

DICTIONARY News



Ya'acov Levy 1952-2004

Ya'acov Levy passed away on his fifty-second birthday, December 23. The cancer he fought against for two years finally overcame him.

Ya'acov was wise, modest, noble, a man of words – a man of few words, a gifted lexicographer. He usually preferred not to talk or write about his dictionary-making but let it speak for itself. Indeed, he has had a decisive impact upon English/Hebrew lexicography in Israel in the past two decades, and his dictionaries are in practically every home.

Ya'acov Levy died young, before having realized his finest dictionary dreams. His legacy shall live on.

Ya'acov Levy | Dictionary pages

- 5 Menahem ben Saruq's *Mahberet*: The first Hebrew-Hebrew dictionary | **Aharon Maman**
 - 10 Review of *Wörterbuch Deutsch-Hebräisch Philosophische, wissenschaftliche und technische Termini*
| **Ora R. Schwarzwald**
 - 12 Towards Hebrew FrameNet | **Miriam R.L. Petruck**
 - 13 *Contemporary English-Russian Dictionary*: A New Type of Dictionary in Russia | **Galina Zakharova**
 - 14 A Large-Scale Lexical Database of Danish for Language Technology Applications and Other Purposes
| **Anna Braasch**
 - 17 If dictionaries are free, who will buy them? | **Adam Kilgarriff**
 - 20 Dictionaries and productive tasks in a foreign language | **Paul Bogaards**
 - 23 Longman Słownik współczesny angielsko-polski polsko-angielski: The First Active
Bilingual Dictionary for Polish Learners of English | **Arleta Adamska-Sałaciak**
 - 27 Bilingual Dictionaries for Learners | **Martyn Back**
 - 30 Power to the Learner: An Approach towards Pedagogically-Oriented Bilingual Dictionaries
| **Wolfgang Worsch**
 - 33 Congress Organisation: The academic's burden | **Geoffrey Williams**
 - 34 The Fifteenth DSNA Biennial Meeting, Boston 2005 | **David Jost**
 - 35 The Fourth ASIALEX International Congress, Singapore 2005 | **Anne Pakir**
 - 37 Les Journées Allemandes des Dictionnaires, Klingenberg 2004, 2006 | **Michaela Heinz**
 - 40 The Twelfth EURALEX International Congress, Turin 2006 | **Carla Marelo**
- K DICTIONARIES** | Recent titles



© 2005 All rights reserved.

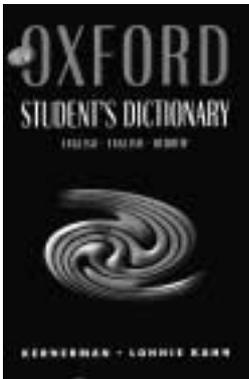
Editor | **Ilan J. Kernerman**

Graphic Design | **Studio Orna Cohen**

Thank you | **Lionel Kernerman, Nili Sadeh, Merav Kernerman**

K DICTIONARIES LTD
Nahum 10 Tel Aviv 63503 Israel
Tel: 972-3-5468102
Fax: 972-3-5468103
kdn@kictionaries.com
http://kictionaries.com

Ya'acov Levy assisted Joseph Reif in editing the first semi-bilingual Hebrew edition of *Oxford Student's Dictionary of Current English* published in 1986. The two worked together on the second edition published in 1993, and Raphael Gefen joined Ya'acov Levy on the third edition published in 2003.



**OXFORD STUDENT'S
DICTIONARY FOR
HEBREW SPEAKERS**

First Edition 1986

A.S. Hornby

with the assistance of:

Christina Ruse

Hebrew editor: Joseph A.

Reif

with the assistance of:

Ya'acov Levy

824 pages

ISBN 965-307-000-2 /

965-307-003-7

Second Edition 1993

English: Christina Ruse

Hebrew: Joseph A. Reif

and Ya'acov Levy

904 pages

ISBN 965-307-021-5

Third Edition 2003

Editor: Raphael Gefen

Translator: Ya'acov Levy

984 pages

ISBN 965-307-034-7

Kernerman Publishing

and Lonnie Kahn

Tel Aviv

middle

534

middle² adj. (only before a noun) in the middle; I wear my ring on my middle finger. אמצעי

middle 'age noun [U] the time when you are about 40 to 50 years old; a woman in early middle age מיל המעבר

middle-aged adj. a middle-aged man גיל המעבר

the Middle 'Ages noun [pl] the period of European history from about 1,100 to 1,500 AD שני הביניים

the middle 'class noun [singular] (also **the middle 'classes** [pl]) the group of people in a society who are neither very rich nor very poor and that includes professional and business people המעמד הבינוני

middle 'class adj. They're middle class. + a middle-class background. Compare the upper class and the working class. טן המעמד הבינוני; בורגני

middle 'ear noun [singular] the central part of your ear behind your eardrum (= the thin piece of skin that allows you to hear sound) Look at inner ear. האוזן המימית

the Middle 'East noun [singular] an area that covers SW Asia and NE Africa Look at the Far East

middleman noun [C] (pl. -men) 1 a person or company who buys goods from the company that makes them and then sells them to somebody else סוחר

2 a person who helps to arrange things between two people who do not want to meet each other סוחר

middle-of-the-'road adj. (used about people, policies, etc.) not extreme; acceptable to most people לא קיצוני; של אמצע הדרך

'middle school noun [C] (pl. -schools) a school for children aged between 9 and 13 בית-ספר יסודי-ליניארי

midge noun [C] a very small flying insect that can bite people תגית מרימסטה; יכחוש

midget noun [C] a very small person הוסיף Be careful. Some people find this word offensive. גמד; ננס

the Midlands noun [singular, with singular or pl verb] the central part of England around Birmingham and Nottingham מחוזות המרכז

midnight noun [U] twelve o'clock at night; They left the party at midnight. + The clock struck midnight. Look at midday. חצות

midriff noun [C] the part of your body between your chest and your waist. סרעפת; בטן

midst noun [U] the middle of something, among a group of people or things: The country is in the midst of a recession. + They realized with a shock that there was an enemy in their midst. אמצע; בקרב; בתוך

midsummer noun [U] the time around the middle of summer; a beautiful midsummer's evening (שנה) אמצע הקיץ

midway adj., adv. in the middle of a period of time or between two places: The village lies midway between two large towns. (מרחק) half-way בתצי הדרך

midweek noun [U] the middle of the week (= Tuesday, Wednesday and Thursday) חמצע השבוע

midweek adv.: If you travel midweek it will be less

crowded. באמצע השבוע

the Midwest noun [singular] the northern central part of the US המערב התיכון

midwife noun [C] (pl. midwives) a person who has been trained to help women give birth to babies מילדת

midwifery noun [U] the work of a midwife מילדות

midwinter noun [U] the time around the middle of winter (שנה) אמצע החורף

might¹ modal verb (negative **might not**; short form **mightn't**) 1 used for saying that something is possible; 'Where's Vinny?' 'He might be upstairs.' + I think I might have forgotten the tickets. + She might not come if she's very busy. יעקב; ככל להיות ש...; עשוי; עלול

2 (BrE formal) used to ask for something or suggest something very politely; I wonder if I might go home half an hour early today? ואם אפשר ל...; יכול

3 used as the form of 'may' when you report what somebody has said; He said he might be late (= his words were, I may be late). עשוי; עלול

[may/might as well (do something)] → well! you, etc. **might do something** used when you are angry to say what somebody could or should have done; They might at least have phoned if they're not coming. (לחצות) היית...

I might have known used for saying that you are not surprised that something has happened; I might have known he wouldn't help. אי לא הייתי יודע

might² noun [U] (formal) great strength or power; I pushed with all my might, but the rock did not move. סת עוצמה

mighty¹ adj. (mightier; mightiest) very strong or powerful עקב; רב עוצמה; עוצם

mighty² adj. (AmE informal) very; That's mighty kind of you. סאה; טרא

migraine noun [C,U] very bad pain in your head that makes you feel sick; a severe headache סיקחה; דרעה

migrant noun [C] 1 a person who moves from one place to place looking for work; migrant workers אועל; מדינאי

2 a bird or an animal that moves from one place to another according to the season יעור; סודת; בער-חיים; סודד

migrate verb [I] 1 (used about animals and birds) to travel from one part of the world to another at the same time every year לנדוד

2 (used about a large number of people) to go and live and work in another place: Many country people were forced to migrate to the cities to look for work. Look at emigrate. לנדוד; ריכור

migration noun [C,U] פרוק; הגירה

migratory adj. (used about animals and birds) travelling from one part of the world to another at the same time every year נודד

mike (informal) = microphone קישאן

milage = mileage

mild adj. 1 not strong; not very bad; a mild winter + a mild punishment עדין; לא קשה; קל; סתון

After preparing a semi-bilingual Hebrew edition of *Oxford Elementary Learner's Dictionary of English*, published in 1991, Ya'acov Levy wrote *Passport Junior English Learner's Dictionary*, edited by Raphael Gefen. The Hebrew edition appeared in 1996 and has since been followed by versions for Chinese, Czech, Danish, Estonian, Greek, Italian, Latvian, Lithuanian, Romanian, Slovak, Slovenian, Spanish, Thai, Turkish and more to come.



**PASSPORT
ENGLISH HEBREW
LEARNER'S
DICTIONARY**
Lexicographer and
Translator: Ya'acov Levy
Editor: Raphael Gefen
509 + 144 pages
ISBN 965-307-029-0
Kernerman Publishing
and Lonnie Kahn
Tel Aviv 1996

plague 288

plague *n.*
 1 פְּשִׁיטָה; דָּבָר
 The plague killed thousands of people.
 2 סָבָה (יִמְנִיחִי)
 a plague of mosquitoes

plain¹ *adj.*
 1 פְּשִׁיט, בְּבִגָּה אֶחָד
 a plain black dress
 2 פְּשִׁיט, בְּגֵי דוֹמָמָה
 a plain carpet (= a carpet without designs)
 3 פְּשִׁיט (בְּגֵי קִישוּיָיִם); בֶּטָעִים טְבִיעִי
 plain food ◊ plain yoghurt
 4 (נִיזָר) בְּגֵי שׂוֹרָת
 plain paper
 5 (= clear) פְּשִׁיט וְכָרוּר
 to speak in plain English
 6 (= obvious; clear) כָּרוּר; גָּלוּי לְעֵינַי
 It was plain that he was angry.
 7 לֹא יָפָה (אִישׁוֹת, מַעֲרָה)
 She was a plain girl.
in plain clothes (בְּבִגָּה אֶנְדְרוּזִי, בְּגֵלֶט, שׂוֹטֵר)
 a detective in plain clothes
plain chocolate *n.* (Brit.) שׂוֹקוֹלֵד טָרוּר
 + American English: **dark chocolate**

plain² *n.* מִישׁוּר (נִישׁוֹנְרִפִּית)
plainly *adv.* (= clearly) בְּנִירוּר
 We could see him plainly.

P **plait**¹ *n.* בֶּטָח
 She wears her hair in a plait.
 + American English: **braid**



plait (Brit.)/braid (Amer.)

plait² *v.* (plaits, plaiting, plaited) (Brit.)
 לְקַלֵּץ (נִצְמוֹת, תְּכֵל, וְכֵדִי)
 She plaited her hair.
 + American English: **braid**

plan¹ *n.*
 1 תּוֹכְנִית
 What are your plans for tomorrow? ◊ to make plans for the future ◊ Everything went according to plan.
 2 טָפָה (שֶׁל עִיר); תַּרְשִׁיִּים; תּוֹכְנִית
 a street plan (= map) of the city ◊ a plan for a new office building

plan² *v.* (plans, planning, planned)
 1 לְתַכְנֵן
 He planned the military operation.
 2 לְתַכְנֵן / לְתַכְנֵן (לְעֻשֵׂת מַשׁוּם)
 I'm planning to visit them in Paris.
planned *adj.* תּוֹכְנֵן
 It wasn't planned.
planning *n.* תַּכְנִין
 The planning took a month. +

plane *n.* טַסָּס
 What time does the plane land? ◊ The plane will take off in 30 minutes. ◊ It's time to board (= get on) the plane.
 + See **aeroplane**; **airplane**; **aircraft**

plane crash *n.* חֲתִירָסוֹת טַסָּס
 Six people died in the plane crash.

planet *n.* פְּלָנֵטָה
 the planet Earth

plank *n.* קֶרֶב שִׁטוֹחַ (לְעַבְדוֹת בְּנִיָּה, וְכֵדִי)

plant¹ *n.*
 1 צָמַח; שְׂתִיל
 to water the plants ◊ tomato plants
 2 סַפְּגָל
 a chemical plant

plant² *v.* (plants, planting, planted)
 לְשַׂתֵּל; לְסַפֵּץ
 to plant flowers/trees

plantation *n.*
 טַקְעַת (שֶׁל תֵּה, כּוֹמֹת, סֶבֶק, וְכֵדִי)
 a cotton/sugar plantation
 + טַקְעַת שֶׁל עֵצֵי פְּרִי = **orchard**

plaster *n.*
 1 קִיט
 to cover the walls with plaster
 2 נְבֶסֶט
 His leg was in plaster (= Amer. in a cast) for four weeks.
 3 (= **sticking plaster**) (Brit.)
 פְּלַסְטֵר; אֶגְדַּת מִיִּדְבָּק
 Put a plaster on the cut.
 + American English: **Band-aid**®

plastic¹ *adj.* פְּלַסְטִיקִי; סַפְּלַסְטִיקִי
 plastic cups
plastic bag *n.* שְׂקִית פְּלַסְטִיקִי / נִיילוֹן
plastic² *n.* פְּלַסְטִיק
 The flowers are made of plastic. ◊ sheets of plastic

The bilingual *Oxford English-Hebrew Hebrew-English Dictionary*, published in 1995, was a simple translation dictionary that Ya'acov Levy would have turned into a comprehensive bilingual one. A reduced but updated edition appeared in 2003, *Oxford Pocket Dictionary English-Hebrew Hebrew-English*, and an extracted *Oxford Mini Dictionary English Hebrew* is just being published.



**OXFORD
ENGLISH-HEBREW
HEBREW-ENGLISH
DICTIONARY**

Editor: Ya'acov Levy
600 + 310 pages
ISBN 965-307-026-6 /
965-307-027-4
Kernerman Publishing
and Lonnie Kahn
Tel Aviv 1995



**OXFORD POCKET
DICTIONARY
ENGLISH-HEBREW
HEBREW-ENGLISH**

Ya'acov Levy
472 + 253 pages
ISBN 965-307-033-9
Kernerman Publishing
and Lonnie Kahn
Tel Aviv 2002

topsy-turvy

522

topsy-turvy adj. adv.	הפוך; בבלבול; סתוּר (כִּירוּה) כִּי אֵיךְ	total adj.	מוחלט; קלא; כולל
tor n.	(כִּירוּה) כִּי אֵיךְ	total n.	סה; סכום כולל
torch n.	נבִּיעָה קטָנוֹת; מִסְקֵה סִלְעוֹת	total n.	לְסִמָּה; לְחִסְקָם בִּי; לְהַסְדִּיר לְסִמָּה
torchlight n.	פְּנִים-כִּים; לְפִיד; מִנְעֵר רִיטוֹן	totalitarian adj.	מִדְּסִלִּיקָרִי
tore v. past of tear	לְקַצֵּית (בִּדְרוֹן)	totalitarianism n.	מִדְּסִלִּיקָרִיּוֹת
tormentor n.	אֲרִיבָּס; אֲרִיבָּיִדִים	totality n.	כֻּלְמוֹת; מְכֻלָּה; קְרִימָל
torment n.	טִרְדָּאוֹת; חֹסֶם שְׂדוּחִים	totalizer n.	מְכַסֵּס חִשׁוֹב חִשׁוּמִים
- be in torment	עִשָּׂה יִסּוּרִים	total loss n.	אֲבֵדוֹן מוֹחֵלֵט (בְּבִישׁוֹת)
torment n.	לְהַתְּעוֹת; לְהַחֲזִיק בְּכַאֲבָם	totally adv.	לְסִמָּה
tormentor n.	לִיבָּר; לְהַבְרִיק לִי	tote v.	לְשִׂאת נֶשֶׂאָה; לְסוּבֵב
torn v. 3rd form of tear	מִיבָּר; מִיבָּר	tote n.	מְכַסֵּס חִשׁוֹב חִשׁוּמִים
tornado n.	סוּפֵט מְרִדָּד	tote bag n.	סֵל קְנוּיָה; אֶרֶקֶת כְּדוּלָה; חִיק כֶּתָף
torpedo n. v.	טוֹרְפֵדוֹ; לְקִרְפֵד	totem n.	טוֹטֵם; סֵמֶל שִׁבְטֵי וְאִנְדִּיָּאָנִים
torpid adj.	עֲבֵלָנוּ; טִרְפָּה; רָדוּם	totem pole n.	יָמוֹר; קֶצֶם נֶשֶׂרֶץ
torpor, torpidity n.	עֲבֵלָנוּת; אִטְרָיוֹת	totter v.	עָבַד טוֹטֵם
torque n.	אֵינְדִּיָּאוֹת; כִּחוֹם	tottery adj.	לְהִלֵּךְ בְּאִי-יָצִיבוֹת; לְקַטֵּם קְפוּסֵי; לְמוֹט
torrent n.	בְּכִימוֹת; מַלְטוֹת מְהִימוֹת	touch n.	לִישׁוֹל; לְהַתְּעוֹדָד (בְּעִסְקָד לִישׁוֹל)
torrential			לֹא יָצִיב; שֶׁמֶד לִישׁוֹל
torrid			טוֹחַ וְטוֹחַ
torrid adj.			
torision n.			
torso n.			
tort n.			
tortilla n.			
tortoise n.			
tortoisesh			
tortuous			
tortuously			
torture n.			
torturer n.			
Tory adj.			
Tory n.			
toss n.			
toss n.			
- not give a			
toss n.			
- toss about			
- toss and			
- toss and			
- toss back			
- toss off			
- toss out			
- toss up			
toss-up n.			
tot n.			
tot n. (tote			

138

תְּרִיקָה

with traffic lights	תְּרִיקָרִי ת.	mortar (military)	תְּרִיקָה ג.
exalted, uplifted,	תְּרִיקוֹם ת. (תְּרִיקוֹמָה)	daisy	תְּרִיקָתִי ג. (תְּרִיקָתִים)
elevated	תְּרִיקוֹן ת. (תְּרִיקוֹנָה)	margarine	תְּרִיקָתִי ג.
restrained;	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	thrilling, exciting,	תְּרִיקָתִי ג. (תְּרִיקָתִים)
checked	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	moving	תְּרִיקָתִי ג. (תְּרִיקָתִים)
sprayed	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	feeling	תְּרִיקָתִי ג. (תְּרִיקָתִים)
crushed,	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	to rebel, revolt	תְּרִיקָתִי ג. (תְּרִיקָתִים)
smashed	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	(against)	תְּרִיקָתִי ג. (תְּרִיקָתִים)
broken line	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	rebellion, revolt, mutiny	תְּרִיקָתִי ג. (תְּרִיקָתִים)
upholstered,	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	apoplectic, narcotic;	תְּרִיקָתִי ג. (תְּרִיקָתִים)
padded	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	anesthetist	תְּרִיקָתִי ג. (תְּרִיקָתִים)
shabby, tattered	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	rebellious person	תְּרִיקָתִי ג. (תְּרִיקָתִים)
race	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	rebelliousness,	תְּרִיקָתִי ג. (תְּרִיקָתִים)
satisfied, pleased,	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	recalcitrance	תְּרִיקָתִי ג. (תְּרִיקָתִים)
contented	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	rebellious	תְּרִיקָתִי ג. (תְּרִיקָתִים)
running	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	chase, pursuit	תְּרִיקָתִי ג. (תְּרִיקָתִים)
- in the course of time	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	bile; gall-bladder	תְּרִיקָתִי ג. (תְּרִיקָתִים)
paved, tiled (floor)	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	brehtaking, splendid,	תְּרִיקָתִי ג. (תְּרִיקָתִים)
Moroccan	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	spectacular, glorious	תְּרִיקָתִי ג. (תְּרִיקָתִים)
Morocco	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	interviewee	תְּרִיקָתִי ג. (תְּרִיקָתִים)
emptied, empty	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	many-	תְּרִיקָתִי ג. (תְּרִיקָתִים)
Moroccan	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	numerous, multiple, multi-	תְּרִיקָתִי ג. (תְּרִיקָתִים)
bitter herbs (for Passover)	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	quadrilateral, quadrangle,	תְּרִיקָתִי ג. (תְּרִיקָתִים)
(usually lettuce)	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	square	תְּרִיקָתִי ג. (תְּרִיקָתִים)
negligent,	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	angry	תְּרִיקָתִי ג. (תְּרִיקָתִים)
careless, sloppy, slack	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	excited	תְּרִיקָתִי ג. (תְּרִיקָתִים)
wicked, evil,	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	furnished	תְּרִיקָתִי ג. (תְּרִיקָתִים)
vicious	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	sage (plant); salvia	תְּרִיקָתִי ג. (תְּרִיקָתִים)
impoverished	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	thirst-quenching	תְּרִיקָתִי ג. (תְּרִיקָתִים)
networked, wired	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	roomy, spacious	תְּרִיקָתִי ג. (תְּרִיקָתִים)
authority, discipline	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	spread, smeared	תְּרִיקָתִי ג. (תְּרִיקָתִים)
confined	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	remote, distant,	תְּרִיקָתִי ג. (תְּרִיקָתִים)
(military); tied; riveted, glued (to)	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	far-off	תְּרִיקָתִי ג. (תְּרִיקָתִים)
gutter (of roof), drainpipe	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	tom apart,	תְּרִיקָתִי ג. (תְּרִיקָתִים)
to spread (butter, etc.); smear; rub (into skin)	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	mangled, mutilated; retouched	תְּרִיקָתִי ג. (תְּרִיקָתִים)
open space, expanse	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	concentrated	תְּרִיקָתִי ג. (תְּרִיקָתִים)
living space	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	softened; mollified;	תְּרִיקָתִי ג. (תְּרִיקָתִים)
leeway	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	toned down	תְּרִיקָתִי ג. (תְּרִיקָתִים)
room for maneuver	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	top, height; sky, heaven	תְּרִיקָתִי ג. (תְּרִיקָתִים)
spatial, regional	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	in the heavens	תְּרִיקָתִי ג. (תְּרִיקָתִים)
far-reaching	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	cheated,	תְּרִיקָתִי ג. (תְּרִיקָתִים)
far-sighted	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	deceived	תְּרִיקָתִי ג. (תְּרִיקָתִים)
bath	תְּרִיקוֹס ת. (תְּרִיקוֹסָה)	hinted, implied	תְּרִיקָתִי ג. (תְּרִיקָתִים)

Menahem ben Saruq's *Mahberet*: The first Hebrew-Hebrew dictionary

Aharon Maman

Menahem ben Jacob ben Saruq was born at the beginning of the tenth century in Tortosa, Spain, and moved in his youth to Córdoba, the capital of Andalusia and the center of the Umayyad rule and of Jewish learning in Spain in those days. He served for a time as the secretary of Isaac ibn Shaprut, and then of his son, Hisday ibn Shaprut (915-970), the leader of the Jews of Andalusia, who served in the court of the Caliph 'Abd al-Rahmân III (912-961).¹ With Hisday's encouragement, Menahem² compiled the *Mahberet*,³ the first Hebrew-Hebrew dictionary, and perhaps the first systematic Hebrew dictionary of any kind. Close to the time when the *Mahberet* was being compiled, a comprehensive Hebrew-Arabic dictionary of biblical terms was written in Jerusalem by the Karaite scholar David ben Abraham Alfâsi,⁴ but it is not clear whether the two were acquainted with each other's work. Despite a certain resemblance between these dictionaries,⁵ their compilers may not have copied from each other but instead drew from common sources – for example, both have similar definitions for the entries טוטפות ([*totafot*] *phylacteries*), גדי ([*gdi*] *kid*, in the context of גדי בחלב אמו) and others.⁶

Prior to the compilation of these dictionaries, works of a lexicographic nature had been compiled, usually bilingual glossaries for the Bible or parts of it or for particular tracts of the Mishnah⁷ and the Talmud⁸. Also compiled were *Aruch* by Zemah ben Paltoiy Gaon in Pumbedita (Babylon), around 732 CE,⁹ but it was lost, and Rav Sa'adiah Gaon's *Agron*, but it was limited in its scope, providing only a list of about 1,000 nouns recommended to poets.

