A NEW MODEL FOR A FOREIGN LANGUAGE LEARNER'S DICTIONARY

by

Philippe R. M. Humblé

A thesis submitted in partial fulfilment of the requirements for the degree of

Doutor em Letras

Universidade Federal de Santa Catarina

1997

Approved by Prof. Dr. José Luiz Meurer

Chairman of the Supervisory Committee

Pós-Graduação de Literatura e Língua Inglesa

November 1997

Esta tese foi julgada adequada e aprovada em sua forma final pelo
Programa de Pós-Graduação em Inglês para a obtenção do grau de

DOUTOR EM LETRAS

Opção Língua Inglesa e Lingüística Aplicada ao Inglês

Drª. Barbara Oughton Baptista
Coordenadora

Dr. José Luiz Meurer
Orientador

Banca Examinadora:

Dr. José Luiz Meurer
Orientador

Drª. Rosa Weingold Konder

Dr. Jean Binon
(Univ. Católica de Leuven, Bélgica)

Drª. Silvana Serrani
(UNICAMP)

Dr. Richard Malcolm Coulthard
(Univ. de Birmingham, Inglaterra)

Florianópolis, 19 de dezembro de 1997.

Every author may aspire to praise, the
lexicographer can only hope to escape reproach,
and even this negative recompense has been
granted to very few.

Johnson in the *Preface* to the *English Dictionary.*

Voor Fernand Humblé, *In memoriam*

UNIVERSIDADE FEDERAL DE SANTA CATARINA

A NEW MODEL FOR A FOREIGN LANGUAGE LEARNER'S DICTIONARY

by Philippe R. M. Humblé

Chairperson of the Supervisory Committee: Professor Dr. José Luiz Meurer

Pós-graduação em Língua e Literatura Inglesa

Abstract

This thesis deals with the problem of foreign language lexicography proposing a new dictionary model. It is concerned with the lack of adequacy between, on the one hand, current reference works and, on the other, the learners' needs. These needs have been insufficiently investigated and this thesis suggests that the lack of substantial results in the area is due to a flawed research methodology. It is suggested that a qualitative type of research, instead of a quantitative one, should yield better results. A sample of this kind of qualitative analysis, carried out on dictionary examples, bears out that dictionaries would gain with a separation of the encoding and decoding parts. This had already been suggested by a few researchers in the past at a theoretical level. A series of consequences of this separation are then considered and, as a result, recommendations are formulated for the compiling of dictionaries in the future. These recommendations mainly tend to improve the encoding part of the dictionary. They were partly induced by new technologies, without being dictated by them. The relevance of this work lies mainly in the possible practical outcome of what is here proposed: a new kind of foreign language dictionary.

Resumo

Esta tese é um trabalho de pesquisa lexicográfica voltado para o ensino de uma língua estrangeira. Propõe-se aqui um novo modelo de dicionário, melhor adaptado às necessidades dos usuários modernos. Segundo o autor, estas necessidades são hoje muito mais de produção e não mais unicamente de compreensão. O fato dessas novas necessidades serem insuficientemente pesquisadas deve-se a uma metodologia de pesquisa inadequada. Sugere-se que um tipo de pesquisa qualitativa, no lugar de uma quantitativa, levaria a melhores resultados. Um exemplo deste tipo de pesquisa qualitativa e aplicada a exemplos, mostra que os dicionários ganhariam em eficiência se as funções lexicográficas de *produção* e *compreensão* fossem consideradas separadamente. Alguns pesquisadores já defenderam esta posição no passado a um nível teórico, mas suas conseqüências práticas nunca foram consideradas. É precisamente o que se tenta fazer aqui. Algumas recomendações são formuladas para a compilação de dicionários de língua estrangeira no futuro, mais do que nada para produção. A informática, em parte responsável para a maior necessidade de dicionários de produção, é considerada uma ferramenta essencial. A relevância deste trabalho está também nas eventuais conseqüências práticas do que se propõe aqui.

Table of Contents

List of figures.

Acknowledgements

When time comes to write the acknowledgements, one is surprised to see how many people have in some way collaborated to bring an undertaking like this to a happy ending.

I would like to thank Prof. Malcolm Coulthard of the University of Birmingham, UK, who assisted me in the first stages of this thesis and provided me with a few of the fundamental insights which lay at the basis of this research. His suggestions showed the direction to take and his criticisms were always extremely helpful.

I would like to thank Prof. José Luiz Meurer for his assistance in the later stages of my research. His advice as an extremely good reader and our conversations on the structure of this research were decisive to help me bring it to conclusion.

I particularly appreciated the help and criticisms of my ex-colleague in Belgium, now co-supervisor, Prof. Jean Binon in whom I found a kindred spirit in terms of lexicography. I have no doubts that our collaboration only just started.

I would like to specially thank Liz Potter from COBUILD for reading and commenting on the text.

A number of people read all or part of this thesis and assisted me with their criticisms. I would like to thank Walter Carlos Costa, Gloria Gil, Silvana Serrani, Carmen Rosa Caldas-Coulthard, Jean Esther Langdon, Silvia Bernardini, Ana Cristina Ostermann, and Christine Everley.

I am particularly indebted to the people at COBUILD who assisted me so well during the fifteen months in which I had the pleasure to be their guest. Special thanks go to Gwyneth Fox who made my stay possible, to Rosamund Moon with whom I had many a fruitful discussion, and the whole COBUILD team now unfortunately scattered around the four corners of the world of lexicography: Ramesh Krishnamurthy, Ceri Hewit, Roz Combley, Laura Wedgeworth, David Morrow, Sue Crawley, Michelle Devereux, Jeremy Clear, Tim Lane, Gill Francis John Todd, Christina Rammell, Jenny

Watson, Mick Murphy, and Sean Lynch.

Special thanks go to my Japanese teachers, here and in Britain: Tsukamoto Chitose, Hukamachi Hideyo and Kimura Mutsuko.

I would like to thank my colleagues at the Universidade Federal de Santa Catarina for having taken over my classes while I was working on this research: Alai Garcia Diniz, Liliana Reales de Ruas, Luizete Guimarães Barros, Rafael Camorlinga, Vera Aquino, and Walter Carlos Costa.

I am much obliged to Paul Meara whose hospitality and good advice I was able to enjoy while his guest in Swansea.

Finally, I would like to thank the Universidade Federal de Santa Catarina, for giving me a four year study leave without which this research would have been impossible to carry out, and the CNPq for awarding me a scholarship which allowed me to spend a year at Birmingham University in excellent conditions.

Of all the *Bolsa de Iniciação Científica do CNPq* and *Monitoria* students who worked with me at the *Dicionário de Uso Português-Espanhol* and whom I wish to thank here, my special gratitude goes to Daniela de Carvalho whose enthusiasm for a long time supplied her grant.

Finally, I can recommend my mother and Tante Mieke to anyone having a bibliography to revise, as for a great number of other things.

Glossary

Bilingual dictionary. A bilingual dictionary translates lexical items from one language into another. Depending on who is the user, a bilingual dictionary translates from an L1 into an L2 or from an L2 into an L1. A normal bilingual dictionary has two possible audiences: the L1 of one audience is the L2 of the other audience.

Bilingualised dictionary. A bilingualised dictionary is a kind of pedagogical dictionary. It necessarily includes information in two languages without being a bilingual dictionary in the traditional sense of the word. Often bilingualised dictionaries are a partial or total translation of a monolingual learner's dictionary. In some cases, the definition of a word in a foreign language is given in the foreign language and a translation in the language of the learner added. In other cases, The definition is translated into the language of the learner. There are hybrid forms such as the *COBUILD English-Portuguese Bridge Bilingual Dictionary*, where the definition is partly in Portuguese and partly in English. In contrast with the normal bilingual dictionary, the bilingualised dictionary has a clearly defined audience: those who speak the language into which the headwords or the definitions were translated.

COBUILD. Abbreviation of Collins Birmingham University International Language Database.

Collocate. A *collocate* is a word which occurs very frequently with another word so that the chance of both occurring together is very high.

Context. Context is the surroundings, in terms of meaning and not only of words, of a lexical item. The *context* includes the *co-text* insofar as it is specifically relevant to understand the meaning of a particular lexical item. The context does not have to be linguistic and can include the kind of medium in which a text was published or uttered.

Corpus. A *corpus* is 'a collection of naturally-occurring language text' (Sinclair on the Corpus Linguistics Homepage http://www-clg.bham.ac.uk/glossary.html) In this thesis, I will refer mainly to the *Bank of English*, corpus of COBUILD.

Co-text is the immediate linguistic surroundings of a lexical item. Often this co-text will indicate a nine-word span, four words on each side of the word under consideration.

