majority of users have no need and would not miss if they were not included in dictionaries. Into this category would come grammatical information including part-of-speech labels, etymology and perhaps pronunciation” (p. 244 here).

Besides a short introduction and an equally short conclusion, the book contains ten chapters: 1. “Dictionaries and the Dictionary,” 2. “A Brief History of English Dictionaries,” 3. “The British Tradition of the Scholarly Dictionary,” 4. “The American Tradition of the Utility Dictionary,” 5. “A New Tradition: The Dictionary for Foreign Students,” 6. “English Dictionaries of the Twentieth Century: The Cultural, the Functional, and the Scientific,” 7. “The

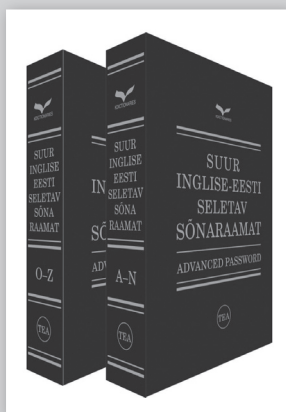


#### PIX

English Spanish Picture Dictionary for iPhone  
Marquee Publishing  
Toronto, Canada  
September 2010  
Compatible with iPhone, iPod touch, and iPad.  
Requires iOS 4.0 or later.  
<http://marqueepublishing.com/app-pix.html/>  
<http://itunes.apple.com/il/app/pix-english-spanish-picture/id389305736?mt=8#/>

Published in collaboration  
with K DICTIONARIES





#### ADVANCED PASSWORD

Suur inglise-eesti seletav sõnaraamat  
 Password Advanced English Dictionary for Speakers of Estonian  
 TEA Kirjastus  
 Tallinn, Estonia  
 April 2011  
 2,318 pages, 250 x 170 mm  
 Hardcover, 2 volumes  
 ISBN 978-9949-24-031-9 (vol. 1)  
 ISBN 978-9949-24-032-6 (vol. 2)  
 ISBN 978-9949-24-033-3 (vol. 1+2)  
<http://tea.ee/>  
<http://kictionaries.com/products/advanced/kaedee.html/>

From the series  
 KERNERMAN SEMI-BILINGUAL DICTIONARIES

Study of Dictionary Users and Uses,” 8. “Lexicography and Linguistics”. 9. “Computers and Corpora in Lexicography,” and 10. “A Theory of Lexicography?”. The book has been written for lexicographers, but at least two chapters (Six and Seven) will provide enjoyment to anyone who cares about language and culture. To give some idea of how chapters are organized, I will reproduce the rubrics of Chapter 6: 6.1 Cultural and Functional Dictionaries. 6.1.1 The mouthpieces of a culture (6.1.1.1 Politics and religion, 6.1.1.2 Ethnicity, 6.1.1.3 The taboos of sex and excretion, 6.1.1.4 Gender, 6.1.1.5 Proper names and culture, 6.1.1.6 What can the lexicographer do?); 6.1.2 Some dictionaries are more cultural than others; 6.1.3 The common features of modern dictionaries (6.1.3.1 Dictionaries are more user-friendly, 6.1.3.2 Dictionaries represent more varieties of English, 6.1.3.3 Dictionaries have more terms), 6.2 The end of a period? (6.2.1 Lexicographers: slaves or masters? 6.2.2 Lexicography: an art, a craft, or a science?). Each of the other chapters contains many more rubrics.

In Chapter 6 we read, among other things, about the difficulties of defining words and remaining, if not politically correct, at least politically neutral. In some cases this goal is unattainable (think of *God*, *communism*, and so forth). As regards Béjoint’s own formulations, I was amused only once, when I read that for a teenager the dictionary is a book to help her to do her homework (p. 227). Even in our progressive age I hope that boys also use dictionaries to prepare for classes. Some humor is unconscious. On p. 104, Béjoint explains what the term *derivative* means; his opinion of his readers’ level of linguistic sophistication becomes abundantly clear. On p. 221 he quotes a respectable author in whose opinion the Grimms came to lexicography from literature and lets it go without a word of comment. The section on students’ inability to use a dictionary would have been funny if it were not so sad.