Mahberet Menahem is, as noted above, the first systematic dictionary of Hebrew. It contains all of the vocabulary of the entire Bible, providing some 2,500 roots, arranged in approximately 8,000 lexemes, in alphabetical order by root, according to Menahem's perception of the root. It is systematic in that the entries are arranged in alphabetical order with the root as a main entry and its derivatives as sub-entries. The alphabetization is usually preserved in relation to the first two letters of the root, but not always in relation to the third. For example, the root חר"ף [*h-r-f*] comes after the root חר"ץ [*h-r-ts*].

Roots comprising four letters or more appear among the regular roots (and are

not assigned a separate section, as was done by Menahem's successors, Rabbi Jonah ibn Janah and Rabbi David Qimhi, in their dictionaries). For example, חרצב [*hartsab*] and חרצן [*hartsan*] appear immediately after the root חר"ץ [*h-r-ts*] (and not at the end of all of the entries for the letter *het*). The root is presented with all of its derivatives below it, verbs and nouns alike, and this example was followed by lexicographers throughout the Middle Ages, until modern times, when nouns began to be listed according to their initial letter even if it is not radical. Eliezer Ben Yehuda¹⁰ seems to have been the first Hebrew lexicographer in modern times to separate nouns from verbs. For example, he listed the root כל"ל [*k-l-l*] including its verbal conjugations in the *kaph* section, but placed the noun מכללה ([*mikhhlala*] *college*), from the same root, in the *mem* section. In more recent years, other editing systems, such as that used in *Milon ha-Hoveh*,¹¹ have been developed.

The original name of *Mahberet Menahem* was, apparently, *Sefer Pitronim* (*Book of Interpretations*) and Menahem's pupils referred to it by that name. But in the body of the book, Menahem calls the list of entries that begin with the letter *aleph* – *Mahberet Aleph*, the *bet* section – *Mahberet Bet*, and so on, for a total of 22 *mahbarot*. Thus the name of the entire dictionary became *Mahberet* very close to the time of its completion, and Rashi (Rabbi Shlomo Yitzhaki, 1040-1105) already refers to it by that name.¹² The name *Mahberet* (from חב"ר [*h-b-r*] *to unite, associate, link up*) reflects contemporary trends in names of dictionaries; it is similar to *Agron* (from אג"ר [*'-g-r*] *to compile, collect*), the name of the dictionaries compiled by Rav Sa'adiah Gaon and Alfâsi; the name used for Arabic dictionaries, *Jâmi'*, literally, a collection of words; and the modern Hebrew expression *otzar millim*, the counterpart of the Latin term *thesaurus*. In modern times, due to the tendency to prefer one word to a phrase, the word *millon* was devised, apparently by Eliezer Ben Yehuda (in 1880), on the basis of the word *millah* (*word*), as the equivalent for *dictionary*.

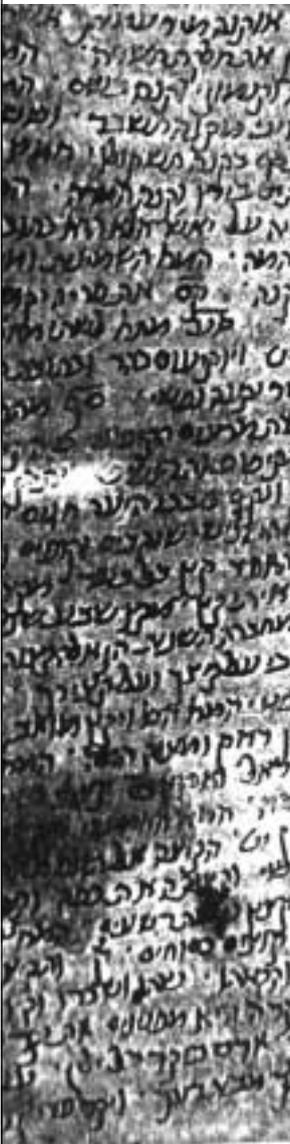
Menahem devoted his dictionary solely to words in the Bible, whose language he considered exemplary, and though he did not compile a grammar book at the same



Aharon Maman was born in Morocco in 1947, immigrated to Israel in 1964, and studied Hebrew and Arabic at the Hebrew University, where he is Professor. In the years 1992-93 and 1997-98 he was Chair of the Department of Hebrew at HU. He is director of the Jewish Oral Traditions Research Center and the Hebrew Linguistic Literature Project, member of the Academy of the Hebrew Language and of academic committees of research centers and journals. He specializes in Medieval Hebrew Philology. He won the Fulbright Program Fellowship by the Council for International Exchange of Scholars in 1993-94, was visiting scholar at the University of London, taught at Yeshiva University in NY and carried out a long-term research on manuscripts preserved in the Jewish Theological Seminary Library. He also published on Karaite Hebrew and grammar, lexicography, and traditional Magrebian Hebrew and Arabic, and is co-editor (with S. Fassberg) of *Mehqarim be-Lashon* (Hebrew Studies) and *Massorot* (Hebrew Traditions).
maman@h2.hum.huji.ac.il



Mahberet Menahem page
(above) and extract (below)
from the Cairo Genizah
ms # JTSLA, ENA 3758 1a



time, the dictionary also includes numerous discussions on grammar, in excursuses or in brief remarks in a few entries. Apart from these explicit remarks, the method used in editing the dictionary is instructive with regard to his grammatical approach, at least in the area of root theory. The theory underlying *Mahberet Menahem* advocates an abstract root that can be composed of one consonant, or of two consonants or more, a theory that also governed the works of Alfāsi, his contemporary, and Judah ben Quraysh, who belonged to the previous generation.¹³

For example, the main entry שׁב [sh-b] includes roots that have been considered as roots in their own right ever since the time of Judah Hayyuj (Fez, Morocco - Córdoba, ca. 1000): נשׁב [n-sh-b], ישׁב [y-sh-b], שׁוׁב [sh-w-b], שׁבׁי [sh-b-y], שׁבׁב [sh-b-b].¹⁴ According to the method used by Menahem and his contemporaries, and the method used by some of their predecessors, every letter that is not used in every inflection of a word, such as *yod*, *waw* and *nun* in these roots, is not radical. After all, for the meaning ישיבה ([yeshiva] sitting) we say ישבתי ([yashviti] I sat) in the past tense, but אֶשֶׁב ([eshev] I'll sit) in the future tense and שֵׁבַת ([shevet] sit) in the infinitive, omitting the initial *yod*. For שׁיבה ([shiva] returning), we say אשוב ([ashuv] I'll return) in the future tense but שבתי ([shavti] I returned) in the past tense, omitting the *waw*. And as for שׁיבׁי ([shivya] captivity), the Bible contains the statement וַיִּשְׁבּוּ מִמֶּנּוּ שְׁבִי ([vayishbu...], Num. 21:1), omitting the final *yod*. Menahem classified as radical only a letter that exists in every inflection. This approach may have been influenced by the concept of *essence* and *accident* in Aristotelian philosophy, according to which only an attribute that is a constant is an essence and attributes that are variable are merely accidents, which are not fundamental to the definition of the “nature” of the object in question.

But this is not to suggest that Menahem attributed the same meaning to שׁביה, שׁביה, שׁביה and נשיבה ([neshiva] blowing). In the entry שׁב [sh-b], he arranged them in “departments” (מהלקות [mahlaqot]), that is, sub-entries, with each sub-entry containing a different definition. He may have considered them homonymic roots, but it is possible that he was thinking mainly of the root as an organizing entity. And yet, there are some words in the dictionary that are treated as polysemic and deriving from one root, whereas according to his method they should be treated as homonymic.¹⁵

In any event, all of the roots that modern methodology considers ‘defective’ or geminate are, according to Menahem’s

system, biliteral. For example, on the basis of וַיִּט מִשֶּׁה אֶת יָדוֹ ([vayet...], Exod. 9:22), he determined that the root of the verb was *tet* alone, and not נטׁה [n-t-h] or נטׁי [n-t-y] as Hayyuj had determined, a generation after Menahem.

The monoliteral roots were not placed in their expected position in Menahem’s dictionary, at the beginning of the entries for each letter that serves as the initial letter of a monoliteral root according to his system, but in a concentrated discussion, apparently because they are very few. This discussion is located in the first opportune place – at the beginning of the entries for the letter *bet*, the first letter in the alphabet that serves as the initial letter of a monoliteral root.¹⁶ And since he is already explaining this particular type of root, he launches into a discussion of root theory in general.

Mahberet Menahem is a compact dictionary, and most of its entries are very short. Usually the entry is structured as follows: it comprises a main entry, and if this main entry has a few definitions, they are presented as “departments”. Each department includes the relevant words, a few examples from the Bible and brief definitions, usually by means of a synonym in a general mode. When summing up the meaning of several entries that he views as belonging to the same semantic family, Menahem uses the following formula: ענייני המה X (they are from the meaning of X). For example, all of the citations presented in the entry בׁך [b-kh], department 1, were defined ענייני הנעה המה – i.e., “they are from the meaning of movement”.

The *Mahberet* contains some embryonic entries that offer only an example from the Bible and no definition, on the assumption that the educated reader will draw the meaning from the example. There are also entries with an alternative definition in the form of במשמעו, i.e., “in its literal meaning” or “as is known”,¹⁷ but without a concrete definition. Hence, the dictionary was not written for laymen but rather for the learned who already had a good basic knowledge of the Bible.

The overwhelming majority of entries contain no grammatical, etymological or semantic discussions (for example, no antonyms are listed); no clues to the declension of the word, neither of the verbs deriving from it nor of the nouns; and, no comments on syntax. The *Mahberet* is mainly a dictionary for the *peshat* (literal meaning) philological significance of the word.¹⁸

The entries are in most cases spelled as they appear in the Bible; the plene is plene and the deficient is deficient,

and the entries and the citations are presented without vocalization. As was the custom of language scholars of the time, no references are provided for the biblical citations, on the assumption that the context of the citation is a sufficient indication of its location in the Bible (and in a few cases also in the Mishnah). However, modern editors (see below) added references as needed.

The *Mahberet* is a dictionary devoted to the entire Bible, including the Aramaic words in the biblical text. Menahem does not assign a special section to the Aramaic words, as was done by Rabbi David Qimhi (Provence, thirteenth century) in his *Sefer ha-Shorashim (Book of the Roots)* and as is customary in modern biblical dictionaries. Instead, he integrated them among the Hebrew entries, as if it were all one language. For example, the Hebrew citation *וַיִּמְלֶךְ לְבִי עָלַי* ([*vayimalekh...*] Neh. 5:7) and the Aramaic citation *מְלִכֵי יִשְׂפָר עֲלֶיךָ* ([*milki...*] Dan. 4:24) appear in one entry in the second department of the root מל"ך [m-l-kh].

It was not that Menahem did not distinguish between the two languages, but his approach to the dictionary was more philological than linguistic, and since the dictionary concerns a particular book – the Bible, in its entirety – he did not separate the two languages used in it. In this sense, separation would be merely a technical matter.

Even though the *Mahberet* includes all of the words in the Bible, it does not include names, neither of people nor of places. (Such names were usually dealt with in concordances, encyclopaedias and lexicons for the Bible, but not in dictionaries.)

Menahem was a methodical scholar who was consistent to an extreme in his opinions, and had a critical, scientific way of looking at things. Since root theory was a focus of great interest in this period when the philological interpretation of the Bible was coming into its own, Menahem and the other members of his generation considered the status of each letter in a word – that is, whether it was radical or servile. In one of the early excursuses in the *Mahberet*, in the entry אב"ה [-b-h], Menahem leveled instructive criticism at Judah ben Quraysh. Ibn Quraysh¹⁹ explained the unique occurrence of this root in the phrase אֶבְחַת הַרְבַּ [‘*ivhat herev*] Ezek. 21:20; *slaughter, massacre of the sword*) as if it were deriving from the cognate אבַעַת הַרְבַּ [‘*iv'at herev*], that is, from בַּעַתָּה [‘*be'ata*] *trouble, fear*; the word אֶבְחַת [‘*iv'at*] is, of course, an invention of Ibn Quraysh). Menahem contended

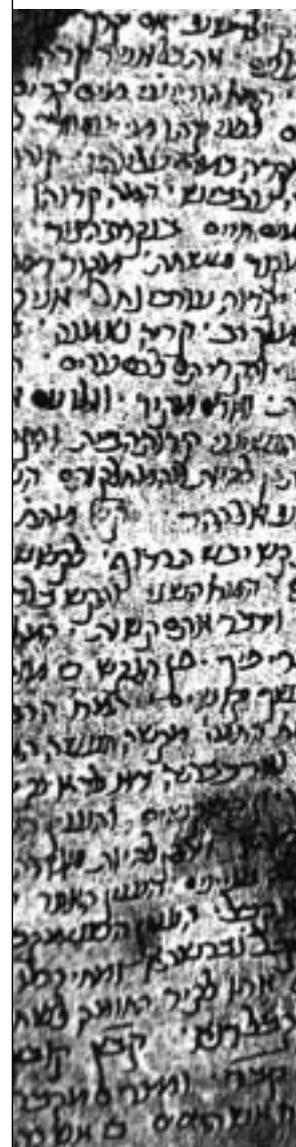
that Ibn Quraysh had thereby committed several grammatical transgressions. First, he had replaced the letter *het* with *'ayin*. Menahem vehemently objected to letter changes and considered such an act to be arbitrary, irregular, unpredictable and ungrammatical, because if we replace one letter with another, the order of the language will be destroyed. The second transgression was that Ibn Quraysh omitted the letter *aleph* in אבַעַת, because if the root is בַּעַת/בַּחַת [b-'t/b-h-t], the *aleph* of אבַחַת is unnecessary, while Menahem considers it radical. The third was that he had radicalized the *taw* in אבַחַת, whereas, according to Menahem's system, it is only the feminine morpheme.

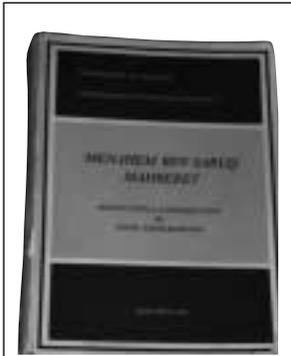
Menahem, too, interprets the word אבַחַת as אֵימָה ([‘*eima*] *great fear*) but although there is no difference between the two philologists concerning the semantics of the word, they differ in how this conclusion was reached and in their grammatical approaches. Menahem's criticism implies that if there is no convincing proof for determining the root, it must be put in its proper perspective: since most *taws* at the end of words were meant to indicate the feminine gender, this can also be assumed with regard to אבַחַת, especially since it is part of a construct phrase. Therefore, he determined that its root was אב"ה [-b-h]. And since this word is unique in the Bible (*hapax legomenon*) and there is nothing with which to compare it, we have no choice but to determine its meaning on the basis of the context alone.

As soon as the *Mahberet* was published, it met with strong reactions, first in Spain and subsequently in France and the East as well. The poet and philologist Dunash ben Labrat, Menahem's contemporary and fellow Córdoba, wrote a book in which he presented 180 remarks (or, as Dunash put it, *responses, objections* תשובות [*tshuvot*]) challenging many of Menahem's decisions, both in general and with regard to individual points. For example, he claimed that there was no need to seek a root for a particle such as מה ([*ma*] *what*) and ruled that only a word that had a grammatical declension could have a root. Dunash also rejected Menahem's determination that יְכַרְסֵמְנָה ([*yekharsmenna*] Ps. 80:14) is an acronym of יְכַרְסֵמְנָה ([*yekhars mena*] itself derived from יְמַלֵא כְרִסוֹ מִמְנָה ([*yemalle kreso mimmena*]); *He [= the swine, which symbolizes the enemies of the Jewish people] would fill its belly from it [= from the vine, which symbolizes the Jewish people]*);²⁰ he claimed that it was only a quadrilateral verb equal to שְׂדֵה... [ש]קְרַסְמוּהָ ([*sade...[she]qirsemuha nemalim*] *a field nibbled by ants*; Mishnah, Pea 2:



Mahberet Menahem page (above) and extract (below) from the Cairo Genizah ms # JTSLA, ENA 3758 1b





*Menahem ben Saruq
Mahberet*
edición crítica
e introducción de Angel
Sáenz-Badillos
Universidad de Granada e
Universidad de Salamanca
Granada 1986



*Tešubot de Dunaš ben
Labrat*
edición crítica e traducción
española de Angel Sáenz-
Badillos
Universidad de Granada e
Universidad de Salamanca
Granada 1980

7), with a replacement of the *qof* with the *kaph*. But Menahem could not agree with the idea of letter changes, even though Rav Sa'adiah Gaon and other language scholars who preceded Menahem ruled that letters could be assumed to have been changed in certain cases. In general, it can be said that we are touching upon a major question here: Can we assume that phonetics sometimes gnawed away at etymology and morphology? Menahem utterly opposed the system of letters' substitution, because he maintained that it was necessary to leave each letter as it was and interpret it as it was, and not destroy the foundations of the language by means of replacements. He did not even agree to changes that were apparently already common before his time (apart from changing the vowel letters אהו"י ['*-h-w-y*]). For example, he believed that although the roots על"ץ ['*-l-ts*], על"ז ['*-l-z*] and על"ס ['*-l-s*] all mean *joy*, they are three separate roots and not three appearances of one root in which one consonant was varied by means of a change in the position of the articulation or the way in which it was articulated.

The question of כרסם/קרסם [*kirsem/qirsem*] is not confined to the matter of letter substitution. It raises another fundamental question: Should the language of the Bible be compared with the language of the sages? Is the language of the sages equal in status to the language of the Bible? This is not a question of having faith in the sages or accepting their rulings. The question that troubled Menahem ben Saruq as a poet was the purity of language. What is the nature of the language of God, which certainly is more exalted than the language created by humans? What is the language that should serve as a model for poetry?

For example, is the creation of denominative verbs from nouns permissible and considered pure language? Menahem ben Saruq objected to this practice, not only in the poetic language of his contemporaries and the language of liturgy in the centuries just prior to his time, but even in the language of the Mishnah. One thing is clear: more than dealing with the question of what was possible from the standpoint of language, he was dealing with the question of the limits of extrapolating from linguistic forms and phenomena that are documented in the Bible, and with the question of the boundaries of good taste in Hebrew morphology.

Menahem therefore related to a very sensitive point – the boundary of good linguistic taste. He did not consent to making small compromises – he viewed the

phenomenon as completely inadmissible. If a practice was not appropriate in ten cases, then it was not appropriate in a single case, and not only in poetry – we have seen that he took his campaign into the field of prose as well, and even dealt retrospectively with ancient literature, the Mishnah! His interest was the language usage in his time and afterward, and defining the concept of purity of the language.

One of Dunash's strongest complaints against Menahem's work was his fundamental ideological objection to the practice of comparing biblical words with Arabic. Menahem compared Hebrew with Aramaic, which he also viewed as being a holy tongue, but even here he did so only when absolutely necessary, and in any event, he did not compare Hebrew with Arabic, which he viewed as a secular language. In order to prove his point, Dunash presented some 160 examples of comparisons with Arabic.

What increased the resonance of the *Mahberet* was the continuation of the polemic over etymological, semantic and grammatical issues that arose in it and Dunash's responses to the dictionary. Three of Menahem's pupils, Isaac ben Gikatilla, Isaac ben Kaprun, and Judah ben David, joined forces to formulate answers to Dunash's responses, and in turn, a pupil of Dunash, Yehudi ben Sheshet, composed answers to their answers. The amazing fact is that even after Judah ben David Hayyuj began to become prominent when he published his theory on the universality of the triliterality of the root with regard to the Hebrew verb,²¹ the dispute and the mention of *Mahberet Menahem* did not abate. The sages of France, who could not read the writings of Hayyuj, which were written originally in Arabic, continued to use *Mahberet Menahem* for many generations. The person who did the most to spread *Mahberet Menahem's* fame was the great French commentator Rashi, who quoted it overtly and covertly hundreds of times in his commentary on the Bible and the Talmud. And since Rashi's commentaries became popular throughout the Jewish world, and continue to be so today, Menahem's outdated theory of grammar and the philological interpretation arising from it that took root in Rashi's commentaries are still widely known. At the end of the twelfth century, Rashi's grandson, Rabbenu Tam, who was one of the foremost sages of the *Tosafot*,²² wrote a book meant to settle the disagreements between Menahem and Dunash.²³ Several decades later, Rabbi Joseph Qimhi, the first of the philologists of the Qimhi family, wrote *Sefer ha-Galuy*

(*Book of the Overt*) in his own effort to settle the disputes, this time in light of Hayyuj's theory.

The *Mahberet* came down to us in many copies in manuscript form and in 1854 it was published for the first time by Zvi Filipowski,²⁴ but his edition got very unfavorable reviews. A new critical edition was published by Prof. Angel Sáenz-Badillos,²⁵ one of the greatest modern scholars in the field of Hebrew philology of medieval Spain, based on the best manuscripts and the fragments from the Cairo Genizah.²⁶

Many other aspects of the perception and interpretation of the Bible, the status of the rulings of the sages of the *Masorah*²⁷ in the new philological commentary, and Menahem's method of interpretation in general are reflected in the *Mahberet*, but space does not permit a discussion of these aspects in this article.²⁸

The fact that Menahem's root theory became outdated within the span of a generation does not negate his originality and achievements. He can also be credited with achievements in the formulation of Hebrew grammatical terminology, notwithstanding the fact that some of it evidenced signs of groping in the dark. For example, the root is called both יסוד ([*yesod*] *element, foundation, base*) and עיקר ([*'iqqar*] *essence, principle*) as well as שורש ([*shoresh*] *root*). The act of attributing a particular letter to a root is called להשריש ([*lehashrish*] *to strike/take root*).