Decode. To *decode* is to translate from an L2, target language or foreign language, into an L1, source language, usually mother tongue. *Decoding* is, in the context of this thesis, the process of understanding a, usually, written text in a language which is not the *decoder*'s L1. It is a commonly used term in lexicography and, as far as I am concerned, does not imply any specific linguistic theory on language as a code. It does not mean that when somebody reads or hears another language, they are *decoding*. One is only *decoding* when one does not spontaneously understand the word in the L2 and has to look it up or go in search of some kind of linguistic information in order to understand the message.

Encode. To *encode* is to translate from an L1 into an L2. *Encoding* is the process of producing a text, written or spoken, in a language which is not the *encoder*'s L1. This does not mean that when somebody speaks or writes another language, they are *encoding*. One is only *encoding* when one does not spontaneously know the word in the L2 and has to look it up or go in search of some kind of linguistic information to be able to do so.

Fixed expression. A series of words people feel belong together and express one concept. I avoid using *idiom* in this sense, although this term is widely used to indicate a fixed expression. (See also *idiomatic expression.*)

Foreign language lexicography. By *foreign language lexicography* I understand all types of language reference works which take non-native speakers as their audience. These reference works can be written in the target language (L2), or include both the source (L1) and the target (L2) language.

Idiomatic expression. An idiomatic expression is a group of words whose meaning cannot be predicted from the meanings of the constituent words *(Collins English Dictionary).* Idiomatic expressions are sometimes called *fixed expressions* or *idioms.* There exists no consensus on this terminology and I give preference to the term *fixed expression,* not to be confused with *idioms.*

Idioms. Idioms are 'semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments' (Sinclair, 1991:110) 'The individual words which constitute idioms are not reliably meaningful in themselves, because the whole idiom is required to produce the meaning. Idioms overlap with collocations, because they both involve the selection of two or more words. At present, the line between them is not clear. In principle, we call co-occurrences idioms if we interpret the co-occurrence as giving a single unit of meaning. If we interpret the occurrence as the selection of two related words, each of which keeps some meaning of its own, we call it a collocation.' (Sinclair, 1991:172)

L1. The *L1* is the source language, normally the learner's mother tongue.

L2. The *L2* is the target language, normally a foreign language which a learner is in the process of mastering.

Learner. A *learner* is someone who is in the process of mastering an L2. Any person using a dictionary in order to obtain information on a language other than his/her mother-tongue, is a 'learner'. Even if people who consult dictionaries are not always doing this with the conscious purpose of learning, the fact is that they do learn and, consequently, at the moment in which they consult a dictionary, they are learners.

Learner's dictionary. A learner's dictionary defines the words of a particular language in that same language for people for whom this language is not the mother tongue. Learner's dictionaries only exist for the languages most studied by foreigners: English, French, Spanish and German. A learner's dictionary aims at giving a more extensive treatment to the most frequent words of a language, although these 'most frequent words' may include up to 90.000 lexical items *(Longman Dictionary of Contemporary English).* Usually the vocabulary used in this kind of dictionary is

'controlled'. Only the, say, 2.000 most frequent words are used in the definitions. Learner's dictionaries have a negatively defined audience. They are not intended to be used by native speakers.

Lexical item. A lexical item consists of at least one word but can consist of more than one. I use the term when what I am referring to applies equally to items including one or more words, or forms of words. *Swimming pool, found* and *at first glance* are all *lexical items.*

Lexis. The *lexis* of a language is the set of all its word-forms. (Sinclair, ibid.)

Monolingual dictionary. In this research I consider two kinds of monolingual dictionaries: *native speaker's dictionaries* and *learner's dictionaries.*

Multi-word item. A multi-word item is a number of words which is considered to refer to one single reality (e.g. *operating table*).

Native speaker's dictionary. A native speaker's dictionary defines the words of a particular language primarily to those for whom this language is the mother tongue. All complex societies, in the anthropological sense of the word, seem to have their dictionary.

Open-choice principle. This is a concept I borrowed from Sinclair (1991:109). The normal way of seeing language according to a 'slot and filler' principle. According to this principle almost any word can occur after any word, respecting syntactic constraints.

Pedagogical dictionary. Pedagogical dictionaries are conceived primarily as teaching materials and secondarily as reference works. They are different from other kinds of dictionaries in that they are as much 'tool' as 'study material', and often concentrate on one particular area of the lexis of a language (e.g., *français des affaires*). As a consequence, pedagogical dictionaries are often published together with exercise material.

Source language. The L1, normally the learner's mother tongue.

Target language. The L2, normally a foreign language which a learner is in the process of mastering.

User. A *user* is a learner using a dictionary.

Word. A *word* is anything between two white spaces.

CHAPTER 1 INTRODUCTION

1.1. STATEMENT OF THE PROBLEM

1.1. The purpose of foreign language lexicography is to help learners with *decoding* from and *encoding* into a foreign language. Until recently and for a number of historical reasons, dictionaries privileged *decoding*. People basically needed to *understand* messages and texts in a language other than their own and the bilingual dictionary would be their main tool. However, over the last half century more emphasis has come to be laid on *encoding*, particularly in the case of international languages. People increasingly have the possibility, and therefore the need, to communicate with each other in real time, and not only by reading each other's books. This change in habits calls for new tools, certainly in the field of foreign language teaching and learning.

1.2. Current dictionaries are still better suited to *decode* than to *encode*. Certainly, bilingual dictionaries have improved to remedy the lack of balance between the two activities, and new initiatives like bilingualised and learner's dictionaries have been taken in the course of the last fifty years. Even so, no steps were taken to change the face of dictionaries radically. There is still room for improvement and present-day language learners would like to have dictionaries which better assist their *encoding* needs.

1.3. The fact that the ideal language tool is, in my view, still not on the market is due to imperfect knowledge of what the needs of non-native speakers are and how they would naturally proceed to meet them. In spite of some very useful research, the process has not been adequately described. The reason for this is at least partly a flawed conception

of how research in the area should be conducted. Until now, research has been done basically by means of tests and questionnaires. These are quantifiable, that is, they yield results which can be put into figures and statistics. There is no question that this kind of research has value, but there is now a need for a more qualitative approach, consisting of introspection on the part of the researcher, and a careful analysis of actual dictionaries in order to discover what their flaws are. It is this qualitative kind of research which I aim to give a sample of.

1.2. Objectives and Hypothesis

The objective of this thesis is to formulate suggestions to improve dictionaries for learners of foreign languages, and this by means of a qualitative kind of research which concentrates on an analysis of existing dictionaries on the one hand, and on an analysis of the learner on the other.

I have been carrying out this research for the last four years on a daily basis, and for the last twenty five in an informal way. The final outcome of my research is a number of suggestions which, if applied, could lead to an improvement of foreign language dictionaries. Although I have constantly in mind the possibilities of electronic methods of storage and retrieval of data, and am convinced that the computer is the future of the dictionary, the suggestions I will formulate will have some possible application in paper dictionaries.

This research aims at solving at a theoretical level a concrete problem which a number of people are confronted with on a daily basis: the inadequacy of dictionaries to

present-day needs. An analysis of these needs and an analysis of dictionaries are therefore the two poles between which this research moves.

It is my hypothesis that foreign language lexicography is in crisis and that the reason for this is that learners have an increased need for a dictionary which can help them with *encoding*. In order to design such a dictionary, one has to find out what the needs of *encoding* learners are.

Needs are not always easy to identify as is shown by the literature on the subject. After an analysis of this literature, I am convinced that research should turn from a quantitative to a qualitative methodology: analyse specific dictionaries and specific look-ups.

The ultimate value of this research is practical. The starting point was my discontent with reference works as they are currently available on the market and I would like to contribute to their improvement.

Synopsis

In the first chapter, I state that foreign language lexicography aims at helping learners with *decoding* and *encoding* and that over the last fifty years emphasis has shifted from almost exclusively *decoding* to *encoding*. However, current reference tools have still not adapted to this new trend. In order to do this, it is important to know what are precisely the needs of dictionary users. I take the view that there is a lack of understanding of the issue and that this is due to inadequate research techniques.

In the second chapter, I situate foreign language dictionaries in the broader field of

historical lexicography. Although not very suited for *encoding*, I suggest that bilingual dictionaries still fulfil an obvious need. This is, however, not the case for monolingual dictionaries, which serve ideological needs alongside linguistic ones. Through learner's dictionaries, these ideological elements have penetrated the field of foreign language lexicography.

In the third chapter of this thesis, I review the literature on foreign language lexicography. To discover the dictionary needs of the foreign language learner, researchers have used tests and questionnaires. In my opinion, these techniques have not yielded the desired results and I take the view that they should be replaced by qualitative forms of research.