Béjoint’s attitude toward scholarship and scholars is commendable. He does not want lexicographers to jump on the bandwagon of every linguistic theory (cf. “...the more information the linguists come up with, the more difficult it is to include it in a dictionary,” p. 345; “Linguists are interested in the definition of words, but they have not produced much,” p. 331). He is not enamored of Halliday’s theory to such an extent as to support its application to lexicography: “For Lexicographers, the distinction between syntax and lexicon is important because it determines what goes into the dictionary, as opposed to what

should be left to the grammar, and the current rapprochement is not much help: if there is no difference between lexical information and grammatical information, then the dictionary can contain anything, and the limit is only practical,” p. 40. Béjoint is aware of the circumstance that “[t]he corpus revolution has not solved all the problems of lexicography.... The corpus shows what is used, not what is not used but is part of the language.... It provides data, but cannot give explanations.... It is about performance, not competence” (p. 369). He has shown how dictionaries have progressed over the centuries and what excellent reference tools they have become. And yet “...one cannot help thinking that modern dictionaries, although more sophisticated than ever, have also become less imaginative, less exciting. In this sense, one can regret the old days, when dictionaries were much worse, and also much better” (p. 222).

*The Lexicography of English. From Origins to Present* is an eminently readable book. A set of illustrations (samples of dictionary pages) enhances its value, as do a copious bibliography and indexes.

The book is practically free from typos. I have noticed only a few: p. 7, top, *slov* -, not *slov*; p. 15, middle, Norwegian has  $\emptyset$ , not  $\ddot{o}$ ; p. 191, second paragraph, *Millennium*; p. 277, top: “In English, nouns are lemmatized in the masculine singular form....”: masculine?; p. 349, *Häufigkeitwörterbuch*: the first umlaut is missing. The reference to Hausmann et al. in Osselton 1999 should be to 1989. Somewhere Shcherba’s name appears as *Scerba* without hačeks, but I have lost the page reference. The only circumstance that impedes reading is a great mass of acronyms cluttering the text, but they were probably unavoidable.

**Anatoly Liberman**

University of Minnesota, Twin Cities  
 aliber@umn.edu

## Invitation to ASIALEX 2011 in Kyoto

A catastrophic disaster struck the Tohoku region in north west Japan on Friday, 11th March 2011. Thousands of people died or went missing, and hundreds of thousands became homeless in a matter of minutes. Such was the power of the initial earthquake and the tsunami it spawned. The Japan Self-Defense Force has been actively involved in providing assistance to the disaster-struck region, and various countries and international organisations have extended their support for the reconstruction of the disaster area. They have been working diligently day and night to help the victims. I sincerely convey my wholehearted sympathy to the victims and bereaved families, and express my gratitude to everyone for such enormous support. Thanks to this eager support, the Tohoku region will rise from the rubble, slowly but steadily.

Now, several months later, a variety of news reports on the aftermath of the earthquake and the status of the nuclear power plants has spread around the world. However, it is important to note that these plants are in Fukushima, which is located well north of Tokyo, and they have little impact on people around the country. Furthermore, the venue of ASIALEX 2011, Kyoto, is more than 500 kilometres away from Fukushima. Therefore, Kyoto remains unaffected and ASIALEX 2011 will be held there as originally scheduled.

ASIALEX 2011, the 7th biennial conference of the Asian Association for Lexicography (ASIALEX), will take place at the Kyoto TERRSA conference hall (a fifteen-minute walk from Kyoto Station, <http://www.kyoto-tersa.or.jp/>), from 22nd to 24th August. The conference theme is 'Lexicography: Theoretical and Practical Perspectives'. We are pleased to invite three keynote speakers. The first is Professor Zafar Iqbal (President, ASIALEX; University of Management and Technology, Lahore), who will deliver a speech entitled 'Lexicography: Building Bridges of Understanding, Communication and Universal Brotherhood'. Professor Robert Lew (Adam Mickiewicz University, Poznań) will give a speech entitled 'User Studies: Opportunities and Limitations'. And Professor Kikuo Maekawa (National Institute for Japanese Language and Linguistics, Tokyo), will deliver a speech entitled 'Development of Japanese corpora at the National Institute for Japanese Language and Linguistics'. In addition, we plan to conduct the following symposiums: E-dictionary (Organiser, Yukio Tono); User studies (Organizer, Yuri Komuro; Commentator, Robert Lew), and Phraseology and Dictionaries (Organisers,

Katsumasa Yagi and Ai Inoue). Furthermore, a workshop on software for lexicography is planned, including a demonstration of the TLex Suite (Organizer, Gilles-Maurice de Schryver). For more information please visit the website, <http://www.asialex2011.org/>.