One thousand years later, there is no doubt that Menahem was right with regard to many aspects, but was wrong with others. For the latter he was subjected to incisive criticism from his adversary, Dunash ben Labrat. Although these disputes were difficult and unpleasant on the personal level, on the scientific level they were extremely fruitful and there is no doubt that they greatly advanced and benefited Hebrew language research. Moreover, thanks to Menahem it is possible to understand the enormous changes that took place in the generations that followed.

Notes

1. See E. Ashtor, *Korot ha-yehudim bisfarad ha-muslemim*, Jerusalem, Qiryat Sefer 1960, Chapter 5, p.103 ff.
2. Menahem is usually referred to by his first name and not by his appellation because his name became attached to the name of his dictionary, *Mahberet*, which is usually called *Mahberet Menahem*.
3. See Y. Blau's introductory article on MENAHEM BEN JACOB IBN SARUQ in *Encyclopaedia Judaica*, vol. 11, p.1305

(Jerusalem 1971) and the bibliography there.

4. See S.L. Skoss, *The Hebrew-Arabic Dictionary of the Bible Known as Kitab Jâmi' Al-Alfâz (Agron) of David ben Abraham Al-Fâsi the Karaite (Tenth Century)*, vol. I-II, Yale University Press, 1936-1945.
5. See Simcha Pinsker, *Likutei Qadmoniyot le-toledot benei Miqrah voha-literatur shellahem*, Vienna 1860, p.172.
6. See A. Maman, 'Peshat and Derash in Medieval Hebrew lexicons', in *Studies in Memory of Naphtali Kinberg (1948-1997)*, *Israel Oriental Studies* 19, 1999, pp.343-357.
7. The Mishnah is the collection of oral law compiled by Rabbi Judah ha-Nasi at the beginning of the third century CE; also, a single paragraph of this.
8. The Talmud is a compilation of the Mishnah and its Amoraic commentary (*Gemara*); there are the Jerusalem (or Palestinian) Talmud (beginning of the fifth century) and the more authoritative Babylonian Talmud (end of the fifth century).
9. See S. Morag, 'Reshit ha-millona'ut ha-ivrit ve-ha'aravit', in *Molad* 3, NS 1970-71, pp.575-582.
10. Eliezer Ben Yehuda (1858-1922) is considered to be the renovator of the Hebrew language. His *Complete Dictionary of Ancient and Modern Hebrew*, in 16 volumes, was published from 1908 to 1959.
11. See M. Mishor, 'Milon ha-Hoveh and Milon Sapir', in *Kernerman Dictionary News* 12, 2004, pp.23-26.
12. For example, in Rashi's commentary to Leviticus 19:19, entry ובגד כלאים, to Isaiah 19:7, entry ערות, and to Jeremiah 4: 12, entry תקרעי בפור.
13. Rav Sa'adiah Gaon had a different approach to the root, according to which the root is the actual morphological basis in the noun form (an approach that resembles the concept of the root in Latinate and Anglo-Saxon languages).
14. See *Mahberet*, the entry ש"ב, Badillos edition, pp.359*-360*.
15. For details of the analysis of the examples, see Maman A., 'The Flourishing Era of Jewish Exegesis in Spain: The Linguistic School: Judah Hayyuj, Jonah ibn Janah, Moses ibn Chiquitilla and Judah ibn Bal'am', in *Hebrew Bible – Old Testament: The History of its Interpretation I/2: The Middle Ages* (ed. Magne Saboe), Goettingen, Vandenhoeck & Ruprecht 2000, chapter 31.1, pp.261-281. With regard to the subject discussed above, see *ibid.*, pp.264-265.
16. Badillos edition, pp.75*-76*.



Title page of the *Books of Replies*, Salomo Gottlieb Stern edition, Vienna 1860



Title page of the *Books of Replies* (Hebrew), Salomo Gottlieb Stern edition, Vienna 1860



Street signs in Tel Aviv,
Menahem ben Saruq Street
(above) and Hisday ibn
Shaprut Street (below)

17. For the meaning of the term כַּמְשָׁמְעוּ, the meaning attributed to it by Dunash ben Labrat, Menahem's contemporaneous, and the disagreements about it in the nineteenth and twentieth centuries, see Maman A., *Comparative Semitic Philology in the Middle Ages from Saadia Gaon to Ibn Barun (10th-12th cent.)*, Leiden, Brill 2004, pp.276-283.

18. With regard to the concept of philological commentary and Menahem's method of interpretation in general, see Sáenz-Badillos A., 'Early Hebraists in Spain: Menahem ben Saruq and Dunash ben Labrat', in *Hebrew Bible – Old Testament: The History of its Interpretation I/2: The Middle Ages* (ed. Magne Saboe), Goettingen, Vandenhoeck & Ruprecht 2000, chapter 25.5, pp.96-109.

19. Sometimes the man's surname begins in the form of "ben" (i.e. *son*) and sometimes it begins with its Arabic counterpart, "ibn", but in most cases a fixed form is used, either Hebrew or Arabic.

20. David Alfāsi was also of this opinion.

21. Hayyuj, who lived following Menahem's generation and worked in the same field as his, is regarded as the greatest Hebrew grammarian in the Middle Ages. He discovered the nature of triliterality of the Hebrew verb even for "weak" roots and composed two large books to prove his theory. Among other innovations, he also suggested a theoretical notion of *sakin layyin*, a soft unvocalized morpho-phonemic entity, in order to solve all kinds of morpho-phonological Hebrew

problems. Hayyuj's views are accepted up to date.

22. *Tosafot (Addenda)* refers to the comments made on Rashi's commentary to the Babylonian Talmud, by his nephews, Rabbenu Ya'aqov Tam and the Rashbam (twelfth century), followed by other French and Ashkenazi scholars up to the fourteenth century. In the famous Vilna edition of the Talmud, the Tosafot are printed in the external margins of the Talmudic text, opposite Rashi's commentary, which is printed in the internal margins.

23. 'Sefer teshuvot rabbenu tam', in *Sefer teshuvot dunash ben labrat... 'al sefer mahberet harav menahem*, Z. Filipowski edition, London and Edinburgh, 1855.

24. Zvi Ben Yehezkel Filipowski, *Mahberet Menahem*, Edinburgh 1854.

25. A. Sáenz-Badillos, *Menahem Ben Saruq, Mahberet*, Granada 1986. On this edition see I. Eldar, 'Askolat ha-diqduq ha-Andalusit: teqfat ha-reshit', in *Pe'amim* 38, 1989, p.24, n.12.

26. Geniza material refers to 250,000 fragments from ragged Hebrew books and documents which were piled for centuries in a special room in the Cairo Ezra synagogue and are now preserved in several libraries around the world.

27. *Masorah* is the philological apparatus and literature meant to keep the text of the Bible untouched. It is assumed that this kind of literature emerged soon after the canonization of each book of the Bible.

28. And see in Sáenz-Badillos's article and the essays noted above.

Review of B. Katz-Biletzky, *Wörterbuch Deutsch-Hebräisch Philosophische, wissenschaftliche und technische Termini*

When I first started studying at university, an amusing adage in common usage was: "the most important Semitic language is German". Scholars of Hebrew and Semitic languages, Biblical and Judaic studies, and indeed any of the scientific disciplines recognized that German had been the principal language of research from the nineteenth century onward. It is hardly surprising that when the Haifa Technion, the most prominent technological institute in Israel, was founded in 1914 on the initiative of the Ezra Organization from Germany, it was decided that teaching should be conducted in German. It was only a consequence of public opposition that prompted the institute to adopt Hebrew as its official language of tuition.

Katz-Biletzky's dictionary is composed of an abundance of Hebrew equivalents for some 25,000 German terms, many of which have been in existence and dispersed throughout the extensive canon of Hebrew philosophical and scientific literature since the Middle Ages. The target audiences for this book are scientists and translators.

The dictionary itself has 720 pages. Following an introduction given in both Hebrew and German is a list of publications used in the compilation of the dictionary. This includes 113 Hebrew sources, books and articles, and 61 mainly German books and dictionaries. The final part of the dictionary includes a list of terms that the author himself has used his considerable scholarship to innovate.

The book's lexical entries cover all science-related areas that a well-informed individual is likely to need: philosophy, psychology, anatomy, economics, zoology, biology, chemistry, geography, linguistics, and so on. Each term has Hebrew equivalents gathered either from one of the sources mentioned above or in some cases coined by the author.

The author retains a purist attitude towards translations and definitions throughout the work. He cites Hebrew terms before loan ones, avoids loan words as much as possible, and invents Hebrew terms and cites them as the first translated option.

The directory of abbreviations makes it possible for each Hebrew term listed to be placed in its relevant field e.g. *'eti(ka)* 'ethics', *mahsh(evim)* 'computer sciences', *dat* 'religion', etc. Some entries are cited with a Hebrew source immediately after the first one or two terms, although not necessarily referring to them, e.g. German *Cubus* – Hebrew *me'ukav* 'cubic' (*mat(ematika)*) [k/b/240] which refers to Jacob Klazkin's *Thesaurus of Philosophical Terms*, volume 2 (New York, Feldheim 1968), p.240; *Modulation* – *silum (mus(ika))* [lmv/249] referring to Aviva Shelah's *Dvir Musical Lexicon* (Tel Aviv, Dvir 1990) (*silum* is in fact used within social sciences in modern Hebrew to mean a hierarchical building scale and it also appears in the translation of *Skalierung* in statistics). *Nominalisierung* 'nominalization' is first translated into Hebrew as *shimuy* based on Ora R. Schwarzwald and Michael Sokoloff's *A Hebrew Dictionary of Linguistics and Philology* (Even Yehuda, Reches 1992), p.146, but this term is not used there at all, but rather *nominalizatsya* or *ha'atsama*. These terms are conveyed secondarily in Katz-Biletzky's dictionary, only after the citation. The same applies to *Tautologie*, translated as *yitur lashon* which is based on Klazkin and the same linguistic dictionary, even though both sources do not mention this definition. In the linguistic dictionary 'pleonasm' is translated as *yitur milim* (not *yitur lashon*), or as *pleonasma* and *yitur*, while 'tautology' is translated as plain *tautologya*. Had the author included references at the end of each lexical entry, this misleading information could easily have been avoided.

It is Klazkin's thesaurus that is seemingly the authentic source for the last word in the Hebrew title of the book: *Milon germani-ivri lemunhey filosofya, mada va'asut*, i.e., *A German-Hebrew dictionary for philosophic, scientific and technological terms*. The word 'asut, 'technology', does

not exist in contemporary Israeli Hebrew nor indeed in any Hebrew dictionary; the word *tekhnoLOGYa* is the commonly used term. Even though it is the author's personal innovation to make use of this specific term, he attributes it to Klazkin.

Although the dictionary is German-Hebrew, many English and French terms are also listed, but are only given some clarification when they happen to be the same in German. Each of these words is referred to the German term where it is translated. For instance, *Langue* ---> *Sprache*; *Sprache* {*Langue*} is translated as *leshon haklal* (in linguistics), *lashon, lang, leshon hahevra*, etc. *Binary* ---> *binär*; *binär* is translated as *shniyoni, duhelki, binari*, etc. It might have been more helpful to translate the English terms as well, as it is English that has become the international language of science since the middle of the twentieth century.

Most of the entries are explained in addition to their Hebrew equivalents, e.g. *Bezugsgruppe* – *kvutsat hityahasut* 'reference group' (psychology, social sciences) is explained in Hebrew as "any group in which the individual sees in its ways of behavior, values and goals criteria for his personal evaluation, and according to which he positively or negatively directs his own behavior, values and goals".

The author's innovations are linguistically interesting. He uses only Hebrew elements and many words are derived from Semitic root and pattern construction, e.g. *schmettern* – *himtikh* 'sounded like metal' (root *m-t-kh*, *hif'il* pattern, from *matékhet* 'metal'). Many words are linearly derived either by prefixing or suffixing, or by blends, e.g. *global* – *kadratsi (kadur ha'arets* 'globe' + *-i* 'adjectival suffix'; *globáli* in modern Hebrew). The formation of new roots using an initial *shin* is extensive, but the number of compounds is sparse.

What makes this dictionary so useful is its wide range of terms, some of which have yet to be officially translated into Hebrew. The explanations in the book are useful too. The book's chief disadvantage derives from the fact that as it is a research dictionary, albeit extremely extensive, the first Hebrew term for each lexical item is determined by the author's personal preference or invention and not by practice. This makes the dictionary less reliable for the average Hebrew speaking user.

Ora R. Schwarzwald

Hebrew and Semitic Languages
Bar-Ilan University
oschwarz@mail.biu.ac.il



Wörterbuch Deutsch-Hebräisch Philosophische, wissenschaftliche und technische Termini

Benjamin Katz-Biletzky
757 pages
ISBN 3-87548-265-4
Helmut Buske Verlag
Hamburg 2003
www.buske.de

Towards Hebrew FrameNet

Miriam R. L. Petruck

Miriam R. L. Petruck holds a Ph.D. in Linguistics from the University of California, Berkeley. She has been connected with FrameNet, at the International Computer Science Institute in Berkeley, CA, since 1998. Her current work to develop Hebrew FrameNet brings her back to her earliest interest in modern Hebrew semantics, syntax and the lexicon.
miriamp@icsi.berkeley.edu

Thanks to a publicly available Hebrew language (newspaper) corpus¹, as well as other web-based resources, such as Rav-Milim² and Hebrew WordNet³, the creation of Hebrew FrameNet (HFN) has become possible. Moreover, there are good prospects for cooperation and collaboration with the computational linguistics community in Israel (e.g. access to larger corpora for research and evaluation purposes or use of software for lemmatization and search).

In the context of research that investigates the universality of the semantic frame, initial steps have been taken towards the development of HFN, an on-line lexical resource for contemporary Hebrew which will provide the semantic and syntactic combinatorial possibilities, or *valences*, for each item analyzed, through the manual annotation of example sentences in a newspaper corpus (and eventually, the automatic capture and organization of the annotation results for web-based viewing and querying). With advances in computer technology and the existence of (searchable) corpora, the work of lexicography has changed dramatically. The fine-grained semantic classification and syntagmatic information of the sort to be provided by HFN will make its database an invaluable resource for lexicographers and advanced language teachers/learners, as well as researchers in linguistics and natural language processing (NLP).

In accord with FrameNet, the first computational lexicography project of its kind⁴ (Fontenelle 2003), HFN is based on the principles of Frame Semantics (FS; Fillmore 1977, 1985, Petruck 1996), at the heart of which is the *semantic frame*, an experience-based schematization of the speaker's world against word meaning can be understood. In FS, a linguistic unit evokes a frame, whose *frame elements* (FEs), or participants and props in a scene, indicate the semantic roles that need to be filled. An FS analysis of a lexical item relies on the identification and definition of the frame(s) in which the word participates, along with the frame-specific FEs that are recorded as triples of information about the semantic role, the phrase type and the grammatical function of the annotated constituent.

Initially, HFN will produce full annotation for frame evoking elements⁵ in the newspaper corpus. This serves as a means of (1) creating the infrastructure

for using the FrameNet DeskTop (FNDDT; a suite of tools used for defining frames, FEs, and words, and annotating example sentences) for the analysis of Hebrew texts and (2) determining the level of linguistic description and computational representation at which the lexicon of modern Hebrew can be characterized in terms of existing FS concepts. Adapting FNDDT for HFN will demonstrate the feasibility of using the software for a non-Indo-European language.⁶ Investigating the linguistic expression of events and scenarios through the same or different frames will also document the different lexicalization patterns of Hebrew and English (Talmy 2000).

As with FrameNets for other languages (e.g. Spanish⁷), the HFN database will function as both a dictionary and a thesaurus. The dictionary-like features include definitions, tables summarizing the patterns of syntactic realizations of FEs occurring with a word, and sets of annotated sentences from the corpus showing the semantic information associated with each syntactic pattern. Like a thesaurus, words are linked to the semantic frames in which they participate, and frames are linked to other collections of words and to related frames. FS analyses are useful for research in crosslinguistic lexicology (Subirats and Petruck 2003) and in the advanced foreign language classroom (Sato 2004). The availability of such information via the internet will facilitate studies in Hebrew linguistics as well as Hebrew language teaching/learning. Ultimately, HFN data will serve the needs of research in NLP for Hebrew, contributing deep semantic information for a variety of tasks, including word sense disambiguation, machine translation, information extraction and question answering (Likowski 2004).

Notes

1. <http://mila.cs.technion.ac.il/website/english/resources/corpora/2000sentences/index.html>
2. <http://ravmilim.co.il>; see *Kernerman Dictionary News* 12, 2004.
3. <http://multiwordnet.itc.it/online>
4. <http://icsi.berkeley.edu/~framenet>
5. A frame evoking element is any sense of a word that brings to mind a frame.
6. Similar efforts are under way for Japanese, see <http://nak.ics.keio.ac.jp/jfn>.
7. <http://gemini.uab.es/sfn>

The complete version of this article, including the full references as well as a practical example, is available online:
<http://kdictionary.com/kdn/kdn1304.html>

Contemporary English-Russian Dictionary: A New Type of Dictionary in Russia

Galina Zakharova

English-Russian dictionaries have always been in great demand in the Russian book market, as a means to help the study and use of the English language for world communication.

In the former USSR, the Russky Yazyk publishing house was the leader, and monopolist, in the field of publishing dictionaries, English-Russian ones in particular. This was a huge publisher with old traditions and a well-known lexicographic school, which formed and trained several generations of lexicographers.

During the last 20 years, since the perestroika, we have witnessed great changes in our economy, including the book publishing industry. Now the dictionary market is no longer uniform, Russky Yazyk has lost its leading position, and it produces fewer dictionaries than before and does not have the national school of lexicography any more.

Nevertheless, the dictionaries compiled and edited by lexicographers of the former Russky Yazyk are highly esteemed and are of great demand in the market. These are *Modern English-Russian Dictionary* (by V.K. Müller, V.L. Dashevskaya, V.A. Kaplan, Moscow, Russky Yazyk – Dropha 2002), *New English-Russian Dictionary* in 3 volumes (edited under the supervision of E.M. Mednikova and Y.D. Apresyan, Moscow, Russky Yazyk, 1st edition 1993) and the recently published *New Comprehensive Russian-English Dictionary* (by D.I. Ermolovich and T.M. Krasavina, Moscow, Russky Yazyk, 2004).

There are also other major publishing houses involved in the local dictionary trade, for example AST, whose *English-Russian and Russian-English Dictionary for Schoolchildren* (2004) contains grammar notes and was compiled by professionals from the old school. However, many of the dictionaries currently available in bookshops are reprints of old editions, without proper revision and editing, just new colour jackets.

Our publishing house, RUSSO, was established in 1993 on the basis of the scientific and technical dictionaries department of Russky Yazyk. We publish specialist and scientific dictionaries on various subjects, including exact sciences

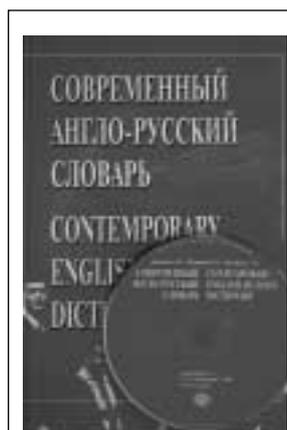
(physics, chemistry, biology, etc) and applied sciences (medicine, economics, law, transport, agriculture, computer sciences, etc).

It was the idea of Ludmila Popova, one of the prominent Russian lexicographers and former chief of the department of Germanic languages at Russky Yazyk, to publish an English-Russian dictionary of a new type, principally different from those that existed. We decided that this dictionary should feature popular terminology of science and technology as well as everyday language, while making room also for fields that were traditionally weakly represented in the dictionaries of the past, such as religion, sociology, psychology, sexology, youth slang, new and extreme sports, modern music and fashion.

The *Contemporary English-Russian Dictionary*, published in 2004, contains about 50,000 words and 70,000 phrases. It combines general, everyday language with up-to-date terminology from the fields of technology, medicine, biology, transport, sports, etc. Special attention was paid to computer terminology, slang and pop music terms. The dictionary can help in reading and translating various texts – from classical literature to newspapers, magazines and popular science. It includes literary words, even obsolete and dialectal, which are necessary for our users in order to read nineteenth century literature, alongside slang and expressions commonly used in modern literature. It makes use of numerous labels (e.g. functional, regional, stylistic) to help distinguish different shades of meanings. We have aimed, in particular, to make this dictionary as comprehensive and user-friendly as possible, and have thus accompanied the book by a CD-ROM.



Galina Zakharova is the director of RUSSO



Contemporary English-Russian Dictionary

Ludmila Popova, Natalia Mokina, Galina Zakharova
944 pages
ISBN 5-88721-274-8
Russo Publishers, Moscow
2004

RUSSO Dictionaries cover a wide range of specialist areas, including:

advertising · agriculture · alimentation · animals · art · automobiles · banking · biology · chemistry · computers · construction · economics · electronics · engineering · food · foreign exchange · forestry · geology · gourmet · industrial automation · information technology · law · marketing · mass media · medicinal plants · medicine · metallurgy · mining · nuclear power · optics · patents · petroleum · politics · psychology · religion · science · security · sociology · solar energy · technology · telecommunications · trademarks · viticulture

www.russopub.ru

A Large-Scale Lexical Database of Danish for Language Technology Applications and Other Purposes

Anna Braasch



Anna Braasch was born in Budapest in 1946, studied languages at ELTE University, and has a Master in German from both Gothenburg and Copenhagen universities with Minors in Computational Linguistics and Russian. She is a senior researcher at the Centre for Language Technology (CST), University of Copenhagen, and managed the STO project. Her expertise is in computational lexicography and terminography, her main research interests are phraseology and lexical semantics, and her language skills include Danish, English, German, Hungarian, Russian and Swedish. She was on the boards of the Danish and Nordic lexicographer associations (respectively LEDA and NFL), was President of NFL, and is currently the Vice-President of EURALEX. anna@cst.dk

1. Background

SprogTeknologisk Ordbase (STO, Lexicon for Language Technology Applications) is the most comprehensive computational lexicon of Danish, primarily developed for Natural Language Processing (NLP), including commercial language technology products and computational linguistic research purposes. STO was created within the framework of a national project, led by the Centre for Language Technology (Center for Sprogteknologi, CST) at the University of Copenhagen. The work was carried out as a result of a contract with the Danish Ministry for Science, Technology and Development. The three-year project ended in February 2004.