In the fourth chapter, I give a sample of such a piece of qualitative research tackling the question of examples. I investigate a few dictionaries and conclude that there emerges no clear policy for examples in most of them, and that in others examples have not been utilised to the full. Finally, I analyse the question of authentic examples as opposed to made-up ones, concluding that both have advantages and disadvantages.

In the final chapter, I outline a new kind of dictionary. The recommendations and suggestions I make take into account the capabilities of modern computers, but are not dictated by them. So much so that my suggestions can be of some use in the compiling of paper dictionaries as well. In that chapter I start from a basic distinction between a *decoding* and an *encoding* part of the dictionary. I then consider both these activities from the point of view of beginning and advanced learners respectively. I discuss the utility of labels, synonyms, examples and definitions for the disambiguation of

translation options in the *decoding* process. In the case of *encoding*, which I consider more difficult, I first discuss how learners can choose their item, then how they can learn to use it. Examples again play an important role in this.

In the conclusion, I sum up the results of my research and suggest a few topics for further research.

1.3. METHODOLOGY

1.3.1. Data of this Research

The data I worked on are dictionaries. On a few occasions I did a test with subjects both at the *Universidade Federal de Santa Catarina* and at COBUILD. I considered three main categories of dictionaries: *bilingual, learner's*, and *native speaker's*. The first two categories are the ones most directly concerned with foreign language learners, whereas I include native speaker's dictionaries mainly because of their relationship to learner's dictionaries. The choice of these dictionaries was essentially guided by the languages of which I have at least a minimal command. Most attention was dedicated to reference works in English: English being the most studied language in the world, more people learn it and more people work on improving its dictionaries. In addition, some varieties of dictionaries only really exist in English.

The dictionaries listed below constituted my main data. A full list of the dictionaries I consulted can be found in the bibliography at the end.

Bilingual Dictionaries

1. *Basic Japanese-English Dictionary*. 1994. The Japan Foundation. Oxford University Press. Oxford.

2. *Collins Gads Danish Dictionary.* 1994. HarperCollins. London.
3. *Collins German Dictionary.* 1991. HarperCollins. London. Pons. Stuttgart. Dresden.
4. *Collins Robert French Dictionary.* 1994. HarperCollins. London. Robert. Paris.
5. *Collins Sansoni Italian Dictionary.* 1995. HarperCollins. London.
6. *Collins Shubun English-Japanese Dictionary.* 1993. HarperCollins. London.
7. *Collins Spanish Dictionary.* 1994. HarperCollins-Grijalbo. London.
8. *Dicionário Inglês-Português.* 1982. Record. Rio de Janeiro.
9. *Dictionary of English Synonyms and Antonyms.* 1992. Penguin. Harmondsworth.
10. *Euroglot Professional V.1.* 1994. Linguistic Systems. Nijmegen.
11. *Genius English-Japanese Dictionary.* 1988. Taishukan. Tokyo.
12. *Il Nuovo Dizionario Hazon-Garzanti.* 1991. *Hazon-Garzanti, Milano.*
13. *Longman Dictionary of Contemporary English.* 1995. Longman. Harlow.
14. *Martin's Concise Japanese Dictionary.* 1994. Charles Tuttle Company. Rutland. Tokyo.
15. Morsbach H. and Kurebayashi K. 1994. *Japanese Phrase Book and Dictionary.* HarperCollins. London.
16. *New Proceed English-Japanese Dictionary.* 1994. Benesse Corporation.
17. *Oxford Wordpower on Diskette.* 1993. Oxford University Press. Oxford.
18. *Oxford-Duden, Oxford-Hachette and Oxford Spanish Dictionary on CD-ROM* 1996. Oxford University Press. Oxford.
19. *Portuguese Dictionary.* 1991. HarperCollins. London
20. *Portuguese-English Dictionary.* 1963. Record. Rio de Janeiro.
21. *Proceed Japanese-English Dictionary.* 1988. Fukutake.
22. *Van Dale Nederlands Engels Woordenboek.* 1991. Van Dale. Utrecht.
23. *Word Selector Inglés Español.* 1995. Cambridge University Press. Cambridge.

Learner's Dictionaries

1. *Cambridge International Dictionary of English.* 1995. Cambridge University Press. Cambridge.
2. *Collins* COBUILD *English Language Dictionary.* First edition. 1987. HarperCollins. London.
3. *Collins* COBUILD *English Language Dictionary.* Second edition. 1995. HarperCollins. London.
4. *Collins English Dictionary and Thesaurus.* 1992. Electronic version 1.5. HarperCollins. London.
5. *Longman Dictionary of Contemporary English.* 1995. Longman. Harlow.

6. *Longman Language Activator*. 1993. Longman. Harlow.
7. *Oxford Advanced Learner's Dictionary of Current English.* 1972 (ed. 1963). Oxford University Press. Oxford.
8. *Oxford Advanced Learner's Encyclopedic Dictionary.* 1989 (ed. 1993). Oxford University Press. Oxford.
9. *Oxford Wordpower* 1993. Oxford University Press. Oxford.

Bilingualised Dictionaries

1. *Collins* COBUILD *Students' Dictionary. Bridge Bilingual English-Portuguese.* 1995. HarperCollins. London.
2. Konder R. 1982 (ed.1993). *Longman English Dictionary for Portuguese Speakers.* Ao Livro Técnico. Rio de Janeiro.

Monolingual Native Speaker's Dictionaries

1. *Bibliorom Larousse* 1996. Microsoft / Larousse.
2. Buarque de Holanda A. *Aurélio Eletrônico* V. 1.3. Nova Fronteira. Rio de Janeiro.
3. *Cassell Dictionary of Appropriate Adjectives.* 1994. Cassell. London.
4. *Chambers Dictionary on CD-ROM. V. 2.1.* 1993.
5. Covarrubias S. (ed. 1989). *Tesoro de la Lengua Castellana o Española.* Ed. Alta Fulla. Barcelona.
6. *CyberDico. Le Dictionnaire Multimédia de la Langue Française.* 1996. Goto Informatique. Tourcoing.
7. *Larousse Référence Électronique V1.11.* 1992-93. Larousse. Paris.
8. María Moliner 1996. *Diccionario de Uso del Español.* Edición en CD-ROM. Gredos. Madrid.
9. Meldau R. 1972. *Schulsynonymik der Deutschen Sprache.* Julius Groos. Heidelberg.
10. Mikhail E. H. 1994. *Cassell Dictionary of Appropriate Adjectives.* Cassell, London.
11. *Petit Robert sur CD-ROM* 1996, Société du nouveau Littré, Paris.
12. *Random House Webster's Electronic Dictionary and Thesaurus.* V.1. 1992.
13. *Roget's Thesaurus of English Words and Phrases.* 1852 (ed. 1987). Penguin. Harmondsworth.
14. *Van Dale Groot Woordenboek der Nederlandse Taal.* 1976. Martinus Nijhoff. 's Gravenhage.
15. Zingarelli N. 1998. *Lo Zingarelli 1998 in CD-ROM. Vocabolario della lingua Italiana.* Zanichelli. Bologna .

Others

1. *Bookshelf '94.* Microsoft Corporation, Redmond.
2. *Collins* COBUILD *English Collocations CD-ROM.* 1995. HarperCollins. London.

3. *Encarta '95*. Microsoft Corporation, Redmond.
4. *The Software Toolworks Multimedia Encyclopedia* 1992. version 1.5. *Grolier.*

1.3.2. Data Analysis

To analyse what are the needs of learners and their relationship to current dictionaries I consulted the literature on the subject searching for a methodology. I subdivided this literature into a few main groups: lexicography in general; history of lexicography; dictionary use and users; bilingual dictionaries; learner's dictionaries; and examples.

The research I considered to have the closest relationship to my topic was the one on *dictionary use and users*. It appeared to me that there is currently no agreement on what is a proper methodology for the analysis of user's needs. In this area, several methods have been used, none of them leading to very conclusive results. It occurred to me that the methods used were in some way flawed when reading an article by Hillary Nesi (1996) which I will discuss in greater detail in section 3.1.5. After coming to some very unexpected results in a test on dictionary examples, she herself suggests that shifting from quantitative to qualitative methods may in the future be a more reliable way of conducting research on the topic of dictionary use.

> Further investigations of the value of examples in learner's dictionary entries will need to develop a more sensitive method of measuring almost imperceptible developments in word knowledge resulting from dictionary consultation. Qualitative rather than quantitative methods may prove more appropriate for this purpose. (Nesi, 1996, no page number, electronic version.)