Kyoto, the venue of ASIALEX 2011, was the capital of Japan for thousands of years and many elements of its precious heritage remain. You can visit temples and shrines (some are designated as national treasures and others as World Heritage Sites) and will also have the opportunity to enjoy Japanese gastronomic culture. I believe the participants will be able to learn about this beautiful city as well as about lexicography, both of whom are unique, and am confident that ASIALEX 2011 will be a fruitful conference. It is rather difficult to explain the wonderful experience awaiting you in Kyoto and ASIALEX 2011 in the limited space here. Seeing is believing! Please attend ASIALEX 2011 and visit Kyoto. We look forward to meeting you this summer!

### *Ai Inoue*

Department of Foreign Languages  
National Defense Academy of Japan  
narudo24@hotmail.com

The Asian Association for Lexicography (ASIALEX) was inaugurated at the Dictionaries in Asia conference in March 1997, at the Hong Kong University of Science and Technology.

The association was set up with a view to act as a catalyst for future cooperative research and to promote lexicographical development in Asia. It has since held the following international conferences:

1999 Guangzhou, China  
2001 Seoul, Korea  
2003 Tokyo, Japan  
2005 Singapore  
2007 Chennai, India  
2009 Bangkok, Thailand  
2011 Kyoto, Japan  
<http://asialex.org/>

### Selected publications by Shigeru Takebayashi

#### Books

- 1981 *Eigo no fonikkusu* [English phonics]. Tokyo: The Japan Times.  
1982 *Eigo onseigaku nyuumon* [An introduction to English phonetics]. Tokyo: Taishukan Shoten.  
1984 *Eigo no tsuzuriji to eigo no hatsuon no shidoo*—Lighthouse eiwajiten wo tsukatte [Teaching English spelling and English pronunciation—using the *Lighthouse English-Japanese Dictionary*]. Tokyo: Kenkyusha.  
1996 *Eigo onseigaku* [English phonetics]. Tokyo: Kenkyusha.

#### Dictionaries

- Kenkyusha's New English-Japanese Dictionary*, Enlarged Second Edition. Tokyo: Kenkyusha. 1951.  
*Kenkyusha's New Concise English-Japanese Dictionary*. Tokyo: Kenkyusha. 1955. [pronunciation]  
*Kenkyusha's New Pocket English-Japanese Dictionary*. Tokyo: Kenkyusha. 1957<sup>1</sup>. [pronunciation]  
*Kenkyusha's New English-Japanese Dictionary*. Tokyo: Kenkyusha. 1960<sup>4</sup> [pronunciation], 1980<sup>5</sup>, 2002<sup>6</sup>. [editor]  
*Kenkyusha's New Collegiate English-Japanese Dictionary*. Tokyo: Kenkyusha. 1967<sup>1</sup>, 1968<sup>2</sup>, 1971<sup>3</sup>, 1977<sup>4</sup>, 1985<sup>5</sup>, 1994<sup>6</sup> 2003<sup>7</sup>. [member of the editorial committee]  
*Kenkyusha's Union English-Japanese Dictionary*. Tokyo: Kenkyusha. 1972<sup>1</sup>, 1978<sup>2</sup>. [editor]  
*Kenkyusha's Lighthouse English-Japanese Dictionary*. Tokyo: Kenkyusha. 1984<sup>1</sup>, 1990<sup>2</sup>, 1996<sup>3</sup>, 2002<sup>4</sup>, 2007<sup>5</sup>. [editor]  
*Kenkyusha's Luminous English-Japanese Dictionary*. Tokyo: Kenkyusha. 2001<sup>1</sup>, 2005<sup>2</sup>. [editor]

## In Memoriam Shigeru Takebayashi, 1926-2011

Professor Emeritus Shigeru Takebayashi, phonetician and lexicographer, passed away on 10 March 2011, one day before the devastating earthquake and tsunami hit Japan.