CST initiated the project and was responsible for its management and the coordination of the work, including tasks such as software development, the creation of linguistic specifications and definition of guidelines for encoding.

The lexicon material was produced in cooperation with the Institute for Computational Linguistics of the Copenhagen Business School, the Institute for General and Applied Linguistics of the University of Copenhagen and the Department of Business Communication and Information Science of the University of Southern Denmark.

The project members had various skills relevant to the task comprising theoretical linguistics, terminology, lexicology, corpus linguistics, computational lexicography and linguistics, and database knowledge. This wide range of expertise ensured that the lexicographic work carried out was of high quality, which is reflected in the lexicon.

The need for a computational lexicon

The main objective of the STO project was the development of a computational lexicon of Danish for a broad practical application area. Language industry and research into computational linguistics often experience the lack of a large-scale comprehensive lexicon as a bottleneck problem for the development of most applications and software tools for language engineering. In particular, for less widely-spoken languages such as Danish, it is essential to develop some multipurpose and flexible language resources in order to optimize the cost/benefit ratio.

In this sense, the STO serves as a basic lexical data collection from which various dedicated modules can be derived for particular applications, such as lemmatizers, inflection analyzer/generators, shallow parsers or Danish modules for Machine Translation (MT) systems.

2. The lexicon

The most important features of the STO lexicon are, besides its size, being theoretically well-founded and empirically supported. The descriptions are very detailed, each piece of information is labelled explicitly and precisely, and any item is easily accessible as the entire lexicon is structured and stored in a relational database (ORACLE). Thus, it is straightforward to extract, for example, all syntactic frames of a lemma, or all lemmas sharing the same syntactic frames or a particular syntactic construction, for research purposes.

The composition of the lexicon

The STO lexicon contains over 81,000

Lexical Category	No. of Lemmas	Morphology only	Morphology & Syntax	Morphology & Syntax & Semantics
noun	64735	47%	41%	12%
adjective	9773	32	55%	13%
verb	5775	2%	81%	17%
adverb	771	81%	19%	
interjection	158	100%		
preposition	80	100%		
conjunction	60	100%		
pronoun	44	100%		
misc.	128	100%		
TOTAL	81524			

Table 1: The composition of the entire STO vocabulary

lemmas, of which approximately 14,000 come from six different domains of language for specific purposes (LSP). All lemmas are provided with lexical category information and exhaustive descriptions of their inflectional properties, and 45,000 of them with a fine-grained syntactic description as well. Tables 1 through 3 show the composition of the vocabulary covered in detail. The STO database is not intended to cover highly specialized terms, but focuses on words of the domain languages that laymen will have to read and understand as part of everyday life. This is considered a kind of transitional area between the general language and specialized expert languages.

General language and domain language vocabulary

Table 1 shows the composition of the entire STO vocabulary classified by the feature 'Lexical Category' (in other terms: 'word class' or 'part of speech') and to which extent the different word classes have been provided with either (a) only morphological information, or (b) morphological and syntactic information, or (c) morphological, syntactic and semantic information.

The large number of lemmas with only morphological information is especially useful in applications such as shallow parsers, taggers, spell checkers, etc.

The words for the syntactic encoding were selected on a frequency basis; all verbs are provided with syntax, whereas only nouns and adjectives above a certain frequency threshold are provided with syntactic information.

Tables 2 and 3 specify the vocabulary from the selection point of view (as originating from general language and domain language texts).

Lexicon model

The establishment of the descriptive model and the linguistic specifications for STO greatly benefit from the experience

acquired at CST within the framework of the multilingual (LE2-4017) PAROLE project (1996-98) of the European Commission. The PAROLE lexicons were built around a generic model, an instantiation of the EAGLES recommendations in an enriched GENELEX model (details about the EAGLES and GENELEX projects are available as part of the PAROLE information: http://hltcentral.org/usr_docs/project-source/parole/ParoleFinal.pdf). Thus, the Danish STO lexicon is well integrated in the multilingual infrastructure of European computational language resources, which ensures its compatibility with other resources developed for Human Language Technology (HLT).

Linguistic description

The STO lexicon is corpus-based as regards both the selection and the description of lemmas. The linguistic descriptions are based on corpus analysis, and all lemma types are treated in a uniform way.

The linguistic information content of the STO lexicon is organized according to the traditional practice in computational linguistics of division into three independent descriptive layers, i.e. morphological, syntactic and semantic. Each descriptive layer is made up of a comprehensive system of the characteristic linguistic properties. The linguistic description of a lemma is structured in different sets of information, the so-called units; each unit represents a particular morphological, syntactic or semantic behaviour of the lemma at the layer concerned. From the computational point of view, a unit is a structured object containing a feature-based description expressed in attribute/value pairs.

The general linguistic features described at the three layers are as follows:

- Morphology: lexical category, inflectional patterns, spelling variants, agreement features, compounding properties, etc.
- Syntax: syntactic patterns comprising subcategorisation frame (categorical

Center for Sprogteknologi (CST)

The CST (Centre for Language Technology) is a research institute at the University of Copenhagen. It employs a staff of approximately 20, including computational linguists, lexicographers, computer scientists and engineers. Its mandate is to carry out and promote strategic research and development in language technology (LT) and computational linguistics in Denmark. In addition to research activities, the CST has considerable experience in the development of a variety of LT applications for both national research projects and commercial purposes, including Machine Translation, XML mark up, company specific ontologies and inter-institutional term databases for EU institutions and agencies, as well as controlled language and authoring tools. <http://cst.dk>

Lexical Category	Number of Lemmas
noun	52840
adjective	8568
verb	5410
adverb	771
interjection	158
preposition	80
conjunction	60
pronoun	44
misc.	128
TOTAL	68059

Table 2: General language vocabulary (all closed word classes belong to this category)

Domain	Nouns	Verbs	Adjectives	Total of Domain
IT	1730	160	115	2005
Environment	1770	50	300	2120
Commerce	1800	60	160	2020
Administration	2430	25	220	2675
Health	2285	40	250	2575
Finance	1880	30	160	2070
TOTAL	11895	365	1205	13465

Table 3: Domain language vocabularies in the STO database with part of speech distribution

STO availability

The lexicon material is now available for both commercial use and research purposes. Starting this summer, the licensing and distribution will be handled by the Evaluations & Language resources Distribution Agency. (ELDA, <http://elda.org>).

Standard packages

- 81,000 entries with morphological description only, provided with full documentation of the morphological layer;
- 81,000 entries with morphological description whereof 45,000 entries are provided with syntactic information including full documentation of both layers.

Standard delivery formats

- Morphology in comma-separated plain text files
 - Syntax in XML format
 - ORACLE dump database files (8.I on request)
- User defined data package types and delivery formats can be produced on demand.

Additional facilities

The STO web interface provides links to other on-line Danish language resources, such as electronic dictionaries for human use (*Retskrivningsordbogen*, the Official Spelling Dictionary, and *NetOrdbogen*, the Internet Dictionary) and corpora (*Korpus2000* and *Berlingske Tidende*, a newspaper corpus). In addition, Danish websites can be searched through a link to Google. These facilities enable direct searches in a user-friendly way, e.g. to compare the STO data with information in the electronic dictionaries and supplement the STO data by corpus evidence. <http://cst.dk/sto/uk>

and functional valence), diathesis and alternation phenomena, reflexivity of verbs, etc.

- Semantics: the information is provided at three specificity levels. Level 1 contains domain reference only (all entries). Level 2 comprises domain information, ontological type, argument structure and selectional restrictions (about 2,000 entries). Level 3 is identical with the SIMPLE semantics. Information types of level 3 are the ontological type, semantic relation, argument structure, selectional restrictions, qualia structure, event structure, domain information, etc (about 7,000 entries). The subdivision of the semantic information into three levels is introduced for practical reasons. Levels 1 and 2 are proper subsets of level 3, representing a relatively lean semantics.

Validation

In a collaborative lexicon project like STO, it is a key issue to ensure the inter-coder consistency in order to achieve homogeneity of the linguistic content. To this end, the lexicographers were guided by detailed encoding guidelines and worked with encoding tools supporting consistency checks. The successive stages of the work were organized in three steps: the lemmas were encoded by one lexicographer/team, then checked/revise by another, finally all data were validated at CST before uploading it to the STO database. The reported experience and comments from external users were taken into consideration during the process.

3. STO data in use

STO is currently the largest and most comprehensive computational lexicon for the Danish language, and the demand for this resource is growing. The material is already being used in a number of projects and applications, for a variety of purposes. According to users' specifications, data subsets were extracted from the lexicon. These were adapted to various format requirements and the linguistic content was exploited for both particular research and development purposes. This way, both the linguistic content and the formal properties of the lexicon were judged from the user's points of view. The examples below illustrate some typical uses of the STO-data.

In research

- evaluation of search engine behaviour in a multilingual environment
- computational analysis and processing of complex sentence structures from the point of view of potential reading speed
- conversion of verb entries into the lexicon

format of the Dependency Treebank

- testing of a computational grammar for Danish
- using the qualia structure information to calculate semantic relations in compounds

In practical applications

- MT for a specific domain
- lemmatizer for Danish
- information retrieval system prototype
- preparatory work with the aim of exploitation of verb descriptions in constructing a dictionary for humans
- ongoing development for speech technology applications, extension with pronunciation of all word forms

Perspectives for further exploitation

Reports on successful experimental applications and positive responses from users provide a promising basis for marketing the STO resource both for the research community and for commercial NLP/HLT tool developers.

Currently, only few industrial products are developed for Danish at all, partly due to the bottleneck problem of lacking a lexical resource. Because of its comprehensive and detailed content, STO can keep up with very different demands and be exploited as a lexicon component in both monolingual tools (parsers, taggers, authoring tools, browsers, spelling/grammar checkers) and in multilingual applications (MT systems, search engines, etc) as well as for HLT tasks such as developing computer-aided language learning tools for Danish as a second language, etc.

4. User interface: looking up single entries

In addition to various NLP applications, STO offers a valuable resource to linguistic researchers, teachers and learners of the Danish language. To facilitate access, there is a web interface that enables various searches and corpus investigations. However, the database contains more linguistic information than shown on the screen for human users.

Search options

- Word Search displays all the inflected forms and syntactic constructions of the lemma.
- Compound Search displays all the compounds containing the search lemma as one of its elements.
- Corpus Search establishes links from each result of a Word Search to direct searches in corpora.
- Parameterized Search uses a combination of the lexical category and the value(s) of all its selected prevalent properties.

If dictionaries are free, who will buy them?

Adam Kilgarriff

If dictionaries are free, who will buy them? The question looms over publishing houses like a slow-motion tsunami. Dictionaries are now free, on the web and bundled with Microsoft Office and other products, so where is the publishers' income stream going to come from?

Articles in the last four editions of this newsletter have addressed the title question. Charles Levine (KDN9) opened the debate in optimistic mode, seeing signs of growth in English-language lexicography despite the web. Joseph Esposito's response (KDN10) was a:

“grim vision ... when I complained about Microsoft bundling a spell checker, with its limited dictionary, into Word ages ago, the techies I knew all laughed at me. Now that most of them have burned through their venture capital after Microsoft “integrated” the gist of their products into Windows, we all cry into our lattes together.”

Then, following Levine's still-optimistic response (KDN11), the villain of the piece bravely entered the fray. Microsoft's Julian Parish (KDN12) argued that Microsoft saw publishers as partners, not competitors. There were many new e-opportunities, for example, as add-ons to Microsoft products.

History: a dictionary in every household

In the twentieth century, a number of European and North American publishers occupied the fertile coastal strip of “a dictionary in every household”. Dependable as the cycle of one generation growing up and handing over to the next, it was a large and enviable market, in harmony with the grand and noble agenda of universal education. To be sure, the coastal strip was sometimes crowded with competitors, but the soil was good: there were always more households to buy dictionaries. They don't need to buy them any more.

There is no use lamenting the lost market. It may disappear with varying speeds: as Esposito notes and Levine confirms:

“In the absence of growth, the old business will be strained for capital, which will beget smaller investments, which will in turn hasten the decline. In the short term, this will rebound to the benefit of market leaders, such

as Merriam-Webster and Oxford University Press...”

But disappear it will.

The market which is collapsing is the monolingual, emblematic “dictionary-at-home” market (the role of which has always been complex: status symbol for spelling, scrabble and – sometimes – schoolwork). Different markets, notably the boom EFL one and bilinguals that people need for travel and language-learning, have different trajectories.

The future

For the regular monolingual centerpiece, away from that lush dictionary-in-each-household coastal strip, what is there?

The key lies in quality. Most free dictionaries are not very good. Most people don't care: a dictionary is a dictionary, good or bad, and one is plenty. Some free ones are even quite good; Esposito and Levine note the quality of the Encarta dictionary, possibly the first of a new breed of market-swamping, “good-enough” dictionaries.

But the minority of people for whom language is their trade do care. They are the translators and academics, etc. The numbers are tiny compared to the golden age but, in this, dictionary publishing is undergoing the same transformation as many other markets with the advent of the internet: the market fractures, and where there were a small number of products selling to millions, there are now millions of products – selling far smaller numbers – to billions. The up side is that customers can be found all over the globe and, once found, they are the right customers for the product so are likely to be willing to spend more.

The nice thing about this is that making good dictionaries, as opposed to bad ones, is what every lexicographer wants to do. There is usually tension between lexicographer and publisher – better vs cheaper – and the change in the market gives more weight to the lexicographer's case. While Esposito despairs at the traditional publishers being left “to focus on the scraps Microsoft leaves on the floor”, we note that the market for the most accurate, the most consistent and the most current account of a language (or source-target pair) is far more than a scrap.

Of course, language professionals will be online. Lexicographically, this is exciting as it means the dictionary can be



Adam Kilgarriff is a Director of Lexicography MasterClass Ltd and Research Fellow at University of Sussex. His specialisms are language corpora, computational linguistics and lexicography. He has worked for Longman and as consultant to Oxford University Press, Cambridge University Press, Bloomsbury and Macmillan. He is on the executive board of EURALEX. adam@lexmasterclass.com

Charles M. Levine, The Coming Boom in English Lexicography: Some Thoughts about the World Wide Web (Part One)
Kernerman Dictionary News
9, July 2001
<http://kdictionaries.com/newsletter/kdn9-1.html>

Joseph J. Esposito, Dictionaries, another Netscape?
Kernerman Dictionary News
10, July 2002
<http://kdictionaries.com/newsletter/kdn10-1.html>

Charles M. Levine, The Coming Boom in English Lexicography – Reconsidered (Part Two)
Kernerman Dictionary News
11, July 2003
<http://kdictionaries.com/newsletter/kdn11-1.html>

Julian Parish, Microsoft and Dictionary Makers: Defining Partnerships
Kernerman Dictionary News
12, July 2004
<http://kdictionaries.com/kdn/kdn12-1.html>

far better than any that went before: it is not constrained by space, and we can open our vision to the dictionary as an object integrated with the underlying corpus resources (as in Word Sketches¹). But that is a different topic: here, our concern is for income streams.

Many of the language professionals are associated with universities and libraries. They are traditional customers for dictionaries, have substantial budgets, and, with physical space ever at a premium, are often enthusiastic about services which do not incur extra demands on space or personnel.

For example, Oxford Reference Online² is an online subscription service, sold almost exclusively as a site licence to institutions, incorporating a wide range of Oxford University Press's reference materials. It is very successful. Extensions which focus on language resources are planned. Of course, OUP has a wonderful brand, and has so many resources that it is able to offer a very broad resource, a one-stop-shop which is attractive to libraries. Others probably need to assemble into consortia (branding according to the best-known brand in each market). It is a route out of the path of the tsunami.

Dictionaries for computers?

All of the above is about dictionaries for people to use. Esposito, writing in 2002, says:

“The real game for Microsoft is using lexical databases within computer algorithms, as in natural-language processing.”

Parish, too, stresses that Microsoft is an energetic customer for dictionaries for NLP (aka Language Technology, Computational Linguistics). As an NLP researcher, I'm a little sanguine here. To be sure, most NLP applications need dictionaries as inputs. In the short term most will probably be derived from dictionaries as we know them, where there are good ones available at reasonable cost. But consider, for example, Prinsloo and de Schryver's spellcheckers for African languages³. The wordlists are corpus derived.

Across NLP, researchers are finding ways of solving problems using corpora. While high-quality, well-structured, hand-crafted resources currently support technologies that corpus-derived resources don't, the list is shrinking. Even three years ago, Esposito's remarks looked right, but now, as NLP has changed, and while it may often be a short-term convenience for Microsoft and others to take publishers' resources, it is not an income stream for

the long term. While post-editing corpus-derived resources is a job that will need doing for some time yet, it is less than a glorious future for the grand old names of dictionary publishing.

Notes

1. <http://www.sketchengine.co.uk>
2. <http://www.oxfordreference.com>
3. D.J. Prinsloo and G.-M. de Schryver 2003. 'Non-word error detection in current South African spellcheckers'. *Southern African Linguistics and Applied Language Studies* 21/4 (Special issue on 'Human Language Technology in South Africa: Resources and Applications'): 307–326. <http://tshwanedje.com/publications/nlp-SPELL.pdf>

Further Comments

Joseph Esposito

This is a very good article and I am delighted to have read it. A few comments, but all in the spirit of appreciation:

First, I do not and did not despair at the demise of Oxford and Merriam. I despaired at the complacency of the managements of these organizations for not addressing the maturity of their product lines and business strategy.

Second, the specialized dictionaries noted in the piece are precisely what I envisioned as the “crumbs” Microsoft would leave behind. As Bill Gates said apropos Sun Microsystems, you can only climb up that tree so far.

Third, the economic success of Oxford Reference Online is arguable. It's a good idea; if it generates a positive contribution to overhead (as I suspect it does), it's a very good idea. But it is not a sustaining idea. Library subscriptions cannot support the overall editorial and technical infrastructure of OUP's dictionary efforts. Other revenue streams must chip in, and those revenue streams are drying up.

Fourth, new product development (of innovative products, of course) is the answer to the business problems – but when is that not the case? The problem for the current crop of dictionary-makers is that the new products require additional investment, but the old products require ongoing maintenance. So the total cost of being in the business continues to rise, even as the market becomes more specialized.

Finally, and this is a matter of semantics, when I said that dictionaries would find a market in computational linguistics, I did not mean the current crop of general-interest dictionaries but new

forms of dictionaries, such as the ones derived from algorithmically generated corpora. The real point here is that there is a future for dictionaries, but there is no future for lexicographers. That's why we have machines. The future is only trivially post-print; more fundamentally, it is post-human. Do I despair? You could have asked that question of Cro-Magnon Man. But wasn't he, like ourselves, doing everything for his children?

Charles Levine

More than two decades into the digital revolution, and more than one decade into the Internet revolution, Adam Kilgarriff's assertion that the monolingual dictionary market is collapsing remains hyperbole – at least from the perspective of the United States.

There is no sign yet that sales of printed monolingual English dictionaries have declined in America; and, in fact, there seems to have been a modest increase over the past few years; possibly the result of an upturn in the economy and the appearance on the scene of more learners of English and more learners dictionaries to serve them.

For example, Merriam-Webster in 2003-2004 had one of its strongest launches ever of a new edition of its flagship Collegiate Dictionary – selling more than one million copies in the first twelve months. Merriam smartly bundled a free CD-ROM with the new print dictionary, but raised the retail price only one dollar.

Another example: An unnamed source at the country's largest book chain confirmed for me that the chain has experienced an up-tick in sales of print dictionaries over the past five years – which she attributed partly to an increase in market share by the chain, and partly to an increase in consumer purchases.

This is not to say that dictionary behavior is not changing rapidly, or that someday (sooner or later) digital lookups won't eclipse print lookups; but at the moment, the print monolingual English dictionary market in the States is chugging along with signs of modest growth.

Kilgarriff says that "Most free dictionaries are not very good," but this is misleading. For example, at the Merriam site (www.m-w.com) a visitor has free access to the previous edition of the Collegiate, which remains more than adequate for most lookups. And Merriam offers purchasers of its new print Collegiate a one-year free subscription to the corresponding updated fee-based Web site. Traffic on all its sites, I am told, has

increased significantly since 2003.

And www.dictionary.com (for whom I do some consulting) currently has free lookups to the latest American Heritage Dictionary of the English Language, which is one of the most popular of American monolingual dictionaries.

Not to nit pick further, but Kilgarriff mischaracterized my optimism about the near-term future of English lexicography, as something that would occur "despite the web." My optimism about English lexicography is because of the Web, and possibly the next phases of the Web, like the Semantic Web, we hear so much about.

By all (anecdotal) accounts, the Web has increased interest in language among students; has increased the worldwide use of English as a lingua franca; has given lexicographers more tools (like Google) with which to research and monitor language; and seems, as in the experience of Merriam, to have buoyed sales of print dictionaries, for the time being, at least, and not to have detracted from them.

Although I agree with Joseph Esposito that the print dictionary business may not be a growth business – i.e. not something you want to invest your money in if you are looking for a healthy return – and that this means that the dictionary business will be strained for investment capital, I am not one to predict the imminent decline of dictionary making – taking into account all sources of revenue, both print and electronic.

But, who can see clearly ahead ten, twenty, or fifty years? By then, we might have forms of e-paper and plastic look-alike-books and pocket gadgets, all utilizing wireless broadband, on which we will be able to read entire libraries of information and content.

To me this suggests that some lexicographers, as much as science fiction buffs, might want to volunteer to be frozen cryogenically and revived in fifty years just to see how people look up words.

Unlike Esposito, I am confident that any such plucky lexicographers, once defrosted in the future, will find numerous co-workers still busily keeping up with a dynamic and changing English language.

Adam Kilgarriff

I hope Esposito's not right about the shrinking market failing to sustain lexicograph*ers* (as no-one will be investing in it), though I must admit there is a grim logic to what he says.