In my interpretation of the facts, most experimental, quantitative, research on dictionaries depends too much on decisions by the researcher which cannot have an objective basis and yields only the forecast results. This is true for *tests* as well as for

questionnaires. Indeed, sometimes one kind of research contradicts the other, even if both seem equally well buttressed by statistics and percentages (Harvey and Yuill, 1992, contradicting Laufer, 1992).

As is secondarily shown by research in closely related areas such as foreign language acquisition (Larsen-Freeman, D. and Long M. H., 1991), little is gained by applying methods borrowed from the exact sciences to what is, to a large extent, a social science. I shall expand on this briefly.

Tests and Questionnaires

In terms of methodology for the research on needs, I will discuss two different strategies: *tests* and *questionnaires* (Section 3.1.). In Chapter 3, I give an overview of the literature on the subject. Here I will limit myself to an evaluation of these techniques as a methodology for discovering the needs of dictionary users.

As regards *questionnaires*, it is my experience these only tell us what subjects think they do when they consult a dictionary, or what the researcher already knew they did. Researchers tend to ask either too obvious, or too difficult questions and the subjects are either unable, or understandably reluctant, to tell the truth.

On the other hand, *tests*, or *controlled experiments*, seem more objective, but in fact are not. So many variables have to be taken into account that doubts are cast on the results. *Tests* are reliable only when asking specific and well circumscribed questions such as 'can subjects tell made-up examples from authentic ones?' In this case,

researchers still have to choose the examples, but that is the only choice they have to make.

Particularly in the case of *tests*, researchers have to make a number of more or less arbitrary choices according to the case. Additionally, their subjects have to work in situations which cannot adequately replicate a real life situation. Researchers are generally aware of this and frequently apologise for it, but that does not necessarily make the results more trustworthy. In her research on the efficiency of dictionary examples and definitions, Laufer admits:

> But it should be remembered that the objectives of the experiment required an artificial situation of testing words out of context. In real life, people seldom learn words from monolingual dictionaries only. In the case of reading, a word is looked up in the dictionary and related to the text. ... In the case of writing, people rarely use words which are entirely new to them. They do, however, look up words which are partially familiar to them. Therefore, in a real-life situation, it would be reasonable to expect better results than in the experiment. (Laufer, 1993:138)

The fact is that these caveats, honestly pointed out by serious and dedicated researchers like Batia Laufer, show how questionable the results of their research can be.

The debate on the use of experimental methods exceeds the boundaries of this dissertation, but some clarification is required in order to dispel doubts that may arise on the nature of my own research. A quote from Popper's essay 'The logic of social sciences' seems appropriate to corroborate my claim on the inadequacy of methods taken from the exact sciences to elucidate lexicographic problems.

> There is, for instance, the misguided and erroneous methodological approach of naturalism or scientism, which urges that it is high time that the social sciences learn from the natural sciences what scientific method is. This misguided naturalism establishes such demands as: begin with observations and measurements; this means, for instance, begin by collecting statistical data; proceed, next, by induction to generalisations and to the formation of theories. It is suggested that in this way you will approach the ideal of objectivity, so far as this is at all possible in the social sciences. In so doing, however, you ought to be conscious of the fact that objectivity in the social sciences is much more difficult to achieve (if it can be achieved at all) than in the natural sciences. For being objective demands that one is not biased by one's value judgements – that is (as Max Weber called it), to be value-free'. But only in the rarest cases can the social scientist free himself from the value system of his own social class and so achieve even a limited degree of 'value freedom' and 'objectivity'. (Popper, 1994:67-68)

It is true that a comparison with the social sciences to which Popper refers is relative. Of course, the urge to free themselves from a social value system seems an irrelevant requirement for linguists. However, it is the possibility of interfering in the execution of scientific methods which counts here. Lexicographers have opinions and these are allowed to interfere in the design of most controlled experiments. It is not my intention to question the expertise or the honesty of the researchers involved, but both questionnaires and tests offer ample opportunities to manipulate the facts. No lexicographer carries out an experiment or writes a questionnaire without already having an opinion on the matter. Contrary to the exact sciences, s/he has tremendous possibilities of influencing the results.

The way dictionary research has been dealt with over the last twenty years gives the impression that the field is, to borrow a concept precisely from the exact sciences, in search of a new 'paradigm' as Kuhn would call it (Kuhn, 1970: passim), or a new 'episteme' as Foucault calls a very similar concept (Foucault, 1966: passim). In the case of lexicography, the old paradigm expressed itself through research questions such as 'is grammar necessary in a dictionary?', 'how to teach dictionary skills', 'problems of

classification of multi-word items', 'the inclusion of specialist vocabulary', 'translation equivalents in bilingual dictionaries'. These were the questions considered 'relevant'. That this old paradigm is now on shaky ground is shown by the fact that in the literature of the last twenty years, articles on relatively minor topics, such as those above, are on an equal footing with articles on the foundations of the discipline: the needs of those who are its prime beneficiaries. This is reminiscent of a paragraph in *The Structure of Scientific Revolutions*:

> When scientists disagree about whether the fundamental problems of their field have been solved, the search for rules gains a function that it does not ordinarily possess. While paradigms remain secure, however, they can function without agreement over rationalisation or without any attempted rationalisations at all. (Kuhn, 1970:49)

'Rules' is the methodology which current lexicographic literature desperately attempts to define. The current attempts to define the needs of dictionary users are a symptom of the urgently required 'rationalisation' of the area. Dictionaries as we know them are the result of material conditions and of ideological considerations of which some have disappeared. Indeed, a number of dictionary features, such as codes or abbreviations, the non-separation of *encoding* and *decoding*, are the result of a lack of space and consequent commercial considerations which ceased to be as inevitable as they used to be. The evolution of society and, most of all, the advent of the computer, which have to a large extent caused the switch in needs of the dictionary audience, also offer the possibilities to meet these new needs. Another example is the issue of alphabetic ordering, so characteristic of traditional lexicography, and which has given rise to research topics such as 'under what entries do learners look up the meaning of idioms'. In a computer programme, all items of a dictionary are now randomly

accessible and programs such as *Euroglot* have solved this problem in a straightforward and simple way.

Even so, articles are still being published on the problem of lemmatisation (Lorentzen, 1996), which is in my opinion a problem which belongs to the past. Furthermore, combining the solution of space and ordering, computers have now made it commercially feasible to realise an old lexicographic dream: split up *encoding* and *decoding*.

Some may argue that the computer is here presented as a panacea. I think it is, to the same extent as the invention of printing was a revolution in the 15th century. It was not the solution of mankind's problems, but it surely changed their appearance and the way they had to be dealt with.

Ideology

I have mentioned the ideological element. A methodology aiming at delineating dictionary user's needs should attempt to disperse a bit of the ideological fog[1] which surrounds dictionaries and was at least partially responsible for their appearance and content. I see the influence of ideology in two ways.

[1] The concept of *ideology* is a highly controversial one, and is used by many philosophical and sociological schools. I will not attempt to go into unnecessary detail here. Thinking of *ideology*, I would retain a fairly Marxist view of ideology in the sense that I understand it to be the system of ideas which justify the position of a particular class within society. This system of ideas may give a correct vision of society, but it might just as well not. In the case of dictionaries, an item intimately linked with the shaping of the modern state, it is difficult not to see any political motivation in their use, beyond the one that is publicly admitted. The introduction to the *Dictionnaire de l'Académie française* claims: 'Le Dictionnaire de l'Académie ne sera pas moins utile, tant à l'esgard des Estrangers qui aiment nostre Langue, qu'à l'esgard des François mesmes qui sont quelquefois en peine de la veritable signification des mots, ou qui n'en connoissent pas le bel usage, & qui seront bien aises d'y trouver des esclaircissemens à leurs doutes'. I consider this to be partially true, partially false, even if those who wrote this introduction were in good faith.

The first one is the more obvious one which can be exposed by placing dictionaries in their historical context. Dictionaries played a role in the consolidation of modern national consciousness by standardising the language. This is why native speaker's dictionaries, for instance, explain simple words to an audience which knows perfectly well what they mean. Seen from the angle of the aim of this thesis, it was through learner's dictionaries which were derived from native speaker's dictionaries, that foreign language lexicography was affected by this ideological factor. Learner's dictionaries were created to meet the needs of foreign language learners, but they were not compiled from scratch. They took native speaker's dictionaries as their model.