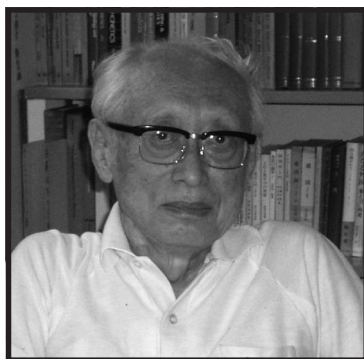
He was born on 22 September 1926 in Asakusa, Tokyo. In 1944, during the Second World War, he was admitted to the English Department of the Tokyo School of Foreign Languages (the present-day Tokyo University of Foreign Studies). Here, English phonetics lessons given by Tamihei Iwasaki fascinated the new student with refreshing insights. After the war, Takebayashi entered the University of Tokyo and enjoyed attending Yoshio Nakano's lectures on Shakespeare and studying linguistics with Shiro Hattori, when he was invited by his former teacher, Iwasaki, to join the team of writers for the Enlarged Second Edition of *Kenkyusha's New English-Japanese Dictionary*. He was the only student among the contributors of this famous dictionary, and later recalled that this experience determined his future course.

When he was a student at the University of Tokyo, Takebayashi suffered from pulmonary tuberculosis and was in and out of hospital throughout his undergraduate and graduate years. After graduation he taught English at the University of Electro-Communications and later moved to the Tokyo University of Foreign Studies, where he taught phonetics and phonology for 23 years until his retirement in 1989.

He was one of the founding members of the Iwasaki Linguistic Circle<sup>1</sup>, which started in 1962 and which will celebrate its jubilee in 2012. As president of this association of nearly 200 members, Professor Takebayashi

was very much looking forward to witnessing next year's publication of the commemorative edition of *Lexicon*, the Circle's journal.

As a phonetician, Takebayashi made immense contributions to the education of English phonetics in Japan by teaching the subject as well as writing textbooks about it, including the widely-read



*Eigo onseigaku nyuumon* [An introduction to English phonetics], Taishukan Shoten, 1982. After his retirement, he wrote the tome *Eigo onseigaku* [English phonetics], Kenkyusha, 1996—a voluminous work covering more or less all aspects of present-day English pronunciation, and comparing it with that of Japanese wherever possible, to accommodate Japanese learners. The book was published in time for his seventieth birthday.

As a lexicographer, his name can be found on the covers of well-known learners' dictionaries published by Kenkyusha of Tokyo, including the Lighthouse titles, which have influenced not only other English-Japanese dictionaries but also bilingual learners' dictionaries of other languages, in and out of Japan. A phonetician, he was always concerned with how to appropriately describe and transcribe contemporary pronunciations of

English spoken in North America and Britain, and how to incorporate the information in dictionaries. He used the 'hooked schwa' symbol [ə̯] to transcribe vowels whose r-sounds are pronounced in North American speech, and was the first to show stress marks on all idioms and phrasal verbs<sup>2</sup>.

Not only was Professor Takebayashi spared from having to witness what happened to Japan the day after his death, he was also spared from feeling the pain of losing his wife, who passed away quietly, it seems, sitting in her chair at breakfast almost exactly 24 hours before her husband died in hospital. Neither knew about the other's (impending) death.

On May 15 the ILC held a party at the Rihga Royal Hotel Tokyo in honor of Professor Takebayashi. 108 ILC members, former students and colleagues gathered to pay tribute to the late professor's life and work.

The professor disliked formalities, loathed war, loved Mozart and was a keen gardener and photographer. The hundreds of his former students will miss him greatly and will always remember him.

### Notes

1. A detailed account of ILC, by Kaoru Akasu, can be found in KDN 15, 2007.
2. *Kenkyusha's Union English-Japanese Dictionary* (1972).

### Hiroko Saito

Tokyo University of Foreign Studies  
qzp11136@nifty.com

A list of selected publications appears on p. 31.