Published by arrangement with
K DICTIONARIES



GLOBALDIX Multilingual Dictionary

**Mobile and Pocket PC
KIELIKONE**

Helsinki, Finland

March 2005

Languages included:

- Danish
- Dutch
- English
- Finnish
- French
- German
- Hungarian
- Icelandic
- Indonesian
- Italian
- Latvian
- Lithuanian
- Norwegian
- Portuguese
- Polish
- Russian
- Slovak
- Spanish
- Swedish
- Turkish

Platforms supported:

- Nokia 9500 Communicator
- Nokia 9300 Smartphone
- Microsoft PocketPC Phone
- Windows CE 3.0
- Compaq iPAQ H3800
- QTek S100 phone

<http://kielikone.fi>

[http://kdictionaries.com/
products/multi/gd2m.html](http://kdictionaries.com/products/multi/gd2m.html)

Dictionaries and productive tasks in a foreign language

Paul Bogaards

Paul Bogaards was born in Leiden, The Netherlands, in 1940. He is an experienced teacher and teacher trainer in French, applied linguistics and didactics, and has taught French, applied linguistics and lexicology at Leiden University from 1976 to 2002. He compiled several Dutch/French dictionaries, published by Van Dale Lexicografie (Utrecht and Antwerp) and Dictionnaires Le Robert (Paris), wrote books on learner characteristics, language learning, and vocabulary learning in a second language, published numerous papers in various journals, and is currently the editor of the *International Journal of Lexicography*. His main research interests concern the acquisition and testing of lexical knowledge in a second language and the use and usability of dictionaries.
p.bogaards@let.leidenuniv.nl

Introduction

The process of making dictionaries involves three main steps: **collection**, **description** and **presentation**. The first step, the **collection** of data, has seen dramatic changes over the past twenty-five years. Over more than four centuries, from the sixteenth till well into the second half of the twentieth century, lexicographers had to read and extract written texts and make paper slips with quotations which were put together in huge boxes, occupying a lot of space. Nowadays, for many languages there are text corpora containing far more numerous and more varied materials. For the English language in particular these corpora are easily available and they contain not so much literary texts as was previously the case. Most of the texts are taken from general newspapers, scientific, administrative and other more or less specialized sources and they include more and more transcripts of radio and television broadcasts and other types of formal and informal oral production. For other languages these corpora may be less rich and less publicly accessible, but many publishing houses do have corpora of several millions of words at their disposal.

At the level of the linguistic **description** of the contents of these immense riches, the last decade has seen a number of significant improvements. KWIC (Keyword in context) software, for instance, which makes it possible to oversee the use of a given word in various contexts, has become quite common and ever more sophisticated during recent years. Moreover, the theoretical notions and the practical techniques that are needed to interpret the language data are much clearer than ten or twenty years ago. The collaboration between experienced lexicographers like Sue Atkins and well known semanticists like Charles Fillmore, for instance, has led to a theoretically sound and practically effective tool that is generally known as the FrameNet approach (see *International Journal of Lexicography* 16.3, September 2003). Insight has been gained into questions concerning the distinction of senses of polysemous words, the status of collocations, idioms and other types of chunks, and there is a much greater awareness of the wide variety of data and phenomena covered by the term 'lexicon'.

As to the third field of lexicography, the **presentation** of the results of the foregoing steps in a dictionary, much less progress can be reported. As a matter of fact we still do not know what type of dictionary and what kind of lay-out is most convenient for which users or how we must proceed in order to improve the success rate in dictionary use, which is traditionally just above chance level.

A number of studies have been conducted on the relative importance of monolingual, bilingual or bilingualized dictionaries for L2 learners. The author of a recent book on this matter, Robert Lew, has investigated this point in a highly systematic way with Polish learners of English. Using a fine-grained design, where more than 700 learners use one of six types of dictionaries in order to find out the meaning of a number of pseudo-English words, the author finds that the monolingual dictionary (with English definitions only) scores far lower than any of the other types, and that some types of dictionaries, especially those with equivalents as well as definitions in English or Polish, tend to be less effective, probably due to an overload of information. In his conclusion Lew states that:

'we must question the validity of the recommendation so popular amongst educators of the presumed superiority of the monolingual dictionary. There is hardly any empirical evidence available to support that supposed superiority, and what little relevant evidence is available, points to the bilingual dictionary as the more effective dictionary for reception. The present study provides further evidence of this type' (Lew 2004: 179).

Another type of research has underlined the importance of lemma structures and the use of guide words, menus or sign posts, especially in long lemmas (Bogaards 1998). A study concerning the relative utility and usability of grammatical information reveals that traditional grammar codes are not used very often, even by advanced learners with a good linguistic schooling, but that syntactic information provided by definitions (like in Cobuild style definitions) and even more in examples helps learners in an effective way to write correct sentences (Bogaards and Van der Kloot 2002). Although other studies could be cited, really relevant experimental evidence on the aspect of presentation is

nevertheless still quite scarce.

As may be clear from the examples of experimental research given above, presentation is intimately linked to the intended users. In contrast to the formal uniformity of most dictionaries all over the world, there probably is not one correct way of presenting lexical data in a dictionary, one that suits all users. Ideally there should be a fairly large number of differently designed dictionaries from which the users could choose. Although this is not a weird idea in this time of computer facilities and on-demand business, it is not clear how learners could make their choice nor how teachers or dictionary compilers could guide them in this choice. A lot of thinking and a lot of experimental research are necessary before we can start compiling the dictionaries that users of different types would really need.

An example of dictionary use

In this paper I would like to reflect on some aspects of the information that learners minimally need when they have to produce a text in an L2. I will not make a distinction between the oral or written nature of such a text, although we all know that dictionaries are only seldom used in oral settings. In order to simplify the situation, I will take an imaginary native speaker of French who is a learner of English and who wants to say or write something in English and who is prepared to look up in the dictionary the elements he does not have at his disposal.

Let us imagine that he wants to express in English something like:

Notre professeur veut qu'on écoute bien ses conseils.

The first word that could cause a problem is *professeur*. We will assume that our learner is a 'collégien' (a schoolboy in a secondary school) and is aware of the fact that instructors on different levels may be called by different names, as is the case in French. He knows the word *teacher* but wants to verify in a learner's dictionary whether this is the right choice here. He opens one of them and finds a definition such as 'someone whose job is to teach' (MEDAL). Not really satisfied, he turns to a bilingual dictionary and finds:

professeur /pʁɔfɛsɔʁ/ [pʁɔfɛsɔʁ] teacher; léc (yɔs, tsWɛʃ) [school] teacher; (lɛvɛ) = lecturer (AmE), instructor (US), (avec chaire) professor + **elle est** - she's a [school] teacher + (Monsieur) le - Durand (lɛvɛ) Professor Durand + - de piano/de chant piano/singing teacher or master (lɛvɛ) or mistress (AmE) + - de droit lecturer in law, professor of law + l'ensemble des -s the teaching staff

Fig. 1: Part of the lemma **professeur** in Robert & Collins (2002)

If he is brave enough to fight his way through all the abbreviations, brackets and parentheses, this will give him the certainty that *teacher* is the right equivalent in his context. So he now can write:

Our teacher wants

But what does the teacher *want*? What is the right construction to use with this verb? For this point our 'collégien' goes back to the monolingual learner's dictionary, where he finds:

want /wɑnt/ verb [T] ★★
 1 to feel that you would like to have, keep, or do something: Do you still want these old letters? • **want sth for sth** What do you want for your birthday? • **want to do sth** Liz wants to see the circus. 1a. to feel that you would like someone to do something or would like something to happen: want sb/sth to do sth Her parents didn't want her to marry him. • **want sth of/from sb** I wish I knew exactly what they wanted of me. • **want sth done** Mrs. Miller wants the entire house repainted. 1b. to feel that you would like someone or something to be in a particular place or doing a particular job: I want those boxes out of the living room by tomorrow! • We want Stephen as chairman. 1c. to ask for someone because you would like to see or speak to them: Does anyone - she's in the kitchen. • You've wanted on the telephone. • The principal wants you in her office after school.
 2 **want sth** to need something: We desperately want rain. • You need to want much money on a camping trip. 2a. **Fr E** If something wants doing, it needs to be done: This bill wants paying. 2b. **Fr E** Informal to deserve something: That race! He wants a good kick up the backside!
 3 **spoken** used for telling someone what they should do: You want to take Pine Street - it's faster.
 4 to feel that you would like to have sex with someone: She wanted him more than ever.
 all sb wants **spoken** used for saying that someone's needs or requests are reasonable: All I want is the truth. • All they wanted to do was play football.
 be/have everything sb wants to have all the qualities that someone thinks a particular person or thing should

Figure 2: Part of the lemma **want** in MacMillan English Dictionary for Advanced Learners (MEDAL, 2002)

or

want (vɔ) [DESER] /wɑnt, \$wɑnt/ v to desire a particular thing or plan of action. Want is not used in polite requests in British English. • I want some chocolate. [T] • She wants a word with you. [T] • He's everything you'd ever want in a man - bright, funny and attractive. [T] • What do you want out of life? [T] • What do you want to eat? [+ infinitive] • What do you want to be when you grow up? [+ infinitive] • I want to be picked up at the airport at about nine o'clock. [+ to infinitive] • Do you want me to take you to the station? [T + obj + to infinitive] • This letter - do you want it sent first class? [T + obj + v-inf] • Do you want this job? [T + obj + adj] • I don't want a load of traffic going past my house all night, awaking me up. [T + obj + v-inf] • To want someone is also to wish them to be present. Johnny, your mother wants you on the phone. [T] • Am I wanted at the meeting tomorrow? [T] • He is wanted by the police (=The police are searching for him) in connection with the murder of teenager Peter Lewis. [T] • Want is not used in the present continuous tense although the continuous form of the verb is sometimes used in other tenses: I've been wanting to see that film for months! [+ to infinitive] • You want - she's been wanting her own horse next! [T] • (DyW) Someone who wants out of/in an activity or plan of action wants to stop or start being involved in it: I want out of the whole venture before it's too late. [T] • Want adj is Acc for classified as Se at CLASSIFY. • "I want to be alone" (Greta Garbo in the film Grand Hotel, 1932) • Compare LIFE [WANT].

Figure 3: Part of the lemma **want** in Cambridge International Dictionary of English (CIDE, 1995)

Although all the information that is needed can be found in each of these lemmas, it is far from certain that our friend will be able to winkle out the right elements. Alternatively, if he goes to the bilingual dictionary, he will be confronted with:

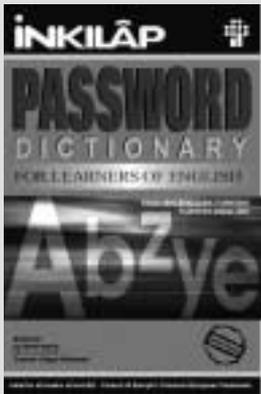
KERNERMAN DICTIONARY RESEARCH GRANTS

In 2002-2004

Kernerman Publishing and K Dictionaries made grants available for research in specified fields of lexicography, with funds distributed through AFRILEX, ASIALEX and EURALEX.

These grants have been discontinued, and new plans are being made on how to promote such research in the future.

From the series
KERNERMAN SEMI-
BILINGUAL DICTIONARIES



**PASSWORD
Dictionary for
Learners of English**

İNKILÂP KİTABEVİ
Istanbul, Turkey
September 2004
Editor: Önder
Renklyıldırım
1234 pages, 194 x 135 mm.
ISBN 975-10-2216-9
[http://kdictionary.com/
products/medium/pwtr.html](http://kdictionary.com/products/medium/pwtr.html)



**PASSWORD
Anglický výkladový
slovník so slovenskými
ekvivalentmi**

SPN
Bratislava, Slovak Republic
November 2004
Translation: Andrea
Cániková
Revision: Elena Monceková,
Eva Malá
861 pages, 240x170x60 mm.
ISBN 80-10-00459-6
[http://kdictionary.com/
products/medium/pwsk.html](http://kdictionary.com/products/medium/pwsk.html)

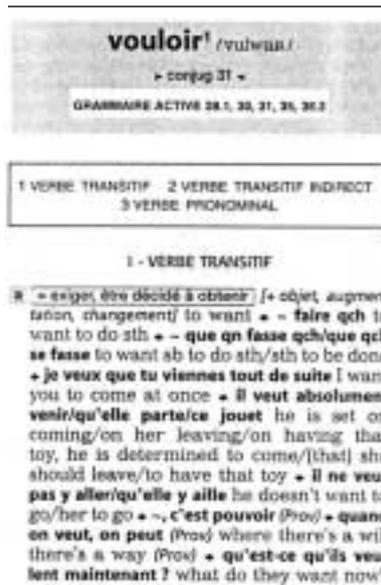


Figure 4: Part of the lemma **vouloir** in Robert & Collins (2002)

But again it is uncertain whether he will be any luckier with this presentation. Although daily practice may show otherwise, let us suppose that our hero eventually comes up with:

Our teacher wants us to

Let us see what happens with the rest of the sentence. *Écouter* = *listen*, quite simple, although the French verb is transitive (*écouter quelque chose*), whereas the English verb uses a preposition (*listen to something*). But what is the word for *bien*? *Listen well*? This does not sound very familiar. MEDAL gives a list of ‘Words frequently used with **listen**’: *attentively, carefully, closely, hard, intently, politely*, some of which can do the job perfectly well. I know of no bilingual dictionary that gives this type of useful information on collocations. So, if no mistakes are made, we now have:

Our teacher wants us to listen carefully to

Not knowing the word for *conseils*, our ‘collégien’ fights bravely on. Only the bilingual dictionary can bring him to the element he needs, which is *advice*. If he can interpret ‘NonC’ as meaning that this word does not permit a plural, he will finally write:

Our teacher wants us to listen carefully to his advice.

Had he checked in a learner’s dictionary, he would have found ‘noun [U]’ (MEDAL) or ‘n [U]’ (CIDE), but, on the positive side, he would have also found a number of examples to take advantage of.

Comments

What I have tried to make clear is that for this virtual dictionary user, the information

that is needed cannot be found in one dictionary and is not always tailored to his needs. I am convinced that the task of the dictionary user is systematically underestimated. Even if the information is there, there remains so much to be interpreted and to be adapted to the context at hand that many users give up or end up with incorrect solutions.

As will be clear, the information that was necessary for the ‘translation’ of the sentence given above was partly available in the **bilingual** dictionary. As these dictionaries have translation as their primary goal, they cannot give a full description or a reasonable number of examples in the target language. Yes, they present the user with the form of the element that is needed in the other language, but in most cases they give insufficient or at best highly coded information on how this element is to be used in the other language. In addition, they do not normally give phonetic transcriptions of the words of the target language at the place where these are most needed: at the right side of the translation equivalence. At the place where these transcriptions are now given, at the left side in the receptive dictionary, they will first of all be helpful for those who already know the word but have to pronounce it, e.g. for reading a text aloud. Those who want to use it for oral production, once they have found the right element, have to make a further step to the receptive part or volume in order to find out how to pronounce it.

The most serious drawback of **monolingual** dictionaries of a foreign language is that the user is often unable to retrieve the word he needs. Although more and more techniques and tools are being created to overcome this problem (see Bogaards 2003), in many cases learners will need a bilingual access mode in order to get at the words they need. A bilingual index to a monolingual dictionary, as has been proposed for some types of bilingualised dictionaries, is not a viable option. For polysemous words (and most frequent words are very polysemous) this would create a very bad kind of bilingual list where users need to make their choice on rather vague grounds and have then to turn to the dictionary proper, a step that many will not make.

Electronic dictionaries can remedy this situation to a great extent. But then they should not just be CD-ROM versions of paper dictionaries. They should be rethought in a fundamental way. Really new results can only be expected when such a rethinking not only takes into account the electronic opportunities that

are widely available nowadays, but, in addition, seeks to apply the results of sound experimental research concerning the behaviour of language learners and dictionary users in various situations and at different levels. At Van Dale (see Bogaards and Hannay 2004) we are taking some cautious steps along this very promising path. It will certainly be a long way.

References

- Bogaards, P. 1998. Scanning long entries in learner's dictionaries, in T. Fontenelle et al. (eds.), *Actes EURALEX '98 Proceedings*, Liège: Université de Liège, Vol. II, 555 – 563.
- Bogaards, P. 2003. MEDAL: a fifth dictionary for learners of English, *International Journal of Lexicography* 16.1: 43 – 55.
- Bogaards, P. and M. Hannay 2004. Towards a new type of bilingual dictionary, in G. Williams and S. Vessier (eds.), *Proceedings of the Eleventh EURALEX International Congress*, Lorient: Université de Bretagne-Sud, Vol. II, 463 – 474.
- Bogaards, P. and W.A. van der Kloot 2002. Verb constructions in learners' dictionaries, in A. Braasch and C. Povslen (eds.), *Proceedings of the Tenth EURALEX International Congress*, Copenhagen: Center for Sprogteknologi, Vol. II, 747 – 757.
- Lew, R. 2004. *Which dictionary for whom? Receptive use of bilingual, monolingual and semi-bilingual dictionaries by Polish learners of English*. Poznan: Motivex.

Dictionaries

- Cambridge International Dictionary of English (CIDE), 1995. Cambridge: Cambridge University Press.
- MacMillan English Dictionary for Advanced Learners (MEDAL), 2002. Oxford: Macmillan Education.
- Robert & Collins, 2002. Paris: Dictionnaires Le Robert.

Longman Słownik współczesny angielsko-polski polsko-angielski: The First Active Bilingual Dictionary for Polish Learners of English

Arleta Adamska-Sałaciak

The following is a brief characterisation of *Longman Słownik współczesny angielsko-polski polsko-angielski* (LSW).¹ After giving some general information about the dictionary, I shall focus on those features which make it the first active dictionary of the bilingual type on the Polish scene.

LSW is aimed at Polish learners of English. Its primary target audience are *gymnasium* (junior high) and *liceum* (high school) students, i.e., people in the 13-19 age group, with a command of English ranging from beginner to upper-intermediate. This does not mean the dictionary has nothing to offer to older or more advanced learners. On the contrary, the quality of the translations (which can only be appreciated when compared – by speakers fluent in both languages – with those in other local bilingual dictionaries) and the wealth of usage information (on which more below) make it a suitable tool also for more ambitious language tasks, especially of the encoding type.

The dictionary is corpus-based. For English, it relies on the Longman Spoken and Written English Corpus; for Polish,

a 10-million word corpus (90 per cent written, 10 per cent spoken) was gathered specially for this project.² The written part of the corpus consists of fiction (mostly for teenagers), fragments of textbooks in various school subjects, and newspaper and magazine articles. For the spoken part, everyday conversations were recorded (with the participants' permission) and subsequently transcribed.

In numerical terms, the scope of LSW is outlined in the table below.³

It should be noted that there are quite a few English-Polish / Polish-English dictionaries of comparable (physical) size. Many of them contain, or claim to contain, substantially more entries. However, these

Arleta Adamska-Sałaciak is head of the Department of Lexicology and Lexicography at the School of English, Adam Mickiewicz University in Poznan. Her academic interests include also: history of the English language, history and philosophy of linguistics, semantics, and English phonetics and phonology. She has co-authored and edited five bilingual dictionaries with English and Polish, including LSW.
arleta@ifia.amu.edu.pl

	English-Polish	Polish-English
headwords	20,554	17,964
run-ons	2,685	2,345
phrasal/reflexive verbs	1,167	777
fixed expressions	8,034	12,927
examples	25,105	33,075
senses	31,402	25,278
translations	39,308	44,240

Table: The scope of LSW in numerical terms³



**Longman Słownik
współczesny
angielsko-polski
polsko-angielski**

Jacek Fisiak, Arleta
Adamska-Sałaciak, Mariusz
Idzikowski, Ewelina Jagła,
Michał Jankowski, Robert
Lew
1039 pages
ISBN 0-582-447755
Pearson Education, Harlow
2004
[www.longman.com/
dictionaries](http://www.longman.com/dictionaries)

are all dictionaries with virtually no depth, where a typical entry consists of one sense with a single equivalent, or sometimes several equivalents with no indication of the differences between them. None of them is based on a Polish corpus. By contrast, the entries in LSW, especially those for high-frequency items, are more elaborate, with clear discrimination of senses and equivalents, rich collocational information, and carefully chosen, corpus-derived example sentences.

The English wordlist of LSW is largely co-extensive with that of the fourth edition of the *Longman Active Study Dictionary of English* (LASDE4), the main additions being recent neologisms (e.g. text message as both noun and verb) and a host of geographical names. The main source for the Polish word list was the Polish corpus, checked against the reversal of the English-Polish side and occasionally supplemented by material drawn from the most recent monolingual dictionaries of Polish.

Although the English-Polish side started out as a bilingualisation of LASDE4, the result is not a semi-bilingual dictionary like, e.g., the Polish version of *Oxford Wordpower* (OW) – a publication similar in size and reasonably popular with Polish students.⁴ In the case of LSW, English definitions have been removed and whole entries reorganised (through splitting and merging of senses), so that what we get is a new bilingual dictionary, with precise contextual equivalents of the different senses of the headwords rather than just broad translational hints typically found in semi-bilingual dictionaries.

All illustrative examples in LSW are English ones. This means that, contrary to expectations, the examples in the Polish-English part are in the dictionary's target language: English, not Polish. The rationale for that is quite simple. LSW is a directional dictionary, in particular, it is explicitly addressed to Polish learners of English, not to native speakers of both languages. It can reasonably be assumed that Poles do not need to be told how Polish words are used, or at least, that it is not the kind of information for which they need a dictionary of English. It thus seemed to make more sense to provide them with English sentences and sentence fragments illustrating the various translations of the individual senses of Polish headwords. This facilitates the retention of new items by repeated exposure; at the same time, the user is given immediate confirmation that the English equivalent that they are offered really works in a particular context. Most importantly, learners can use the examples as models for their own production in

English. The presence of L2 examples in the L1-L2 part of the dictionary eliminates the need to go to the L2-L1 part, or to consult a monolingual dictionary of L2, each time the learner wants to use a newly found L2 item in their linguistic production. Thus, the main criticism levelled against traditional bilingual dictionaries, i.e., that they offer little or no help in encoding, does not apply to LSW.

Like the examples in the English-Polish part, those in the Polish-English part also come from the Longman Spoken and Written English Corpus and from the Internet. This being the case, one might wonder what the role of the Polish corpus was in the project. For one thing, it served as a basis for extracting the Polish word list, proving to be an especially rich source of collocations and other fixed expressions, often too recent or too colloquial (but nonetheless frequent, especially among young people) to feature in the traditionally conservative monolingual dictionaries of Polish. The Polish corpus was also critical in deciding upon sense discrimination and the ordering of senses within entries. Finally, better English equivalents could be provided thanks to the large number of collocates found in the Polish corpus: by comparing groups of collocations across the two languages it was possible to fine-tune the translations.

As a rule, the English examples are not translated into Polish, except in cases where a structural difference between the two languages makes word-for-word translation impossible, thus posing too big a challenge for the user. Even then, it is normally only the difficult fragment itself, and not the whole sentence, that is translated. The reader might wonder what *difficult* means here, or rather: *difficult for whom?* Since the target audience of LSW encompasses learners of different proficiency levels, attempts were made to exercise some control over the complexity of examples. In particular, whenever the corpus afforded a choice, the simplest possible sentences were opted for as illustrations of basic senses of common words, the assumption being that the entries or entry fragments in question would be of interest primarily to beginners. If no simple example was available, rather than making one up, we took an authentic sentence from the corpus (sometimes shortening it) and provided a translation in brackets.