Secondly, dictionaries are not only tools, they are also objects which take their place and function in the 'système des objets' as characterised by a philosopher like Baudrillard (1968, 1970, 1972). Dictionaries are social objects, as emblems of literacy, as guardians of language and culture. In this function, they also have to respond to particular criteria which have little to do with their declared purpose. Instead of having only a usage value, dictionaries also have a significant symbolic value, and are not the mere answer to a mere need. As Baudrillard puts it:

> Aujourd'hui la consommation – si ce terme a un sens autre que celui que lui donne l'économie vulgaire – définit précisément *ce stade où la marchandise est immédiatement produite comme signe, comme valeur/signe, et les signes (la culture) comme marchandise.* (Baudrillard, 1972:178, author's emphasis)

The reasons for the compilation of a native speaker's dictionary are to some extent symbolic. They are produced as 'signe', as a cultural value. In the case of learner's dictionaries, their psychological value is not to be underestimated.

One of the symbolic functions of dictionaries, maybe the main one, is to preserve languages against 'corruption'. The compilers of one of the most famous and oldest national dictionaries, the *Dictionnaire de l'Académie française*, were convinced that the French language had in their century attained the highest degree of perfection and claimed:

> Il (le Dictionnaire de l'Académie) a esté commencé & achevé dans le siecle le plus florissant de la Langue Françoise; Et c'est pour cela qu'il ne cite point, parce que plusieurs de nos plus celebres Orateurs & de nos plus grands Poëtes y ont travaillé, & qu'on a creu s'en devoir tenir à leurs sentimens. On dira peut-estre qu'on ne peut jamais s'asseurer qu'une Langue vivante soit parvenuë à sa derniere perfection; Mais ce n'a pas esté le sentiment de Ciceron, qui aprés avoir fait de longues reflexions sur cette matiere, n'a pas fait difficulté d'avancer que de son temps la Langue Latine estoit arrivée à un degré d'excellence où l'on ne pouvoit rien adjouster. Nous voyons qu'il ne s'est pas trompé, & peut-estre n'aura-t-on pas moins de raison de penser la mesme chose en faveur de la Langue Françoise,... (*Introduction*, no page number, electronic version.)

And just like the members of the *Académie* by whose example he was inspired, Johnson stated in his *Plan of an English Dictionary* (1747):

> It was not easy to determine by what rule of distinction the words of this dictionary were to be chosen. The chief intent of it is to preserve the purity, and ascertain the meaning of our English idiom... (no page number, electronic edition)

However legitimate from a political or other points of view, from the moment it became the main function of dictionaries to preserve the 'purity' of the language, they stopped being tools and became repositories. They inserted themselves again in a long tradition, tracing back at least to the Middle Ages, of books not to be read but to be treasured.

Learner's dictionaries, tools derived from these dictionaries, preserve many of the characteristics of their models and instead of adjusting to needs, they have inherited part of the native speaker's dictionaries' function as symbols. To quote only one example,

the definition of *spanner* in the *Oxford Wordpower* as 'a metal tool with an end shaped so that it can be used for turning nuts', has no other significance, no other *raison d'être* than its mere presence. It has to be there because only by including a definition of *spanner* is this book entitled to be called a dictionary. It will probably not teach anybody what a *spanner* is who did not know it, and certainly not to a foreigner for whom this book was supposedly written. This is not the case for the illustration, of course, which the editors had the common sense to include.

I do not claim that dictionaries should not maintain their ideological function. Ideologies have a function. However, for the benefit of the part of these tools which is dedicated to solving linguistic questions, they only stand in the way. An identification of the 'bare' facts of the problem without ideological considerations is essential for research to be productive. As it is, the strength of the dictionary's symbolic value is one of the elements which has prevented researchers from approaching the question in a more fundamental way. As Kuhn puts it:

> Effective research scarcely begins before a scientific community thinks it has acquired firm answers to questions like the following: What are the fundamental entities of which the universe is composed? How do these interact with each other and with the senses? What questions may legitimately be asked about such entities and what techniques employed in seeking solutions? (Kuhn, 1970:5)

This reads like a striking metaphor of the field of current foreign language lexicography. It is no longer clear what are the fundamental entities that continue to compose the universe of lexicography and those that do not, such is the burden of history, ideology, and past material conditions. The development of initiatives such as bilingualised dictionaries, which are adjustments to a traditional model, indicates the

inadequacy of the old paradigm and how lexicographers are trying to define the new one.

New criteria

The task of defining new criteria appears to be a huge one. Remarkably, publishers have taken over the initiative from researchers and it is no bold statement to say there are at present more electronic dictionaries on the market than articles published on the subject. At the least there should be carried out a sound analysis of the learners' needs and a consequent reshuffling of the different components of existing dictionaries. Over the centuries lexicographers have produced a large amount of fine work and a mere reorganisation of this material would mean a huge step forward.

For a *tabula rasa*, however, techniques borrowed from the exact sciences have proven insufficient[2]. Until now, a *qualitative* approach has hardly been attempted and the accusation of subjectivism is feared by every contemporary researcher. But the objective veneer on research methods such as *tests* and *questionnaires* is thin. Instead of this, I prefer researchers with experience in the matter to interrogate themselves on their practice. In the words of Rousseau: 'La réflexion jointe à l'usage donne des idées nettes...' (Rousseau, 1973, Vol.1:235) I propose to submit dictionaries to an '*analyse de textes*' which is indeed more typical of literary studies and bears some relationship to structural methods as applied by Barthes (1957) and Baudrillard (1968, 1970). On the one hand, I look for the implicit criteria which guide lexicographic practice, as in my analysis of examples (Chapter 4). On the other, I analyse the look-up process as the

[2] This is, again, shown by what occurs in the area of second language acquisition. As Krashen (1996) shows in his account of a very proficient language learner, more can be learned from the experience of successful individuals than from so called scientific tests.

expression of the user's needs. To do this, I draw on 25 years of experience as a daily dictionary user and, why not admit it, the use I made of dictionaries to write this thesis.

In the course of this research I had more than once the feeling that I benefited most of all from consulting dictionaries and monitoring myself. As Descartes put it:

> non pas qu'il ne puisse y avoir au monde plusieurs esprits incomparablement meilleurs que le mien; mais pource qu'on ne saurait si bien concevoir une chose et la rendre sienne, lorsqu'on l'apprend de quelque autre, que lorsqu'on l'invente soi-même. (Descartes, 1951:80)

Due to the unpopularity of any subjective methodology and my rejection of much of past research I am conscious of the risk of giving birth to a 'ridiculus mus'. However, my disappointment with current reference works and the great need I have for them encouraged me to put forward suggestions which some may call unrealistic. There might be some truth in this. However, in *Émile*, and dealing with a much more consequential matter than mine, Rousseau says:

> Proposez ce qui est faisable, ne cesse-t-on de me répéter. C'est comme si l'on me disait: Proposez de faire ce qu'on fait; ou du moins proposez quelque bien qui s'allie avec le mal existant. Un tel projet, sur certaines matières, est beaucoup plus chimérique que les miens; car dans l'alliage, le bien se gâte, et le mal ne se guérit pas. (Rousseau, 1966:33)

As regards the form in which this thesis was written, some may object to its somewhat episodic character. This is a direct consequence of the fact that the bulk of this research consists of an observation of the facts and not of what has been written on them, however valuable this material in itself is. I would like to put forward the observation of Francis Bacon with a quote which has always seemed to me to speak directly to our present age, and which foretells by a few centuries Wittgenstein's work:

> The admiration of mankind with regard to the arts and sciences, which is of itself sufficiently simple and almost puerile, has been increased by the craft and artifices of those who have treated the sciences, and delivered them down to posterity. For they propose and produce them to our view so fashioned, and as it were masked, as to make them pass for perfect and complete. For if you consider their method and divisions, they appear to embrace and comprise everything which can relate to the subject. And although this frame be badly filled up and resemble an empty bladder, yet it presents to the vulgar understanding the form and appearance of a perfect science. (Bacon, 1952:122)

It is surely easier to weave a perfect web when one remains distant from the facts, and it is easier to write a philosophy of dictionaries than attempt to improve them. In this thesis I have examined all publications which were within my material and intellectual reach. But the examination of these convinced me that the confrontation with the different aspects of the problem would inevitably produce a less homogeneous picture than if I had worked on secondary sources.

Finally, I must recognise my debt to the work of John Sinclair which gave me some of the fundamental analysing tools used in this research. Even when I disagree with his views and they seem to me unproved and unprovable, they were sufficiently powerful to put one's brain to work.

Conclusion: What is qualitative lexicographic research?

Two points of view can be taken. The first one is to look at what has been produced and ask oneself: what purpose does this serve? For instance, taking the entry for *midget* in the *Oxford Wordpower*, one could ask what the information '/'mIdZIt/ noun [countable] a very small person' can be good for. The *pronunciation* is clearly aiming at a dictionary user wanting to speak; the *grammatical information* is aiming at a user either wanting to speak or write, and the *definition* aims at somebody who is *decoding,* reading or listening. One can then go one step further and see if the information aimed at the

speaking audience is sufficient and adequate and do the same for the information intended for the listening, writing and reading audiences. In this particular case, one could claim there is too much information for someone who only wanted to know the pronunciation of the word. The pronunciation, in turn, is not what the writing user wanted to know and the place taken up by it could have been used to provide this user with an example.

Doing this kind of exercise one quickly becomes aware of the shortcomings of dictionaries. This is necessary if we want to improve them.

Another type of qualitative approach starts not from the product, but from the person who is supposed to use this product. This kind of approach implies analysing the behaviour of specific dictionary users: the researcher her/himself, or a third person. I did this constantly with myself since it is a question which researchers have tried to answer for years, without much success. I may have deceived myself more than once. I cannot have deceived myself the whole time.

CHAPTER 2 LEXICOGRAPHY AND LEARNER'S NEEDS

2.1. DICTIONARIES IN A HISTORICAL PERSPECTIVE

Dictionaries used to tackle immediately evident, practical needs. In Ebla, around 4400 years ago, clay tablets registered Sumerian words and their Eblaite translations to facilitate communication between the two language communities. In the first century AD, some nine hundred years after the *Iliad* and the *Odyssey* were written, the Greeks found it increasingly difficult to understand the versified account of the eventful travels of their mythical ancestors, and scholars set out to compile commentaries on words that had become difficult to understand. In the Middle Ages this practice of glossing works continued to be common and they began to be collected in separate books, independently of their source texts. They became in this way a kind of hard words dictionary.

At the end of the fifteenth century, trade developed and travelling was becoming less hazardous. Hale sums up the period as follows:

> From the 1480s travellers had begun to include glossaries of useful foreign words as appendices to their narratives. Polyglot vocabularies were published with increasing frequency from 1477; François Garon's *Vocabulary of five languages: Latin, Italian, French, Spanish and German* proved so popular after its publication in Venice in 1526 that by the 1546 edition it had been extended to cover eight languages. Jacopo Strada, a scholarly collector and dealer in antiquities whose portrait was painted by Titian, dies in 1588 while working on an eleven-language dictionary. From the early sixteenth century multi-lingual conversation and phrase books started to appear as simple aids for merchants abroad; from the mid-century they broadened to satisfy those who wished to learn a foreign language in some depth. (Hale:1994, 159)

In fact, in 1447 an Italian-German word list was compiled for travelling purposes which can be considered the oldest still existing 'modern' bilingual dictionary. Up to

now, there was no doubt about what the purpose of these publications was and what the needs of their audiences.

During the revival of classical Latin in the Renaissance, a whole range of Latin, Greek and Hebraic bilingual dictionaries went to press. They supplied primarily a *decoding* need and were intended to facilitate the reading of classical and biblical texts. Secondarily, they assisted people in learning Latin. Paradoxically, the reintroduction of classical Latin was the death of Latin as a lingua franca.

During the same period, languages were expanding their vocabularies through the borrowing of words, mainly from Latin and Greek. It now seems strange to us that words like *involve, exactly, activity, education, sincerity* and *society*, to name but a few, are sixteenth century coinages. Even more astonishing is the fact that they were the feat of one single lexicographer, Thomas Elyot (Green, 1996: 87). It is not surprising that these new words had to be explained, principally 'to the ladies' (Brown, 1994) or to the 'more-knowing women, and the less knowing-men' (Osselton, 1983). It was for those people that the first monolingual dictionaries were compiled: the new words had to be clarified for those who did not know them.

In 1612, the Florentine *Accademia della Crusca* published a comprehensive list of all the words its members, the *crusconi*, considered to be authentically Italian. What they considered to be authentic Italian was the Italian they spoke, of course, and 'wrong' was what the others spoke. The Crusca *Vocabolario* is generally regarded to be the first national dictionary. It sought comprehensive coverage for its own sake. Since there was no point in explaining to an Italian audience what *cane* meant, *dog* was quite sensibly

defined as 'known animal' (Lepschy, 1984:174). The intention of the members of the *Accademia* was to exalt the Florentine language, rather than explain it. As I mentioned before, the purpose of 'purifying' the language, preserving it at a particular stage of its evolution and protecting it from 'corruption' is common to most subsequent 'national' dictionaries.

In 1634, the *Académie Française* was founded by Cardinal Richelieu with the task of compiling a French national dictionary which was eventually published in 1694. The intentions of its makers were in part political. The centralisation of power and the creation of a national feeling demanded a unified language. At that time, there were at least four different languages spoken in France.

After the French initiative, the English would have to wait another 61 years before they had their own national dictionary. Johnson published his *Dictionary of the English Language* in 1755. Although Johnson's dictionary was a private undertaking, it had a similar goal as its French and Italian predecessors: apart from standardising the spelling, it aimed at stabilising the language and protecting it from 'corruption'. It was intended to function as an official record of the language and both obscure and common words had their definition.

Whereas formerly the aim of dictionaries was to assist communication, intralingually or interlingually, this was now true only for bilingual dictionaries. From the *Crusca* dictionary on, the purpose of dictionaries ceased being exclusively communicative and became a matter of national interest. Monolingual dictionaries were now something which represented nationhood.

> The spirit of nationalism has often proved a driving force in the making of dictionaries. Scholars were quick to recognize that the compilation of a reliable and comprehensive dictionary was one sign of the achievement of their country's maturity, just as the lack of grammars and dictionaries indicated the dominance of a foreign power or the weakness of a truly national feeling. (Collison, 1982,18)

In addition, dictionaries embodied literacy and culture, just as the possession of a Bible meant adherence to Christianity. At least partially, the purpose of dictionaries changed from practical to symbolic. Dictionaries had become 'treasures', intended primarily to be owned and not to be used[3]. They began including as much as could be known of words, independently of the usefulness of this information for a general audience of 'cultured people': meaning, etymology (one of the first interests of lexicographers), grammar, synonyms, antonyms.

As a result, it is not clear anymore what the primary function of the monolingual dictionary now is. Surveys report that *meaning* is the first reason why people reach for a dictionary, *spelling* the second. However, this is based on the respondents' opinion, or on what they would like the researcher to think their opinion is. According to my own observation, the main use native speakers make of their dictionaries is *spelling* first, and *meaning* second. I agree this might be hard to admit since it implies that such a considerable piece of work and scholarship can in practice be replaced by a simple list of words. However, since Wittgenstein compiled one of these there should be no dishonour in admitting this.

Monolingual dictionaries are primarily symbols. Bilingual dictionaries, on the contrary, have little symbolism about them. In my experience, and rather surprisingly,

[3] If we look at the titles, this trend had begun early: : *Trésor de la langue française* (Nicot, 1530-1600), *Tesoro de la lengua castellana* (Covarrubias, 1611). The Dutch and German words for vocabulary – *woordenschat* and *Wortschatz*– reflect this, as well as the Latin word commonly used for dictionaries: *thesaurus*.

bilingual dictionaries do not have the aura which monolingual dictionaries have about them. The borderline between the two kinds of dictionaries became blurred from the moment someone, presumably Hornby, started using monolingual dictionaries in foreign language teaching. In 1947, the first learner's dictionary was published.

2.2. Learner's Dictionaries

At the end of the nineteenth century, the focus of foreign language teaching started changing slowly from merely understanding to producing the language. It was the Second World War which gave this tendency a decisive impetus, resulting in dramatic changes in language teaching and learning pedagogy in the sixties (Rivers, 1975). The 'communicative approach' in the eighties made it clear that the emphasis now was on production. In 1947 Hornby had published the *Oxford Advanced Learner's Dictionary.* Compared to normal monolingual dictionaries, the language used for defining was simpler, the number of entries and senses reduced, and there were more examples.

From a commercial point of view, learner's dictionaries were bound to be a gold mine. They are an intralingual bilingual dictionary for anyone studying English, independently of what their first language is: they translate hard L2 into easy L2. There are, however, a few problems with this kind of dictionary, some of them due to the fact that they were derived from native speaker's dictionaries. The most obvious one is that they feel obliged to define every word, even if the words used in the definition are in many cases inevitably more difficult than the word defined.

The market got a bit confused by the sudden appearance of a new kind of dictionary and by 1981 this was not yet clear to the main target audience of these

dictionaries. A. P. Cowie, in a thematic issue on 'Lexicography and its pedagogic applications', illustrates the attitude towards learner's dictionaries at the beginning of the eighties:

> Yet until quite recently there was no correspondingly widespread understanding among teachers that the scale and character of the information provided in a general EFL dictionary was radically different from that to be found in works intended for the native user. In particular, it was not generally understood that the learner's dictionary was designed to help with production as much as interpretation. (Cowie, 1981:203)

Most people using learner's dictionaries, even teachers, would still now be hard-pushed to say what the difference is between a normal monolingual and a learner's dictionary.