The practice of leaving most examples untranslated is one of the features distinguishing LSW from another bilingual Longman publication, the *Longman Dicionário Escolar inglês-português*

português-ingles (LDE). That dictionary gives English examples in the English-Portuguese part, Portuguese examples in the Portuguese-English part, and translates all of them. The example sentences, which come from corpora of English and Brazilian Portuguese (Harmer 2002: iv), appear to have been considerably simplified. The most likely reason is that LDE seems to be aimed primarily at beginner and pre-intermediate Brazilian learners of English, a feature which is also reflected in the dictionary's coverage.⁵

LSW has a number of extra features which are meant to assist in language learning and at the same time increase the book's attractiveness. An insert in the middle contains a section on life in the UK and the US, a communication skills bank (with expressions used in various everyday situations, such as giving advice or extending an invitation), a grammar section (with information on English tenses, the passive, reported speech, etc), a false friends section, and the mandatory list of irregular verbs. The end matter features, among other things, a list of popular e-mail and text-message abbreviations with their Polish equivalents.

Interspersed among entries in the main body of the dictionary are a few dozen grammar boxes (e.g. on the English articles, individual modal verbs or irregular plurals) and well over six hundred usage notes. The latter focus on important lexical and grammatical differences between English and Polish, on meaning and register differences between similar and/or frequently confused English items, as well as on instances where no exact equivalent exists in the target language for a given source language item. In general, the idea was to highlight points of special difficulty for the Polish learner. The authors of LSW have all at some point taught English to Poles. Being native speakers of Polish, we have also learnt English as a foreign language ourselves – and are still learning it. Thanks to this insider's perspective, we hope to have come up with notes that are of more relevance to Polish learners of English than the one-size-fits-all usage advice dispensed by monolingual dictionaries.

LSW is the second Longman dictionary produced by a team of lexicographers at the School of English of Adam Mickiewicz University in Poznan. Its predecessor, *Longman Podreczny Słownik angielsko-polski polsko-angielski* (LPS), came out in 1999. In terms of the number of entries, the English-Polish side of LPS was about a quarter smaller; its Polish-English side was essentially a Polish-English index.

Predictably, experience gained in the preparation of LPS proved invaluable in the course of work on the later project.

To those familiar with recent developments in the area of bilingual pedagogical lexicography, the dictionary presented here may not appear particularly revolutionary or even innovative.⁶ However, in the Polish context it is a huge step forward. We are confident that LSW will serve its users well. One prerequisite for that is that Polish teachers of English must forget what they themselves have been taught, namely, that all bilingual dictionaries are merely poor relations of monolingual ones. The reactions we have had so far, from teachers and students alike, are encouraging. We would like nothing better than for our dictionary to set a standard aimed at by any new arrival on the Polish dictionary scene which aspires to be taken seriously.

Notes

1. I am grateful to Mariusz Idzikowski, Ilan Kernerman and Robert Lew for commenting on drafts of this paper.
2. The corpus is the property of Pearson Education Ltd.
3. The apparent discrepancy in the number of headwords and fixed expressions between the two sides is a result, mainly, of the treatment of Polish noun+adjective compounds of the type *stacja benzynowa* "petrol station". Such compounds are nested in the entries for their headnouns (and included in the above count as fixed expressions), whereas English compounds head separate entries. All run-ons (adjectival adverbs, deverbal nouns, etc) are accompanied by their L2 equivalents. All English phrasal verbs and most Polish reflexive verbs (except those with no corresponding non-reflexive form) are nested within entries for the respective main verbs. If we count run-ons and phrasal/reflexive verbs, the actual number of entries is closer to 24,000 (E-P) and 21,000 (P-E).
4. The first semi-bilingual work of reference in the English-Polish context was the Kernerman dictionary of 1990, followed by its 1996 and 2002 incarnations (see EDSP, EPLD and PS in the References). The later arrival, OW (1997, 2002), has since dominated the semi-bilingual dictionary scene in Poland.
5. These remarks are not meant as an exhaustive comparison of the two dictionaries, but merely as a brief indication of the differences in their design and execution. This may give a better idea of LSW to readers who are familiar with LDE.

From the series
KERNERMAN SEMI-
BILINGUAL DICTIONARIES



PASSWORD Dicționar Englez- Român

EDITURA NICULESCU

Bucarest, Romania

May 2005

Editing: Liliana Scarlat,

Anca Florea, Dragos Trifan

Translation: Stefan Firica,

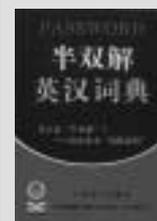
Alexandru Farcas, Raluca

Dumitru

ISBN 973-568-710-0

<http://www.niculescu.ro>

<http://kdictionaries.com/products/medium/pwro.html>



PASSWORD Semi-Bilingual English-Chinese Dictionary

SHANGHAI

LEXICOGRAPHICAL

PUBLISHING HOUSE

Shanghai, PR China

May 2005

Editor: Ling Zhiguo

Translation: Wu Hao, Song

Xiaoman, Ling Zhiguo

1080 pages, 140 x 203 mm.

ISBN 7-5326-1670-3

www.cihai.online.sh.cn

<http://kdictionaries.com/products/medium/pwzh.html>

6. See, e.g., Back (2004) or other papers in the Bilingual Lexicography section of Williams and Vessier (2004).

References

Dictionaries

English Dictionary for Speakers of Polish. 1990. Toronto: Kernerman Publishing Inc. and Poznan: Kantor Wydawniczy SAWW. (EDSP)

English-Polish Learner's Dictionary. 1996. Warszawa: Wydawnictwo Naukowe PWN and Password Publishers. (EPLD; a revised and updated version of EDSP)

Longman Active Study Dictionary of English, 4e. 2000. Harlow: Pearson Education Limited. (LASDE4)

Longman Dicionário Escolar inglês-português, português-inglês (para estudantes brasileiros). 2002. Harlow: Pearson Education Limited. (LDE)

Longman Podreczny Słownik angielsko-polski polsko-angielski. 1999. Jacek Fisiak, Arleta Adamska-Safaciak,

Mariusz Idzikowski and Michał Jankowski. Harlow: Pearson Education Limited. (LPS)

Oxford Wordpower: Słownik angielsko-polski z indeksem polsko-angielskim. 1997 (2e, 2002). Oxford: Oxford University Press. (OW)

Praktyczny Słownik angielsko-polski, polsko-angielski. 2002. Warszawa: Wydawnictwo Naukowe PWN S.A. and K Dictionaries Ltd. (PS; a reprint of EPLD)

Other references

Back, Martyn. 2004. "A new bilingual learner's dictionary format: The Junior Bilingual". In Williams and Vessier (eds.) 451-455.

Harmer, Jeremy. 2002. "Bilingual breakthrough". (Preface to the LDE, pp. iv-v).

Williams, Geoffrey and Sandra Vessier (eds.) 2004. *Proceedings of the Eleventh International EURALEX Congress*. Lorient: EURALEX.

chłopczyk *n* (little) boy

chłopiec *n* **1** boy; a school for boys **2** (sympatia) boyfriend — **chłopięcy** *adj* boyish: a slim, boyish figure

chłosta *n* flogging — **chłostać** *v* flog: Thieves were flogged in public.

chłód *n* **1** (zimno) chill: There was a chill in the early morning air. **2** (obojętność) coldness, coolness: He detected coolness in her tone.

chmiel *n* hops

chmura *n* **1** cloud: stormy clouds **2** **chmura dymu/kurzu** *itp.* cloud of smoke/dust etc

chmurzyć *v* **chmurzyć się** cloud over: The sky's really clouding over; I think we're in for a storm (=czeka nas burza).

chochlik *n* **1** (duszek) gremlin, goblin, imp **2** także **chochlik drukarski** misprint, typo

choć *conj.* **part** **1** **patrz** **choć** **2** **choć** **coś** that's something: At least we have some money left. That's something, isn't it (=choć coś, nie)?

choćaby *conj.* **patrz** **choć**

choć *conj.* **part** **1** (mimo że) (even) though, although: I seem to keep gaining weight, even though I'm exercising regularly. | Although it was raining we decided to go for a walk. **2** (ale) though: I don't really like classical music, though I did enjoy the Pavarotti concert. | I think she's Swiss, I'm not sure though. **3** (przynajmniej) at least: I think you should at least consider his offer (=powinnoś choć rozważyć jego propozycję). **4** **choć raz** (just) for once: Just for once I'd like to see him cook dinner.

choćby *conj.* **part** **1** even: He didn't want to stay, even for a while (=choćby na chwilę). **2** (nawet gdyby) even if: I'll never speak to her again, even if she apologises (=choćby (nawet) mnie przeprosiła). **3** **choćby nie wiem co/gdzie itp.** no matter what/where etc: No matter how hard he tried (=choćby nie wiem jak się starał), he couldn't get her to change her mind.

chodak *n* clog

chodnik *n* **1** (na ulicy) pavement *BrE*, sidewalk *AmE* **2** (dwan) runner

cholera *n* **1** (przekleństwo) shit, damn, goddammit *AmE* **2** (choroba) cholera

cholerny *adj* damned, bloody *BrE*: That bloody fool, Hodges! — **cholernie** *adv* damn, bloody *BrE*: damn good

cholesterol *n* cholesterol

chomik *n* hamster

chorągiew *n* flag — **chorągiewka** *n* flag: The children waved their flags as the Queen went by.

chorąży *n* warrant officer

choreografia *n* choreography — **choreografika** *n* choreographer

choroba *n* **1** disease, illness, sickness: He suffers from (=cierpi na) heart disease. | She missed a lot of school because of her illness. | an insurance policy against long-term sickness **2** **choroba lokomotywna** travel/motion *AmE* sickness **3** **choroba morską** seasickness: **cierpieć na chorobę morską** be seasick

illness, disease, sickness

UWAGA

illness odnosi się zasadniczo do stanu złego samopoczucia i długości jego trwania, **disease** natomiast do przyczyny, która ten stan wywołała. Mówimy więc: She died after a long illness. | How many working days have you missed through illness? | She suffers from a rare disease of the central nervous system. Gdy mowa o nazwach chorób, narządach nimi dotkniętych czy sposobach ich przenoszenia, używamy wyrazu **disease**: Alzheimer's disease | a kidney disease | a sexually transmitted disease | infectious disease. Wyjątkiem jest choroba umysłowa (mental illness), jak również choroba śmiertelna lub bardzo ciężka (terminal/critical/serious illness). Rzeczowniki **sickness** można w powyższych kontekstach używać wymiennie z **illness**: working days lost due to sickness. Ponadto **sickness** oznacza też 'mdłości', w związku z czym występuje w nazwach chorób, które się w ten sposób objawiają, takich jak np. choroba morską. **Sickness** oznacza również chorobę w sensie przonałnym: He said the idea of 'success' was part of the sickness of Western culture.

Extract from the Polish-English part of
Longman Słownik współczesny angielsko-polski polsko-angielski

Bilingual Dictionaries for Learners

Martyn Back

In 2002 Dictionnaires Le Robert produced a new bilingual dictionary designed for pre-intermediate learners of English, the Junior Bilingue (since renamed First in English). Unlike most standard bilingual dictionaries available in France, it is asymmetrical (the “encoding” and “decoding” sides are presented in different ways) and designed specifically for French people learning English (ie, all the metalanguage is in French and the text is built around known problems encountered by learners). The dictionary is now widely used in French schools, and similar texts have been produced for French native learners of Spanish and German, and (published by Vox) Spanish native learners of French and English. What follows is an outline of some key issues that determined the editorial orientations of this project.

1. A production dictionary

This book was designed first and foremost as an aid to target language production – what bilingual lexicographers term “encoding”. The encoding dimension in standard European bilingual dictionaries has been steadily systematised since the publication of the Collins-Robert in the late seventies (1978, 6th edition 2002), with its network of synonymic and contextual indicators, its emphasis on collocational patterns and its wealth of carefully selected, corpus-based example sentences. But Collins-Robert and its rivals are reference works designed for relatively sophisticated users. Their entries are highly coded, much of the information they give is abbreviated or implicit, and they assume a good deal of prior knowledge on the part of their readers (at the very least, an acquaintance with the “how to use” section at the beginning of the book). Using a bilingual dictionary as an encoding tool is a tricky business, even for the experienced user. The encoding user is always, to a greater or lesser extent according to his or her level of linguistic competence, stepping into the unknown; translators know that to lift a foreign language term from a dictionary without further cross-checking is fraught with danger, and most teachers have anecdotes about the hilarious misuse of dictionaries. A user-friendly dictionary – and a fortiori a bilingual learners’ dictionary – must do all it can to reduce this risk factor, lighting the reader’s way as he or she gropes towards proficiency.

The Junior attempts to achieve this in several different ways. Instead of baldly presenting the target-language equivalent of a given word, it first shows the headword ‘in action’ – in a translated example sentence with no distracting metalanguage. The example sentences are designed to emphasize points relating to the headword equivalent (prepositional collocations of verbs, uncountability of nouns, use of articles, grammatical behaviour) which students need to understand and learn. Facing these example sentences, in a separate column, we find the translation of the headword used in the example, followed by notes in French that draw attention to specific points (register, syntax, pronunciation and so on). The presence of this second column means that information that is implicit in standard dictionaries can be made explicit for the learner. In standard dictionaries, for example, French numerals (*deux, quinze*) are translated by English numerals (*two, fifteen*), and the reader is expected to know that where dates are concerned English uses ordinals (the second, the sixteenth). In the Junior, this is made explicit via example sentences and an accompanying note in the right-hand column. The two-column layout is an effective way of uncluttering the dictionary entry and clarifying metalinguistic commentary.

2. Focusing on essentials: the reception/decoding side

Space constraints, lexicographical conventions (not all of them indispensable, or even particularly useful) and economic realities (bilingual dictionaries typically being expected to pay their way in two linguistic communities at once), mean that at least some of the information provided in a bilingual entry is likely to be irrelevant to a given user. Worse, such a surfeit of information can present an obstacle to understanding. The example below illustrates this:

black 1. *n* (= colour) noir (m). 2. *adj* noir. ♦ Black (= person) Noir(e) m(f).

This classically well-formed one-line entry is both basic (in terms of the rudimentary lexical information it contains) and complex (in terms of the highly coded way in which the information is delivered). Riddled as it is with metalinguistic codes, abbreviations and symbols – e.g. no fewer than five pairs of brackets – it contains a

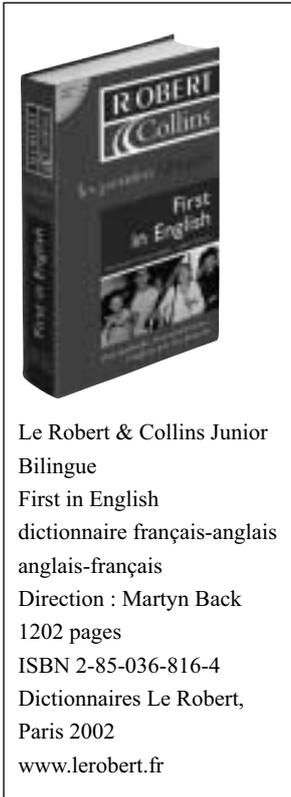


Martyn Back was born near Stratford-Upon-Avon, England. He studied French at Lancaster University and decided to live in Paris for a while when he graduated. 23 years later he has still not found a good reason to leave.

After several years teaching English, translating and acting, he began working on large bilingual dictionaries, first at Larousse, then at Dictionnaires Le Robert, where he is now editorial director in the bilingual department. In 2002 he created the Junior Bilingue, a successful new learners’ bilingual dictionary.

When he is not working, Martyn can be found cooking for friends, hiking through various wildernesses (lately the Sahara in Mauritania and Libya) or chugging along the Mediterranean coast in an old Provençal fishing boat moored in La Ciotat.

mjback@lerobert.com



good deal of superfluous information for the French native who just wants to know what “black” means. Dividing the entry into noun and adjective is only useful for the English-speaking user (it makes it possible to indicate the masculine gender of the French noun). The indicator (= colour) is designed to inform the English user that other senses of the noun black are not covered here. The sub-entered capitalized form Black, with its special symbol and accompanying synonymic indicator, is there to introduce the capitalized French translation and to remind the English reader that “a Black (person)” is either un Noir (if a man) or une Noire (if a woman). But for the French decoding user most of this information is redundant, and the following presentation (adopted in the Junior) is perfectly adequate:

black noir.

Here is another example of how classic bilingual editing can over-egg the pudding for decoding users:

eighteen 1. *adj* dix-huit inv 2. *n* dix-huit m inv 3. *pron* dix-huit.

In the Junior this becomes, with no loss of vital data:

eighteen dix-huit.

3. Translated examples in a learners’ bilingual dictionary

The example sentence in a bilingual dictionary can (in its most banal guise) exist simply to illustrate a given headword translation; it can introduce important collocational information (*we took part in the show = on a participé au spectacle*); or it can show a contextual nuance for which a new translation is required. In “reference” bilingual dictionaries, as in advanced learners’ monolinguals, example sentences help present an accurate “snapshot” of word behaviour, and corpus tools provide examples that have the stamp of authenticity. Phrase translations tread a careful line between naturalness and “generativity” (ie, they must be idiomatic, yet sufficiently banal in stylistic terms to enable the user to re-use them confidently).

In bilingual dictionaries designed for pre-intermediate learners, the status of the example sentence is somewhat different. Examples exist primarily not to provide a snapshot of headword behaviour, but as pedagogical devices. Although natural and idiomatic, they are designed above all to generate a translation that shows the headword equivalent in action. The words they contain are carefully chosen to avoid complicating the issue for the reader, and in

most cases they are deliberately designed to make a point. At *mai/May*, for instance, there is an example sentence that in most reference bilingual dictionaries would be considered superfluous: *Stéphanie est née le trois mai / Stéphanie was born on the third of May*.

This apparently banal example sentence in effect “teaches” three important things. Its primary function is to show the use of the prepositions *on* and *of* with month names. Secondly, it shows the capitalization of the month name in English (compared to lower case in French). It also (though more incidentally) takes the opportunity to show *elle est née / she was born*, which many French native speakers have difficulty with.

Example sentences in the Junior are not drawn from corpora. The level of language in our French lexical corpus is inappropriate (much of its content consists of newspapers such as *Le Monde*, novels and magazines), and it was important for editors to “fine-tune” their examples to maximize their pedagogical value. Corpora provide an accurate and full picture of a language; pre-intermediate learners need a more focused and simplified view. The same is true of phrase translations: editors were encouraged to produce translations that are grammatically well-formed, but to avoid “over-idiomatic” language. At this level, the goal of language learning is to produce intelligible, grammatical sentences, not to imitate the idiosyncrasies of native speakers.

4. Asymmetry

The clear distinction between the functions of encoding and decoding, language production and comprehension, tends to be blurred in the classic bilingual text, which attempts to square the circle by providing information for four user types at once (ie encoding and decoding users in each of the two languages) in a text whose two sides more or less mirror each other. In a learners’ dictionary, where the emphasis is on production, the encoding and decoding sides can (and should, ideally) be treated differently, with lots of examples and guidance on the encoding side and a much more summary presentation in the decoding section. In Junior Bilingue, as we have seen, pages on the encoding side are divided into two parallel columns, with headwords and examples on the left hand side and explicit annotations (in French) on the right. The decoding side, on the other hand, is arranged like a pared-down traditional bilingual text, trimmed of all information designed for the “other” user (including systematic exemplification –

Power to the Learner: an Approach towards Pedagogically-Oriented Bilingual Dictionaries

Wolfgang Worsch



Wolfgang Worsch was DAAD-Lektor at Birmingham University, and lexicographer for Collier Macmillan and OUP. In the last twenty years he has been working as a fulltime lexicographer at Langenscheidt, as a freelance in several Duden dictionaries, and as a lecturer at the Munich University of Applied Sciences teaching lexicography and terminology. His major roles in recent projects include managing editor of Two-in-One Dictionary of English, co-author of Langenscheidt Power Dictionary and managing editor of the new edition of Muret-Sanders Großwörterbuch German-English.
w.worsch@langenscheidt.de

Introduction

Change is not something that people tend to associate with dictionaries. [...] The heavy cost of dictionary production, and the penalty to be paid for errors of judgement, have made it almost impossible for any radically new dictionary to come into being. (Atkins 1996)

A decade ago, Sue Atkins' rather pessimistic sounding statement was not an isolated one – on the contrary. Numerous meta-lexicographers expressed their doubts as to the possibility of bringing really innovative approaches into the vast field of dictionary making, such as Zöfgen, who wrote

It is therefore all the more surprising that in a market saturated with language learning material a bilingual dictionary developed along the lines of pedagogical lexicography and especially designed for the foreign learner is [...] still not available. (Zöfgen 1995)

Looking back at the development in the German market over the past decades, one is tempted to agree with Atkins' and Zöfgen's observations. Of course there were some "new" dictionaries – published by well-known publishers – with titles like *Schulwörterbuch* or *Schülerwörterbuch* [school dictionary]; in fact, these turned out to be more or less conventional dictionaries with some "add-ons" such as warning-signs against the use of false friends, an additional appendix, and so on.

In a way, Atkins was right when she talked about the financial risks that prevented publishers for a long time from pursuing radically new ways of presenting languages in bilingual dictionaries. Traditionally accepted principles of organizing the macro – as well as micro – structure were adhered to. Therefore, competitors had little more to show off with than the number of entries, the number of translations and some catchy neologisms. Quantity seemed to dominate quality.

Also, on the academic level, in-depth reflections and discussions on the topic of bilingual learner-oriented lexicography did not – to my knowledge – start until the early nineteen eighties.

As regards learner's dictionaries,

reflections and discussions by lexicographers as well as meta-lexicographers were of course influenced by the work and results already accomplished worldwide. And yet again, the main interest did not focus on bilingual dictionaries but on monolingual ones – the following definition speaks for itself:

A learner's dictionary is a synchronic monolingual dictionary intended to meet the demands of the foreign learner. (Herbst 1995)

Along with the prevailing didactic approach of teaching a foreign language via the foreign language, the main course of discussion followed – for quite a long time – the monolingual path rather than the bilingual one.

Thanks to institutions such as EURALEX, the relevance of bilingual dictionaries for foreign language acquisition was re-discovered and is now widely recognized, and lexicographers worldwide (either in the function of dictionary-makers working for a publisher or as meta-lexicographers working at a university) have taken cognizance of the need for bilingual learner's dictionaries that meet the requirements of the users.

Or to put it differently: discussions and reflections on monolingual English learner's dictionaries over the past few decades have significantly contributed to the high quality represented in monolingual dictionaries for advanced learners of English, which set international standards for the genre of monolingual learner's dictionaries. In Germany it was Langenscheidt publishing house who were thus determined to maintain the high standard set up by these and other titles when they decided to produce the *Großwörterbuch Deutsch als Fremdsprache*.