When looking critically at a learner's dictionary from a *decoding* point of view, one can affirm that these dictionaries presuppose at least a sufficiently good command of the language to understand the definition. If a learner does not understand simple words such as *cat* and *dog*, it is useless to look them up in a learner's dictionary. One of the easiest learner's dictionaries, the *Oxford Wordpower*, defines *cat* as

> a small furry animal with four legs and a tail. People often keep cats as pets.

In this case, it is obvious that a better command of English is needed to understand the definition than the word it defines. Moreover, from a *decoding* point of view, the reduced number of entries and senses in learner's dictionaries is a disadvantage.

When one looks at learner's dictionaries from an *encoding* point of view, learner's dictionaries have drawbacks as well. First of all, and most evidently, one has to know what to look up. This means that in most cases a bilingual dictionary will have to be consulted prior to the learner's dictionary. The learner's dictionary then is used to

double-check what was found in the bilingual dictionary. However, in a great many cases, somebody with a command of English good enough to understand the above mentioned definition of *cat,* will be interested in usage information, i.e. *syntax* and *collocations*. This implies that features which transmit this kind of information, such as examples, will be very important. However, the ordering of the information within the entry is based on 'senses', and not on usage. This makes the search cumbersome and the results doubtful. And since it appears that information on syntax or collocates is paradoxically no priority for learner's dictionaries, these are neither ideal *encoding* tools nor ideal *decoding* tools. I will discuss the consequences of this as regards examples in chapter 4.

2.3. Bilingual Dictionaries

Bilingual dictionaries have not so far shown themselves to be flexible enough to meet the new *encoding* needs. When one compares present-day bilingual dictionaries to sixteenth century ones, one sees surprisingly little difference. It is true that sixteenth century dictionaries were astonishingly good, and until recently there was not much reason for change. As long as *decoding* was the main concern, traditional bilingual dictionaries tackled their task adequately. Examples, which are primarily an *encoding* feature, are still rare and, of these examples, fixed expressions take up most of the space.

Even so, bilingual dictionaries are also expected to work precisely as *encoding* tools, since their target audience includes native speakers of two different languages. The *decoding* information of one audience is supposed to be the *encoding* information of the other one. Even so, since bilingual dictionaries are most frequently produced by

'mono-national' publishers, emphasis is given to the interests of the national audience. Hence the customary lack of balance between the two halves of these dictionaries in number of pages.

Looking at it from the point of view of dictionaries as commercial undertakings, recent international joint ventures show that publishing companies think it is time for a new approach (*Oxford-Hachette; Oxford-Duden; Collins-Robert*, etc.). New bi-national bilingual dictionaries have certainly changed lexicography in the sense that an equal amount of attention is given to the native speakers of both languages. However, up to now new experiments have mainly made dictionaries thicker and heavier, without introducing radical innovations. An *encoder* –the new audience– needs a large amount of information. If one compiles a French-English dictionary that has to be equally useful to English and French users, the ideal amount of information is immense. As the *Oxford-Hachette* or *Collins-Robert*, though admirable achievements, show the result is cumbersome to use and still incomplete.

2.4. BILINGUALISED DICTIONARIES

A relatively recent phenomenon in foreign language lexicography is bilingualised dictionaries. These are a mix of bilingual and learner's dictionaries. They give a definition as in a learner's dictionary, together with a translation. Sometimes they only translate the definition. This technique has the apparent advantage of settling the question whether it is better to give learners a definition or a translation, since it gives them both. In the case of conscientious learners who read the whole entry, one may say this is an improvement. The constantly reprinted *Longman English Dictionary for*

Portuguese Speakers is an example in its kind. Thanks to a simple Portuguese wordlist at the end of the dictionary, this dictionary comes close to the ideal beginners' encoding dictionary. It is surprising that other learner's dictionaries have up to now neglected this simple and useful device.

However, to some extent the restrictions which I formulated for bilingual and learner's dictionaries apply to this kind of dictionary as well. For the purposes of *decoding*, these dictionaries have too few entries and are only helpful in the case of relatively easy texts. For most hard words, learners will still have to consult another dictionary.

A dictionary is a tool and the main activity of a dictionary user is not usually 'looking up a word in a dictionary'. It is therefore likely that even in dictionaries which provide a translation alongside a definition, users will turn immediately to the translation and not waste time reading both. This translation might not give an exact idea of what the word means, but since the difficult word was found in a context –the text the learner is reading–, this context will help the learner to make the necessary adjustments.

Seen from the angle of *encoding*, the criticisms that apply to learner's dictionaries apply to bilingualised dictionaries as well. Here as well an *encoder* has to know which word to look up and this is a problem these dictionaries do not resolve. On the other hand it should not surprise us that in a test conducted by Batia Laufer (1994) subjects equipped with this kind of dictionary scored better than others who only had a bilingual or a monolingual dictionary at their disposal. Since these dictionaries are often a

learner's dictionary with a translation added, the result is of course a net increase in information.

A somewhat different kind of bilingualised dictionary is the COBUILD *English-Portuguese Bridge Bilingual.* Here the definition of the English word, as it appears in the COBUILD *Student's Dictionary*, was translated into Portuguese maintaining the entry word in English. As the name *Bridge* suggests, it was the intention of the lexicographers to make a dictionary that made a bridge between a bilingual English-Portuguese dictionary and an English learner's dictionary. The information is identical to that found in the COBUILD *Student's Dictionary*, but in Portuguese. Here is an example:

> **begrudge**, se você **begrudge someone** algo, você sente que essa pessoa não merece isso e sente inveja dela por tê-lo.

Probably the purpose was to show the syntactic constraints of the word in the L2, but it is doubtful whether this is effective (Humblé, 1996). Even so, one has to welcome the inventiveness of, again, John Sinclair and his team and the courage they showed in trying out something new in an area where the vastness of the task makes people prone to conservatism.

2.5. The Longman Language Activator

From a strictly *encoding* point of view, the most successful lexicographic undertaking of the last few years was the *Longman Language Activator*. This is a case in which lexicographers started from scratch to try and match a product to the needs of the user. The result is a considerable leap forward. The previous knowledge in the L2 which users are supposed to have is the superordinate of the word they are looking for. From

there on, the dictionary leads the learners to an alternative that expresses their thought more adequately (e.g. find *overdo* by looking for *exaggerate*).

A criticism one can make of the *Activator* is that, in spite of its 1600 pages, its innovations and its size are, all things considered, not ambitious enough. There are, principally, not enough examples and for advanced learners there might be too few entries since only the most frequent superordinates were included. Even so the *Activator* is a large volume and the fact that it was printed on paper must have limited the inclusion of more entries and examples. One can only regret that in a time in which paper dictionaries are being put on CD-ROM as they are, Longman put on paper a dictionary which has all the characteristics of an electronic dictionary.

Conclusion

At the beginning of their history, dictionaries were compiled with a precise purpose in mind: to facilitate communication in the same or in another language. In both these cases, the aim was *decoding*. With the advent of modern states, monolingual dictionaries began to assume a symbolic function. Through learner's dictionaries, a part of foreign language lexicography also participates in this symbolic function, and this confuses the relationship between the dictionary and the user's needs.

Nowadays the needs of foreign language learners are changing and foreign language lexicography is seeking ways to adapt. Learners ask for more efficient tools for *encoding*, which requires more information than *decoding*. Technology, however, offers a solution to lexicography's oldest problem: space. The question lexicographers are now faced with is how to combine needs and the technological means to meet them.

In spite of its obvious benefits, lexicography has as yet made little use of computer technology, because no one has a precise idea of what has to be done. Solving this question implies research as to the kind of dictionary learners need for *encoding*. Once it is accepted that *decoding* and *encoding* have to be separated, the problem is reduced to knowing what an ideal *encoding* dictionary is. Even if one can always think of improvements, a number of bilingual and monolingual dictionaries are perfect *decoding* tools.

To know what an *encoding* dictionary should be like, one has to know what sequence learners follow when they succeed in using a new word correctly. A possible starting point is the L1, but it could be a word in the L2 that is close. Alternatively, they may start from the right L2 word and want more information on its usage. If one succeeds in mapping the learner's look-up sequence, one knows what the structure of the ideal dictionary has to be.

As will become clear in the next section where I review the literature on the subject, over the last few years much research has concentrated on discovering the answers to the question: What do learners do when they consult a dictionary and what helps them best? Once this question is answered, a new kind of dictionary can be designed.