Towards the Bilingual Learner's Dictionary

In the 1990s, renowned international dictionary publishers added a new facet to the monolithic block of monolingual learner's dictionaries: the *semi-bilingual learner's dictionary*. OUP/Cornelsen published *Oxford Grundwörterbuch* [Basic Dictionary] *English-German / Deutsch-Englisch* and Langenscheidt introduced *Two in One Dictionary of English*, along

lines set outside Germany by the *Password* series of Kernerman Semi-Bilingual Dictionaries.

Langenscheidt Power Dictionary Englisch – Deutsch / Deutsch – Englisch

In 1993 a working-group was established at Langenscheidt which was to discuss and design ways of improving one of its best-selling books, the well-known *Schulwörterbuch*, hence the provisional working-title *Didaktisiertes Schulwörterbuch*, or DSW for short. On the team were not only lexicographers, but also experienced teachers and pedagogical experts from universities and teacher-training seminars.

The target-group was identified as German-speaking learners from beginner to intermediate level. Although we were primarily aiming at school students, we realized from the start that the concept should also work for ELT in adult education.

During several brain-storming weekends it became clear that the mere incorporation of new features would get us nowhere. It was the structure as such, the lexicographic identity, which had to be changed. The idea began to invade our minds that we would have to get rid of old and convenient lexicographical conventions if we were to achieve anything. Slogans such as *Less is More!*, *Kill the Swung Dash!*, *No more v/ recip and v/impers!*, *Phonetics for Translations!* and so on covered pinboards and flipcharts. At the end we had a concept to start with, a concept that of course was changed and adapted in the course of compilation, which again was accompanied by permanent discussions in the team. It was a concept which – and we were absolutely convinced about that – would lead us to a dictionary unprecedented in the history of German lexicography.

The main characteristics

- We followed the topos of *less is more*, which means we decided on a lean dictionary concept – a reduction in the number of headwords in favour of an extended, user-oriented microstructure.
- No more swung dashes – each compound or derivative, each headword that is repeated in a phrase or example sentence, is given in full.
- Each headword is given full lemmatization – no more nesting of compounds or derivatives.
- The headwords are printed in ‘Langenscheidt blue’ – phrasal verbs follow after the base-verb in a blue frame.
- In the German-English section, entries that changed considerably due to the new

German orthography are given in a blue frame.

- No more cryptic abbreviations for grammar, but clear, explicit meta-language or phrases and examples to show typical grammatical structures.

• The dictionary is tailored to meet the varying needs of the users depending on the individual context:

- English-German for decoding:

This part concentrates on what students might hear or read. Here they can find a large number of headwords with a lot of idiomatic contemporary vocabulary.

- German-English for encoding:

This part concentrates on what students wish to express in spoken or written English. Here they find a selective range of headwords for active use. For the first time, L2-phonetics are given in the L1-L2 section of a bilingual dictionary.

- usage notes:

A wealth of usage notes were especially designed to meet the needs of the German-speaking users.

- illustrations:

A new kind of contextualized, lively colour illustrations help to extend the learner’s vocabulary and offer scope for classroom work.

- Bearing in mind that there is more to language learning than just grammar and vocabulary, i.e. cultural and communicative competence, we introduced *Info-Fenster* [info boxes] with additional information on language phenomena, cultural or political topics, and other subjects of relevance to learners.

Conclusion

The feedback on the *Power Dictionary* has been overwhelming. In fact, soon after publication we received inquiries from other European dictionary publishers who showed an interest in producing English learner’s dictionaries for their languages in line with the *Power* concept.

Inevitably, we pursued the idea of establishing a *Power Family*, including a range of bilingual learner’s dictionaries for the main European languages. So, the second in this line was *Power Wörterbuch Französisch* (French), published in 1999, followed by *Power Wörterbuch Spanisch* (Spanish) in 2004 and *Power Wörterbuch Italienisch* (Italian) appearing this year. The English and French *Power* dictionaries underwent revisions in 2002 and 2003 respectively.

At the same time, some features of the *Power* concept were incorporated into revised editions of other Langenscheidt dictionaries: blue headwords, info boxes and other didactic features can be found



Langenscheidt Power Dictionary Englisch Englisch-Deutsch Deutsch-Englisch

Konzeption: Vincent J.

Docherty

Projectleitung: Wolfgang
Walther

1078 pages

ISBN 3-468-13113-5

Langenscheidt, Berlin and
Munich 2002

www.langenscheidt.de

in the *Großes Schulwörterbuch* [Concise School Dictionary] as well as in the *Taschenwörterbuch* [Standard Dictionary] series. Even the complete revision of the *Muret-Sanders Großwörterbuch Deutsch-Englisch* [Muret Sanders Comprehensive German-English Dictionary] was infected by the "Power virus".

However, we perfectly realize that the *Power* concept cannot be applied to all kinds of dictionaries. Tourists want, for example, a pocket dictionary with plenty of headwords and a good choice of travel-related vocabulary. Professional users, like translators, need a plethora of

special vocabulary, secretaries may want a dictionary with typical office-oriented phraseology, etc.

In the future, Langenscheidt will continue to offer a wide range of dictionaries in order to meet the requirements of as many users as possible. But with the *Power* concept we are convinced that we have achieved a major breakthrough in bilingual learners' lexicography.

Dictionaries

Das Oxford Grundwörterbuch Englisch-German/Deutsch-Englisch. 1990. Oxford: OUP/Cornelsen.

Kernerman Semi-Bilingual Dictionaries, PASSWORD. 1986-2005. various locations and publishers.

Langenscheidt's Großes Schulwörterbuch. 1999. München: Langenscheidt.

Langenscheidt's Großwörterbuch Deutsch als Fremdsprache. 1993, 2003. München: Langenscheidt.

Langenscheidt's Muret-Sanders Großwörterbuch Deutsch-Englisch. 2004. München: Langenscheidt.

Langenscheidt's Power Dictionary Englisch-Deutsch/Deutsch-Englisch. 1997, 2002. München: Langenscheidt.

Langenscheidt's Power Wörterbuch Französisch-Deutsch/Deutsch-Französisch. 1999, 2003. München: Langenscheidt.

Langenscheidt's Power Wörterbuch Italienisch-Deutsch/Deutsch-Italienisch. 2005. München: Langenscheidt.

Langenscheidt's Power Wörterbuch Spanisch-Deutsch/Deutsch-Spanisch. 2004. München: Langenscheidt.

Langenscheidt's Two in One Dictionary of English. 1997. München: Langenscheidt.

References

Atkins, Sue. 1996. Bilingual Dictionaries – Past, Present and Future; in *Euralex Proceedings Göteborg 1996* (2 Vol), Vol 2, p.515).

Herbst, Thomas. 1995. (Thomas Herbst, Dictionaries for Foreign Language Teaching: English; in *Wörterbücher*, Vol II, p.1379).

Zöfgen, Ekkehard. 1996. (Ekkehard Zöfgen, Bilingual Learner's Dictionaries; in *Wörterbücher*, Vol II, p.2888).

moon

moon¹ ['mu:n] *salopp* den nackten Hintern vorzeigen

moon about oder around [ˌmu:n əˈbaʊt oder əˈraʊnd] *mit. zu Hause*: lustlos herumlungern

moonlight¹ ['mu:nlaɪt] Mondlicht, Mondschein

moonlight² ['mu:nlaɪt] *umg.* schwarzarbeiten, Ⓢ pfuschen

moonlighter ['mu:nlaɪtə] *umg.* Schwarzarbeiter(in), Ⓢ Pfuscher(in)

moonlit ['mu:nlɪt] mondhell; moonlit night Mondnacht

moor [mʊə] *Landchaftsform*: Hochmoor

moose [mu:s] *Pl.*: moose *Hirschart in Nordamerika*: Elch

mop¹ [mɒp] 1. zum Wischen: Mopp 2. auch mop of hair Mähne

mop² [mɒp] mopped, mopped wischen, abwischen

mop up [ˌmɒpʊp] aufwischen

moped ['mɒpɪd] Moped

moral¹ ['mɒrəl] moralisch; moral obligation moralische Verpflichtung; moral

312

noch etwas, noch etwas mehr; can I have a little more? kann ich etwas mehr haben?; what more do you want? was willst du denn noch?

moreover [məʊəvə] außerdem, überdies

morning ['mɔ:niŋ] 1. (= Tagesbeginn) Morgen; good morning! guten Morgen!; in the morning morgens, am Morgen; early in the morning frühmorgens, früh am Morgen; this morning heute Morgen; tomorrow morning morgen früh

2. vor 12 Uhr: Vormittag; in the morning vormittags, am Vormittag; this morning heute Vormittag; tomorrow morning morgen Vormittag

Morocco [mə'reɪkoʊ] Marokko

Moroccan¹ [mə'reɪkən] marokkanisch

Moroccan² [mə'reɪkən] Marokkaner(in)

morose [mə'reɪs] mürrisch

Morse code ['mɔ:s kəʊd] Morsealphabet

morsel ['mɔ:sl] 1. Bissen, Happen 2. übermager Quälchen

mortal¹ ['mɔ:tl] 1. Mensch: sterblich 2. Verletzung *u.w.*: tödlich (to für) 3. mortal fear Todesangst; mortal sin Todsünde; mortal enemy Todfeind

mortal² ['mɔ:tl] Sterbliche(r); we ordinary mortals humorvoll wir gewöhnlichen Sterblichen

779

Komplott

comfortabel 1. Hotel, Wohnung: well-appointed, nur hinter dem Subst.: with all (the) conveniences (BE auch mod cons)

2. (= luxurious) luxurious [lʌg'zjuəriəs]

Komik humour ['bjʊəmə]

Komiker(in) comedian [kə'mi:diən], comic

komisch 1. funny (auch im Sinn von merkwürdig) 2. das Komische daran ist the funny thing (about it) is

Komitee committee [kə'mi:ti] (Δ. Schreibzug)

Komma 1. comma; hier fehlt ein Komma there's a comma missing here 2. drei Komma vier (3,4) three point four (3.4) (Δ. mit Punkt geschrieben) 3. null Komma drei (0,3) (nought, AE O [əʊ]) point three (0.3) (Δ. mit Punkt geschrieben, nought wird mündlich oft weggelassen)

Kommataehler punctuation mistake

Kommandeur(in) commander [kə'mɑ:ndə]

kommandieren (= befehlen) command

Kommando 1. (= Befehl) command [kə'mɑ:nd], order 2. das Kommando haben be* in command (über of) 3. (= Einheit mit Sonderauftrag) commando [kə'mɑ:ndaɪ]

kommen 1. allg.: come*; ich komme!

was! you'll never get anywhere if you go on like that! 19. hinter etwas kommen find* something out

Kommen: Sneakers sind wieder im Kommen sneakers are coming back into fashion

kommend 1. coming. (= zukünftig) auch future ['fju:tʃə]

2. Wendungen: kommende Woche next week; in den kommenden Jahren in the years to come; die kommende Generation the rising generation

Kommentar 1. (= Stellungnahme) comment ['kɒment] (zu on) 2. zu Fußballspiel im Fernsehen *u.w.*: commentary ['kɒmentəri]

3. in Zeitung: opinion column [ɒ'pɪnjən,kɒləm]

Kommissar(in) 1. Polizei: superintendent [ˌsʊpərɪn'tendnt], AE captain ['kæptən]

2. (= Bevollmächtigte, -er) commissioner

Kommode chest of drawers [ˌtʃest əv 'drɔ:z], AE auch bureau ['bjʊərəʊ]

Kommune (= Gemeindef) community

Kommunikationsmittel: ein modernes Kommunikationsmittel a modern means of communication

Kommunion Sakrament: (Holy) Communion

Kommunismus communism [ˌkɒmjʊnɪzəm]

Congress Organisation: The academic's burden

Geoffrey Williams

Every two years the EURALEX Congress comes round again, moving each time to a new venue so that we can explore Europe and its university cities. Why every two years, why keep moving? As the last organiser I shall try to answer these questions and give an overview of how the events are organised.

EURALEX congresses are an important event in the lexicographical calendar as they set off a rare meeting point for all the actors in the field, lexicographer to lexicographer, business to business, academic to academic and all to all. During four days all these essential actors in the lexicographical community from throughout the world are able to meet friends, make friends and exchange views. A lot of planning and unpaid work go into making a successful congress. The aim of this paper is to show the type of constraints under which the organiser is working, and consequently how the community can help make the task easier. Obviously I can only speak for Lorient 2004, but most of the organisational problems are the same everywhere. Before going into detail I must elucidate slightly the title.

The academic's burden? It is true that there are plenty of us in the membership, but the wealth of EURALEX is in the variety of its members, not just in terms of geographical background and the diversity of expertise, but also in terms of professional background with practising lexicographers, in-house or independent, publishing houses, software developers and just plain lovers of dictionaries, as well as academics. The task of congress organisation, though, tends to fall to the academics. This is probably for a number of reasons. Academics are in general relatively stable, physically if not mentally, staying for the time necessary to organise events. They generally have a relatively flexible work schedule into which congress organisation can be fitted as an extra, but handleable task. Congresses are also an occupational hazard for an academic, if you do not publish research you do not exist. One way of publishing is through attending conferences, and if you attend them, why not organise them? The last point is the nature of their employer, as academics also work for institutions that welcome congresses as a means of showing off their facilities and affirming their reputation as a seat of learning.

Whilst it may be natural that the task falls to an academic, it remains useful that others recognise the immense amount of work that goes into an event.

A moveable feast

Like the progression of a medieval monarch, EURALEX wends its way through Europe, although it leaves behind an exhausted, but happy, organiser, rather than devastation and famine. I cannot believe that anyone ever regrets taking on the task, the finale is such fun, but during the two years of preparation doubts can occasionally be voiced as to the wisdom of the undertaking.

I started off with two questions; why every two years, and why keep moving?

The first will be the subject of the rest of this paper. The aim is to give a brief overview of the tasks involved and the time schedule from acceptance to the opening ceremony. The second can be answered more quickly.

We are a European association, so it seems normal that each country gets the chance to show off its own lexicographical practice. If a meeting is always held in the same place it tends to become the property of that place. This is fine for some areas of research where a university seeks to become a centre of excellence in that field, but EURALEX is not a university association. It has a more virtual existence, simply using facilities put at its disposal by member institutions, or the institutions of its members. By moving, the congress remains on neutral ground, allowing both business and academic partners to exchange freely.

This mobile-feast syndrome means that since the founding of the association in 1983 and its first meeting in Exeter, EURALEX has been steadily moving around Europe, each time hosted by a different institution, some big, some small, but all committed to making everyone welcome. Of course, some places are easier to get to than others, but everywhere is difficult for someone. Every country has its own dictionary publishers and institutes, which means that in moving each country can present its own lexical practice. This is vitally important in terms of membership and European cohesion. Lorient offered French dictionary publishing houses the opportunity to receive the rest of the lexical world on its own ground. It meant that many French academics and researchers



Geoffrey Williams is a senior lecturer in the Applied Languages Department of the University of South Brittany, Lorient (UBS). He has been teaching in France for 26 years, most of this in the University of Nantes. After completing a PhD in Corpus Linguistics, he joined UBS in 2000. Since then he has been working to promote the use of corpora in language research in France and is the organiser of the Journées International de la Linguistique de Corpus, held every two years in Lorient. His research interests are special language corpus linguistics and lexicography, and he is particularly interested in using collocations in the development of corpus-based advance learner's dictionaries for use in scientific contexts. He was the organiser of the Eleventh EURALEX International Congress, held in Lorient in July 2004, and is assistant secretary and treasurer of EURALEX.
geoffrey.williams@univ-ubs.fr

**The Dictionary Society
of North America
XV Biennial Meeting
Boston, MA
June 8-11, 2005**

The fifteenth Biennial Meeting of the Dictionary Society of North America will take place at Boston University, in the New England area where American lexicography had its beginnings. In addition to the meeting, members will be able to attend events at the offices of the local dictionary publishers Houghton Mifflin and Merriam-Webster. Papers will be given on a wide range of topics in lexicography and lexicology. Presentations are twenty minutes, followed by five minutes for discussion, and there will also be a forum on dictionary usage open to the public. The conference will include tours of Merriam-Webster and the American Antiquarian Society, and of Harvard University's Houghton Library and the Boston Public Library's Rare Book Room, as well as a reception hosted by Houghton Mifflin. New England has played an important role in North American lexicography and continues to do so today. It is fitting that lexicographers from around the world will convene here and learn about the modern publishers of American dictionaries and have an opportunity to see documentation of the history of lexicography including American lexicography leading up to the present.

David Jost

DSNA President
<http://polyglot.lss.wisc.edu/dsna/DSNABostonMtg.html>

who would not necessarily have attended previous congresses could come along. Many of these will join the association, and will also come to Turin, where Italian lexicographers will bring a new Italian outlook to the next event, and from there on, who knows? Europe has such a wealth of lexicographical practice to share; our wealth of languages underlines our rich cultural diversity. EURALEX is far from an English-speaking club; whatever language someone chooses to present in, there will always be an audience eager to listen.

Organisation

Now we can get down to business. What happens in the two years between congresses?

Acceptance

The first stage is that an offer to host the following event must be accepted by the general assembly, which always meets during the congress. Prior to this, offers will have been discussed by the board so that some arbitration can be done if there is a plethora of potential venues; this also helps planning ahead even if the plans rarely work out.

Once the offer has been accepted serious planning can start. This means setting up a local committee and handing out tasks. This more than takes up a first year as the website and databases are built, the call written, contacts made with local authorities, including the tourist board, contacts established with the press, et cetera. Each organiser knows that the previous organisers can be relied on for assistance, but local needs generate new problems, which may require local solutions. We took on two students to help us through this first year, which culminates, with the summer board meeting.

Launching the machine

By the time of the board meeting, which always takes place at the new host institution in the summer preceding the congress, all the infrastructure must be in place, as by June the first call for papers must go out. The call should be in at least two languages, we used three, with English and Breton in addition to French. The inclusion of Breton was important to us, as being able to include one of Europe's lesser-spoken languages underlines our concerns for all the languages of the EU.

As the call goes out in the *International Journal of Lexicography*, advance planning is essential as paper-publishing deadlines are obviously different from those of electronic lists. The journal reaches a different audience than the electronic

lists, and EURALEX members must be informed first. In addition to the journal, we try to cover all the major electronic lists, the aim is to widen the circle to all those interested in lexicography.

The call will go out at regular intervals up to the final deadline for the receipt of papers. There is obviously a limited degree of flexibility, but the deadlines for the review process are very short. Whereas in some congresses review is done on short abstracts, EURALEX requires full abstracts so that a rigorous selection process can be carried out and the whole proceedings be in print before the event. This puts a lot of pressure on the organiser.

At this point I must reiterate the fact that the organiser is still doing their fulltime job. It is very unlikely that they will get a reduced teaching or research load, in fact the opposite generally occurs as those dynamic enough to organise a major event must be able to take on something else, *n'est ce pas?*

The review process

The reputation of a congress is paramount to success; if the selection is seen as anything but rigorous then the congress will no longer fulfil its purpose of disseminating knowledge. It is all too easy for a congress attached to an association to be seen as a sort of club where it is sufficient to be a member to speak. Such a situation is clearly disastrous, as a small group of people would only ever listen to themselves and over time the circle would inevitably diminish and the debates become sterile. This is far from the case with EURALEX, as we have a strict process of double-blind review, which ensures rigorous selection. For those who do not know the process I shall explain.

As can be seen from the call, every EURALEX congress has a programme committee, these are the people who will make the final selection, but on the basis of reports supplied by the review panel. The review panel consists of experts in different fields of lexicography who agree to read and comment on a maximum of 5 proposals. The organiser has of course the lists from previous congresses, but will also add in more reviewers from their own circle of professional acquaintances, as a new venue will require a greater weight in some language combinations. Management of this panel is complex, as unlike many congresses that only have English as their working language, EURALEX is very multilingual. This means taking into account both speciality and language combinations; a far from

easy task as it is all too easy to have a wealth of reviewers for one speciality or combination and a dearth elsewhere. This was very much the case for EURALEX 2004 as we had a much larger number of proposals in French than at previous meetings; the situation will probably be the same for TurinLEX in 2006.

Bringing in new members on a review panel also leads to complications. EURALEX standards are rightfully high. This means that time must be taken to explain the review procedure to new reviewers. All the reviewers are heavily committed professionals who are taking on the task for free, which means that the time available is also a factor. For EURALEX we use a standard review form with fixed criteria, and also have a space for comments to help decision-making and others that are passed on to the authors who will expect feedback.

When a proposal is received it is given a number and filed ready for review. Proposals must be classified by theme and language. In some cases authors will have ticked a number of themes, in which case the organiser will have to exercise judgement as no paper can be reviewed by 6 reviewers, which would be the case if 3 categories had been ticked. For the smooth running of the review process, discipline and realism on behalf of the authors is essential, but cannot always be counted on.

Deadlines must be respected and late papers cannot normally be considered. The organiser can have well over 200 proposals to manage and cannot send them out until after the deadline has passed, as otherwise managing a review panel is impossible. Proposals that do not satisfy the presentation format cannot be reviewed. Time is another factor. However much you might have to say, the reviewer only has a limited time to read it. A proposal that is too short cannot be properly judged; one that is too long may indicate someone who will not respect speaking times either. In all cases, electronic submission must be preferred so as to reduce handling costs and speed up operations.

Once the proposals have been safely gathered in and rendered anonymous, the review procedure can begin. In normal circumstances, a reviewer will receive a maximum of 5 proposals, although given the number of proposals received some reviewers will accept more. Review is anonymous and each paper will be read by two persons independently. This is double blind review.

The review procedure is not without its problems. There are anguishing cases

where the reviewer fails to deliver; in this case another reviewer has to be found at short notice. The same may apply if a paper is deemed highly acceptable by one but is rubbished by another. We aim to be rigorous but fair, and double blind review ensures this.

Upon receipt of the review forms, numerical values are fed into the database so as to give an initial sorting by clear acceptance and clear refusal. The former do not pose a problem, but the latter do receive individual attention, as all refusals must be justified. At this stage the proposals are still anonymous.

Programme committee meeting

The review process up to this point will have been carried out by the organiser with in-house assistants, all working under great pressure. Once the reviews have been received and classified; the selection can then be carried out by the programme committee.

The programme committee consists of the organiser assisted by a team of four other persons, usually members or ex-members of the board and the previous organiser. This means a team of people who already have experience in organising EURALEX congresses. The committee meets on a Saturday, and has one day in which to go through all the proposals from clear acceptance to absolute rejection and to choose those that will go onto the programme. In other words, they must reject a minimum of 50% of the papers, many of which would be quite worthy of acceptance. This is done anonymously, only after do we reveal names and see who we have offended. Once the meeting is over it is the job of the organiser to send out letters of acceptance or refusal, and wait for the flack to arrive.

The higher the number of proposals, the higher the level of rejection. Having too many parallel sessions is to be avoided and nobody would be satisfied by having an over-rich programme. The event cannot be lengthened either as it would greatly increase costs. This means that the review process is a gruelling one for the committee, and especially the organiser who can expect angry recriminations. However, the high level of selection ensures the reputation of the congress, as participants know they can expect a very high standard of paper. This is important for any congress, especially when the participants are paying a fee and travelling long distances to attend.

There remains the problem of doctoral students. Papers from doctoral students are not given any preferential treatment,

The Fourth ASIALEX International Congress Singapore 1-3 June 2005

It is our great pleasure to host the Fourth ASIALEX International Congress at the National University of Singapore (NUS).

ASIALEX 2005 is one of the year-long academic events celebrating the one hundredth anniversary of NUS and is organized by the Department of English Language and Literature, the Faculty of Arts and Social Sciences and the Asia Research Institute.

With the theme of **Words in Asian Cultural Contexts**

this three-day conference of the Asian Association for Lexicography aims to examine the functions and representations of words, with Asia as its focus, and has strands with related perspectives in sociolinguistics and language pedagogy, information and communications technology, and literary, cultural and postcolonial studies.

The major objective of ASIALEX is to foster scholarly and professional activities in the field of lexicography in Asia.

ASIALEX 2005 aims to bring together scholars of language, linguistics and literature in an interdisciplinary forum.

The papers capture a focus on how words assume distinctive shapes and meanings in different cultural contexts. Particularly in Asia, where different cultures and ethnicities commonly converge and give rise to cultural and textual hybridity, the wider study

of words and language can enhance cross-cultural communication and lead to greater understanding of its peoples.

The plenary speakers at the conference are Charles Fillmore (University of California), Reinhard Hartmann (University of Exeter) and Zhang Yihua (Guangdong University of Foreign Studies).

Featured speakers include Susan Butler (Macquarie Dictionary), Gwyneth Fox (Macmillan Publishing), Gregory James (Hong Kong University of Science and Technology), Ismail Talib (NUS), Yukio Tono (Meikai University) and Lily Wong Fillmore (University of California). There will also be a pre-conference tutorial, *Using Corpora*, by Adam Kilgariff (Lexicography MasterClass).

We are pleased to welcome also all participants from the related associations: Euralex, Afrilex and Australlex. To our visitors, we hope you will enjoy the hospitality of the local residents as well as the sights and sounds of Singapore.

Anne Pakir

ASIALEX

President (2003-2005)

<http://asialex.nus.edu.sg>

some high quality papers will be accepted, others rejected on the same basis as any other proposal. However, I firmly believe in the necessity of opening the congress to a maximum of young researchers who would not be able to attend if not given the possibility of presenting a paper. Because of this we used the database to find papers that had been rejected and offered the possibility of presenting at a pre-congress work-in-progress session, following a new selection process. These short papers would be published in a separate section of the proceedings. I am pleased to say that this formula worked, and was well attended.

Invited Speakers

Another task of the programme committee is to invite the plenary speakers. The organiser will draw up a list based on his or her own wishes, but balancing the requirements of international speakers and speakers from the host country. The list is discussed by the programme committee and advice may be sought from the board. The next problem is one of availability, which means beginning negotiations well ahead. We were lucky in Lorient in that our first choices were both available and willing.

The last laps before the event

Publishing the proceedings

Review process over, final papers safely (including those of the invited speakers) gathered in, so now we can go to press. Not so quick.

In addition to guidelines, we provide an electronic style sheet, but not everyone knows how to use one. Failure to follow the guidelines means an enormous job of standardising the texts, which can come in a variety of formats. Some insist on sending pdf files, and have to be informed politely that we may not be equipped to transform these. Some may have adapted the style sheet to their own format, forgetting that we might in turn need to adapt the style sheet to suit our printer. Some still produce overlong texts, and have to be reminded that size adds to cost. Many will have complex diagrams and space-consuming image files which do not always transfer easily between systems. Our two students worked solidly on the files, updating the style sheets until the printer was satisfied and a print-ready pdf file could be prepared. This quite literally means sleepless nights, as the printer cannot and will not wait.

Drawing up the programme

Whilst all this is going on, the pre-

programme is being drawn up and circulated electronically. This too is no easy task, as managing parallel sessions must be done in such a way that people can easily change rooms whilst ensuring that thematic areas do not clash. Facilities and potential numbers also have to be taken into account. Not all rooms are equally well equipped, some papers will attract more audiences than others and therefore need a bigger room. Then come the cancellations, inevitable with over a hundred speakers. Added to this will be those who cannot arrive on a given day and need to change time slot, or who may have arrived but wish to listen to someone who is speaking when they are. Sometimes individual needs can be accommodated, but not always, as this is a very delicate balancing act.

The point here is not to shoot the piano player, your unpaid and overstressed organiser is juggling with a multitude of problems, and yours is but one of them.

One of the problems with which your happy organiser is trying to cope is the chairing of the individual sessions. Choosing a chair means finding someone who is knowledgeable on the subject area, who can listen and not intervene unless needed, and who can impose discipline when either the speaker or the audience get carried away. Being invited to chair a session is both an honour and a burden. Accepting the honour of a chair means reading the texts beforehand and preparing questions. No paper must go without at least one question, and often this first question will set off a discussion that the chair must then bring to a close when the time runs out. Timing is essential if people are to be able to change rooms without disturbing a paper that has already started. The chair is also stuck for at least one session unable to go and listen to other papers that might be of interest. Of course they get to hear the papers in their session, but they are never really relaxed as they are also note taking and keeping an eye on the clock. Chairs are precious volunteers.

The Congress

The last minute preparations are traumatic. Not only are you wondering what you have forgotten, but there are also practical aspects like preparing the congress bags. The proceedings will have to go in, so will the tourist publicity and other goodies. You know there will be a rush on the day so the bags must be distributed in alphabetical order. All this takes time. The congress starts on Tuesday, but you can expect visitors on Monday who would like to have their bag immediately and cannot

understand why it is not possible and why you may not have time to chat. You keep a smile and carry on rushing.

The big day arrives. What can go wrong? Plenty, and mostly the unexpected.

In Lorient, we did not plan on the storms, nor the power cut that followed, nor for that matter getting locked out on Saturday morning because the security man had overslept. I am used to forcing the doors of the university, my setting off the alarms does add to the early morning atmosphere, but it does not do good for my nerves. Throughout, the team kept its good humour and the participants put up with the minor, and some major, inconveniences.

Day to day management requires a good team; keen student helpers and wonderful staff do help greatly. In Lorient we had both. There are a myriad of problems that the staff will have to deal with, and provided they can do it with a smile all goes well. People may fail to have booked a hotel room, they may have a car accident and need an interpreter, health problems, again requiring an interpreter, clashes may occur which require oil on troubled waters. All of these can happen and must be dealt with. None of these can be really planned for, but if the team is there an answer will be found.

Neither the organiser nor the team can be said to “enjoy” the congress in the same way as the other participants. They never stand still, they rarely finish on time, they never get to listen to a paper all the way through. They must be everywhere at the same time. Thank heavens for mobile phones. And yet, the congress is the climax of two years hard work, so with the adrenaline running high their enjoyment will be there. There is nothing quite as thrilling as the buzz of conversation. In July 2004, from Tuesday to Saturday, our faculty buzzed and everyone could feel the satisfaction. Conference organisers and their team are enablers, and enabling is exciting.

Come Saturday lunchtime the pressure should begin to drop. Saturday afternoon is for excursions, and if the planning has been done the coaches and guides will be there and a new holiday atmosphere will reign. The team can then relax... until Monday morning when the tidying, paying bills, sorting out cancellations et cetera begins.

Financial aspects

Now that we have looked at the organisational problems and their happy outcome, we need to look at the financial side of things. Congresses require time, and time may be money, but other factors

also have their costs.

Running a congress ain't cheap

In the introduction I said that it might seem natural that the task of congress organisation fall to an academic. This is of course not really so. There are many large conventions, much larger and more expensive, but less friendly, than EURALEX that are organised by professional bodies and set up by professional organisers. So why go to an amateur?

The answer is, of course, cost.

Many see EURALEX as an expensive congress. Compared to some it is, compared to others it is not. If the money is coming from your pocket attending a congress is a pricey thing, especially once you add in the cost of travel and accommodation. On the other hand, food is included and you walk away with a couple of volumes of proceedings that reflect what is best in lexicography at the time. The cost aspect is also a cultural one. In France we try to keep our meetings cheap as we get little or nothing from our universities to help with conference attendance, students may get a grant, but lecturing staff do not, so that an active researcher is going to be heavily out of pocket. The northern hemisphere tends to have higher overheads so relatively high fees are not surprising, it may be that laboratories also defray some of the costs involved.

Keeping costs very low works for small events, as grants will cover the basic costs, but as soon as an event goes beyond a critical size, costs inevitably rocket. Whilst not being of the size of vast impersonal conventions, EURALEX congresses do bring together an average 300 participants, which means that some sort of infrastructure is essential. This means that relatively high fees are requested. However, I hold that EURALEX congresses do offer tremendous value for money. Some professional help may be called in, but the vast majority of the workload is undertaken by people working for free. To explain the cost factors involved I shall outline the main sources of income and outgoings.

• Income

Fee structure

Fee structure must take into account the ability of people to pay, which is why we have a variety of fees and deadlines. In building a fee structure a number of factors have to be taken into account: job status, geographical origin, membership and date of payment.

Les Journées Allemandes des Dictionnaires Klingenberg am Main

‘Les Journées allemandes des dictionnaires’ is a new lexicography conference devoted to French dictionaries taking place in Klingenberg am Main, Germany. It is organized by the Institut für Angewandte Sprachwissenschaft of the Friedrich-Alexander University of Erlangen-Nürnberg and is coordinated by Michaela Heinz.

The first meeting was held on 25-27 June 2004 with the theme **Entre définition et citation – l'exemple** featuring the lexicographic example in contemporary French dictionaries as a multiform and plurifunctional lexicographic unit which can be addressed to all types of users for different purposes in monolingual and bilingual dictionary contexts.

The event gathered some sixty participants with papers from Josette Rey-Debove, Bénédicte Gaillard, Jean Pruvost, Louis Mercier, Pierre Rézeau, André Thibault, Michel Francard, Pierre Corbin, Alise Lehmann, Michèle Fourment, Martine Coutier, Christiane Tetet, Peter Blumenthal, Franz Josef Hausmann and Michaela Heinz. The proceedings will be published by Max Niemeyer Verlag.

The second conference will be held on 7-9 July 2006 with the theme **Le Dictionnaire Maître de Langue: lexicographie et didactique** and be dedicated to the memory of Josette Rey-Debove.

It will highlight the didactic aspects of French

▶ dictionaries – their presentation, analysis and critique, and cover monolingual, bilingual and semi-bilingual learner's dictionaries, for native French speakers and for learners of French as a foreign language, for children, elementary French, chrestolexicography, the didactic value of illustrations in dictionaries, dictionaries and grammar, and didactic aspects in big dictionaries.

<http://www.ias.uni-erlangen.de/klingenberg2006> (from August 2005)

● **Michaela Heinz** has a PhD in applied linguistics from the University of Erlangen-Nürnberg. She collaborated on the editing of the *Nouveau Petit Robert* from 1988 to 1993, worked on behalf of l'Agence Universitaire de la Francophonie in lexicographic projects in France and abroad, and currently teaches French linguistics and (meta)lexicography as "Privatdozent" at the universities of Erlangen and Passau.
heinz.michaela@t-online.de

The first, job status, is relatively straightforward as we differentiate between students and non-students. It is obvious that students have a lower income than the fully employed; they are also the lifeblood of the discipline as they represent new ideas and future full members of the association. There is thus a much lower fee in this case. However, increasingly, institutions will only pay a fee if a communication is accepted. This is particularly problematic as the doctoral student gets no preferential treatment in the review process and is up against established researchers in the selection process. This is another reason why student fees must be kept low so as to allow self-financing students to attend, and also why we set up a special pre-congress work-in-progress session at EURALEX 2004 so that students would have a greater chance of getting a paper accepted.

Geographical origin is slightly more complex. Originally this was put in to assist with people coming from behind the iron curtain, or to economies that were gradually adapting to the western European economy. In this case we would offer a low fee, and in many cases participants would be assisted with a grant from the Hornby Trust. This low fee category only concerned European countries, as we are a European association, and is also one that is being phased out as the candidate countries are now full members of the EU.

Membership is the third factor. There is no point in having an association if everyone, members and non-members pay the same fees. Lower fees are a privilege of membership and may even incite some to join. There is more to EURALEX than just the congresses and the more we are, the more we can do to promote all aspects of our discipline.

The final factor is one of time, and that is one that many have trouble understanding. Congress organisation costs, EURALEX does have seed money to help get congress organisation going, but many outlays come before the event, so money must come in from somewhere. In addition, the organiser has to have an idea as to numbers involved. Lunches and receptions are included in the fee, so the restaurant must have advance notice. Food wastage increases costs, and in a hungry world is also immoral. The only way of covering initial costs and calculating numbers is to give a variable fee structure based on date of payment.

Sponsors and grants

Fees may seem high, but they do not cover overheads. Other sources of financing are required. A considerable amount of time

can go into this. National, regional and local authorities may help, either in cash or in kind. Conference bags and contents have a cost, so when someone gives them, it helps greatly. The EU may be a source of funding, but we found that the weight of bureaucracy was enormous and then we got a refusal. So much wasted time for no result is depressing, but then a sponsor comes along.

Some give for general funds, others for more precise areas. We were extremely lucky to have the assistance of at least one major sponsor. The money was important, and their presence throughout the week also added to the event. The great generosity of the Hornby Trust enables many from central and eastern Europe to attend; they also covered the costs of the Hornby lecture, a fitting tribute to a man who did so much for lexicography. As we move around Europe, the local dictionary houses will come forward to help. Le Robert provided the CD-ROM dictionary that was in the Lorient congress bag, Larousse paid for the champagne aperitif at the gala dinner. These are all things that would otherwise have to come out of general funds.

Sponsorship also comes in through the exhibition, the fees are ridiculously low for the service provided. During a whole week, some 300 people representing all aspects of the lexicographical world will see a stand. This is the time to compare what other companies are doing, to show off products, and make contacts. Exhibitors may even find new staff.

This is what the exhibitor gets, but sponsorship is a two-way thing. EURALEX congresses are run by academics who are training the next generation of lexicographers, the EURALEX association is promoting good lexicographical practice and bringing awareness of dictionaries to a wider public. Sponsorship is thus putting back a little of what the publishing industry takes back. Coming forward spontaneously, as our sponsors did, saves time for the organisers and clearly shows a commitment to the event, and to the discipline itself.

● **Outgoings**

Staff

Very many hours of unpaid labour go into organising a congress, but inevitably some paid staff are necessary. Everything depends on the individual institution, some may charge for secretarial assistance, others not. If the assistance is not available, outside staff will have to be brought in.

In Lorient, we are a young dynamic

university which is keen to organise things and has a staff that is keen to participate. This is a privilege for an organiser.

Staff are needed both during and before an event. In our case we chose to employ students from the applied languages department who would help with organisational tasks as part of their mandatory work placement. This meant that I could rely on students who were not only motivated, but had good language and organisational skills.

The permanent staff took on the extra workload as part of their work, but put in much more time than they were ever paid for. In our faculty we have a research centre with secretarial staff who are there to help manage research projects and meetings. Throughout the week Valerie Sauvaire and Danielle Guyomard were always available, as they were during preparatory meetings, weekends and evenings included. To these I must add the many other members of university staff who were around to help, up front and behind the scenes. In our case a special mention must be made of co-organiser Sandra Vessier. She is not an academic, her job is international relations, but she was with us from the outset facilitating and organising. This is a privileged situation, and one which cannot be counted on everywhere, for us it meant major budget savings.

Premises

Again the university provided these for free, and even arranged for a complete refurbishing of the audio-visual material in the amphitheatres. Hiring premises can be a very expensive, but necessary option. We had originally budgeted for hiring a hall for the welcome session, it would have greatly added to costs.

Catering costs

EURALEX 2004 took place after the closure of the university restaurant for the summer. However, the staff came back and produced excellent food with a very friendly service. Catering is an important extra cost, but not one where you can cut corners as a congress marches on its stomach, important business takes place during the coffee breaks and at lunch.

Publication

Although we managed to keep most of our costs under control, publication costs are a very big outlay. Programmes and proceedings can only be printed at the very last minute, this means respecting very tight schedules and finding a printer who would respect them too.

Traditionally the EURALEX proceedings are always in the conference

bag. This means that from receipt of final copy to printing there is only a matter of weeks. Few university printing facilities can cope with the number of copies to be produced in such a short space of time and to a high standard. It is thus necessary to use an outside contractor, and as it will be a one-off contract you cannot expect favours. In our case the tough bargaining was carried out by our communications department, but the price was still high.

Congress organisation: the academic's privilege

By way of conclusion I shall revisit the title. Burden? In some ways, yes, as it generates a tremendous amount of work that cuts into a busy teaching schedule and virtually writes off any research activity. Privilege? Indeed. Research is nothing if not done within a community. EURALEX is more than just a special interest group. Getting involved leads to meetings with remarkable people, it brings challenges, and meeting challenges is what carries us forward. Organising a congress is indeed a privilege.

This paper only touches the tip of a lexicographical iceberg in terms of the time taken to organise an event. The time it took me to type up these notes is nothing compared to the time spent preparing the congress. The workload is tremendous, but the satisfaction great. I did not hear any papers, I always had something to do, but I could feel the exchange of debate, the constant buzz of debate. The staff at UBS enjoyed the week, the students enjoyed it. But we would not tackle anything so big for a while yet.

One of the great privileges of an organiser is to hand on the flame to someone else. I have handed onto Carla Marella who will be bringing us TurinLEX in 2006. She has now had a busy year, she has a busier one ahead. The result will be a very different event from EURALEX 2004, which was different from Copenhagen, from Stuttgart, from Liège, from all the other congresses. I shall have the advantage of being just an ordinary participant, but one who knows what it takes to make a congress work, and why it is that EURALEX is such a successful organisation.



The ancient university courtyard (seventeenth century) of Università degli Studi di Torino where the opening ceremony of TurinLEX 2006 will be held

EURALEX International Congresses

- I** Exeter, UK 1983
- II** Zurich, Switzerland 1986
- III** Budapest, Hungary 1988
- IV** Benalmádena (Malaga), Spain 1990
- V** Tampere, Finland 1992
- VI** Amsterdam, The Netherlands 1994
- VII** Gotebörg, Sweden 1996
- VIII** Liège, Belgium 1998
- IX** Stuttgart, Germany 2000
- X** Copenhagen, Denmark 2002
- XI** Lorient, France 2004
- XII** Turin, Italy 2006



The Twelfth EURALEX International Congress, Turin 2006

Carla Marello



Carla Marello is a Professor of Applied Linguistics at the Faculty of Foreign Languages of the University of Turin and Director of PhD studies in Linguistics, Applied Linguistics and Linguistic Engineering. She has published extensively on lexicography, lexicology, and theoretical and applied linguistics, is associate editor of the *International Journal of Lexicography*, and series editor at Zanichelli and Guerra publishing houses. She has been on the executive board of EURALEX since 1992 and is the organizer of its forthcoming 12th Congress due in Turin in 2006.

carla.marello@unito.it

Turin (Torino, Italy) hosts the XII EURALEX International Congress on 6-9 September 2006. Situated near the French border, Turin has always had a strong lexicographic tradition, reinforced during the nineteenth century in parallel to the unitary policy of the Savoy family.

TurinLEX will have among its organising partners the Florentine Accademia della Crusca, whose *Vocabolario degli Accademici della Crusca* appeared in 1612 and has turned Florence into the very birthplace of modern European monolingual lexicography, a model for dictionaries of French, Spanish, Portuguese, English and German. Now it will welcome in Turin lexicographers from all over the world.

The international character of Turin has both historical and contemporary reasons. As the capital of the multilingual Savoy Kingdom, which spread across the Alps and included Sardinia, it has had since the eighteenth century fine production of dialectal dictionaries, of dictionaries for French, Latin and Greek, and of onomasiological dictionaries. At the end of the nineteenth century, 17 publishing houses operated in Turin, and since most of them were active in the school book market, nearly all had at least one dictionary in their catalogue. The best known were Paravia, Loescher, Rosenberg & Sellier, Petrini and UTET (Unione Tipografica Editrice Torinese), all of whom still produce dictionaries today.

When in 1861 Luigi Pomba began the publication of the great dictionary later known as *Tommaseo-Bellini* (completed in 1879, eight volumes, 7300 pages), he was well aware of promoting “a national monument” of the history of the Italian language and was doing it in Turin, the capital of the newly-formed Kingdom of Italy. One hundred years later, the same publishing house, UTET, started the dictionary known as *Battaglia, Grande dizionario della lingua italiana* with the aim of pursuing and updating the Tommaseo-Bellini work.

Rosenberg & Sellier specialised in Latin dictionaries, but Paravia and Loescher also published good Latin and Italian dictionaries. The recently compiled *Ancient Greek and Italian Dictionary*, by Montanari, was published by Loescher. SEI (Società Editrice Internazionale),

Petrini and Paravia produced bilingual dictionaries for French, German, English and Spanish, which had their share of buyers, constantly eroded in the latter part of the twentieth century by Zanichelli and Garzanti dictionaries. Malcolm Skey's 1977 adaptation and bilingualization of *Oxford Advanced Learner's Dictionary* for SEI and the recent series of Tullio De Mauro's Italian monolingual dictionaries (published by Paravia and UTET) are among the most original and innovative lexicographic works to appear in Italy in modern times.

The humanistic faculties of Turin University (TU) were not without some part in these lexicographic activities, since most members of local in-house editorial teams graduated from TU. In the last three decades there has also been remarkable experimental work – carried out by TU professors together with software and publishing houses – on classroom use of dictionaries and the creation of corpora.

TU is also the seat of Atlante Linguistico Italiano (ALI) and attracts scholars and experts in sociolinguistic surveys. The Turin school of engineering and applied-sciences (Politecnico) is well known and has contributed to the industrial past of the town and to its present transformation. Many scientists teaching at the Politecnico have contributed to encyclopaedic or specialist dictionaries in Turin and elsewhere.

Piedmont, the region at the feet of the Alps in the middle of which Turin is located, will host the 2006 Winter Olympic Games, but in September next year EURALEX participants will enjoy the warm end of summer and the sight of first reddish leaves in the numerous vineyards. The organizers will do their utmost so that the guests can appreciate a part of Italy which has always been a hinge between North and South and which will be an important stopping place on the European East-West corridor.

The congress programme will include plenary lectures, parallel sessions on the usual topics of previous congresses, software demonstrations, a book and software exhibition, and social events. We hope to welcome many of you in a town which offers a lively human and scientific milieu.

Further details

euralex2006@unito.it
www.euralex.org