CHAPTER 3 REVIEW OF CRITICISM

In what follows I review some of the research carried out mainly by Tomaszczyk, Atkins, Nesi, Laufer and Béjoint on dictionaries and their users. Next, I will discuss Sue Atkins's article (1996) 'Bilingual dictionaries: past, present and future' to discuss the potential of new technologies for foreign language lexicography.

3.1. Dictionary Use and Users – Questionnaires and Experiments

Over the past twenty years a number of articles have been published on the subject of dictionary use, and interest in the subject is increasing. Since the emphasis is no longer on *decoding*, and *encoding* is becoming more and more important, the way in which dictionaries are compiled must be reconsidered. To discover the needs of the target audience, researchers have used mainly two techniques: questionnaires and tests. The former ask users what they use their dictionaries for, while the latter test users and dictionaries via controlled experiments. I will now dedicate some attention to each of these research techniques.

3.1.1. Questionnaires

Questionnaires ask users questions of the kind: if they use a monolingual or a bilingual dictionary; if they use it for reading or writing; if they use it to look up meaning, spelling, etymology or grammar. The subjects are always students and almost exclusively study English as a foreign language.

The first comprehensive questionnaire on dictionary use was elaborated by Tomaszczyk (1979). He questioned foreign students at American and Polish colleges,

and foreign language students at a Polish university. His conclusions were, among other things, that students preferred bilingual to monolingual dictionaries, and that *meaning* and *spelling* were the main reasons for using a dictionary.

In 1981, Béjoint conducted a survey which was more detailed, but less comprehensive in terms of subjects. Béjoint wanted to discover how French students of English used their monolingual English dictionaries. He asked 122 subjects 20 questions ranging from 'Do you own a monolingual English dictionary?' to 'Which types of information do you look up?'

According to the results of Béjoint's survey, half of the students used their monolingual English dictionary once a week and *meaning* was the main reason for this. Half of the students said they sometimes browsed through their dictionary without a special purpose and 89% admitted they had read the introductory matter 'less than thoroughly'. More than half responded that they never used the dictionary codes, while 70% affirmed they used examples and quotations, 68% used synonyms and 24% used pictures. Some 50 % of the students thought their dictionaries were too simplified and 45% thought they were just right; 77% claimed satisfaction with them.

Béjoint devotes the introduction to his report to a succinct but keen analysis of the needs of the foreign language learner. He stresses the difference between *encoding* and *decoding* and sums up the matter: 'the best dictionary for *decoding* is the one that contains the largest number of entries' and 'the best dictionary for *encoding* is one that provides the most detailed guidance on syntax and collocation'. (1982:210) Béjoint

knows more about dictionary use than his subjects, and his thoughts on the topic exceed by far the results of his survey.

However, Béjoint's questionnaire is certainly one of the best of its kind. But when looked at from a distance of more than fifteen years one is astonished that the questions make little use of the *decoding/encoding* distinction when the author's analysis shows a clear awareness of how this affects dictionary use. For instance, the question 'Which types of information do you look for most often in your dictionary?' is difficult to answer, or the answer is difficult to interpret, if one does not know the purpose of the look-up. Furthermore, 87% of the interrogated subjects claimed that they looked for meaning, 53% for syntactic information, 52% for synonyms, 25% for spelling and pronunciation, and 5% for etymology. However, it is likely that students look for meaning when they *decode* and for *syntactic information, synonyms* and *spelling* when they *encode*.

3.1.2. Criticisms of Questionnaires

Criticisms of questionnaires are voiced by Bogaards (1988), among others. He mentions the want of homogeneity of the investigated subject samples, the lack of depth of some analyses and the scarcity of interesting answers due to the vagueness of the questions. Finally, Bogaards laments that often the questions are drawn up in such a way as to suggest the answers. He does not fundamentally criticise the method, however, presenting in the same article the results of his own questionnaire.

In 1987, Hartmann, himself the author of a questionnaire (1983), sums up his criticisms as follows:

> Research into dictionary use still tends to be small-scale, often non-representative, non-comparable (even contradictory!), non-correlational, and non-replicable. This partly explains the tentative nature of many of the findings, which frequently have the status of 'informed opinion' rather than valid generalisation. (Hartmann, 1987:27)

Although he does not say so explicitly, Hartmann's criticisms refer essentially to questionnaires since his survey includes two tests only. Besides, he says:

> (tests) are more difficult to devise and possibly therefore rarer... Even more complex techniques like controlled experiments have not been used at all. (ib.)

In my view, questionnaires are subject to caution for two reasons. First, one assumes the subjects have a knowledge of linguistic concepts and a lexicographic awareness which I think they generally lack. Do they know what an 'idiom' is? Do they really, as Béjoint claims, look for 'syntactic information' in more than half of the cases, or is this what they think they do, or remember they did? Do they know what 'syntactic information' is at all? Do 68% look up idioms 'very often' or do they not remember when they looked up, e.g., synonyms but remember idioms as something special?

A second assumption is that the subjects tell the truth. In a questionnaire I personally devised, one of the requirements was to write down the name of one's dictionary without looking at it. Nearly all the subjects appeared to be able to answer this question! When a teacher administers the questionnaires, as is often the case, the danger exists that subjects take it as an exam, and answer what they think the teacher wants them to answer. Consequently, one tends to agree with Glynn Hatherall:

> Are subjects saying here what they do, or what they think they do, or what they think they ought to do, or indeed a mixture of all three? (quoted by Hartmann 1987:22)

Nuccorini's research (1992) on dictionary use among students and teachers is a recent example of how difficult the issue of reliability of surveys is. Repeatedly she catches her subjects out in contradictions, concluding blandly that 'teachers' forms are much more reliable than students' forms'.

In an overview, Battenburg (1991) lists 11 research projects carried out between 1979 and 1990. Most of them consist exclusively of questionnaires while a few have a test element. These studies all come to similar conclusions, first, that learners prefer bilingual to monolingual dictionaries and, second, that they should be taught more dictionary skills.

Questionnaires were the first indication of a changing relationship between the lexicographers and their audience. Seen from a distance of almost twenty years, the results of these first questionnaires seem obvious, although they might not have been so at the time. The conclusions which they reached are now firmly accepted and it is not surprising if more recent questionnaires only confirm those of the first ones. One can safely conclude that this research method has by now yielded all the results it is capable of.

Questionnaires have an obvious limitation: they only investigate what *is* and not what should be or could be. The only practical conclusions to be drawn from them are what dictionary features learners use and should be further developed, and what features they do not use and should therefore be eliminated.

Since questionnaires are heavily dependent on what subjects *consciously* know or think they know, it seems logical that controlled experiments, or tests, would at some point take over.

3.1.3. Controlled Experiments

In this section, I will discuss some of the most important tests carried out to gauge various aspects of foreign language dictionaries. The first experiment testing dictionary use was carried out by Bensoussan, Sim and Weiss (1984). It deals with the impact of bilingual and monolingual dictionaries on reading comprehension. Surprisingly, no noticeable difference was detected between the results of one group of students allowed to use a bilingual dictionary in answering multiple choice questions in EFL reading test, and another one which worked without dictionary help.

In 1992, Batia Laufer tested the efficiency of corpus-based versus lexicographer examples in comprehension and production of new words. The results indicated that the innovative 'authentic examples' as used by COBUILD, did not improve learners' performance. On the whole, lexicographers' examples appeared to be more efficient. Learners, so it appeared, did not write better sentences after the test-words had been illustrated by means of authentic examples and made-up examples were deemed to be more helpful.

In 1993, Laufer investigated 'The effect of dictionary definitions and examples on the use and comprehension of new L2 words'. She started from findings in first language vocabulary development research which show that:

> children below puberty define words by describing and using them in sentences, while the older ones (ages 10-14) tend to define words in abstract and generic terms... These findings would suggest that it is the definition rather than the example that is most beneficial to the average user of the monolingual dictionary, since such dictionaries are most often used by adolescents and adults. (p.133)

In her test, Laufer asked 43 first year Hebrew university students for the translation of 18 words. Some had the help of a definition, some of an example, and some of a definition in combination with an example. Her conclusion was that, for *decoding*, combined entries work best but that an example alone is worse than a definition alone. In addition,

> ...the understanding of new words improves more when a definition is added than when an example is added (...) in production, unlike in comprehension, an example alone will be just as efficient as a definition alone'. (Laufer, 1994:139)

In a 1994 research project briefly quoted above, Laufer and Melamed investigated the efficiency of monolingual, bilingual and bilingualised dictionaries. Two groups of EFL learners totalling 123 subjects were tested on their *decoding* and *encoding* abilities, using 15 low frequency words. Although there were differences depending on the activity, the overall conclusion was that bilingualised dictionaries, in which subjects have access to both definitions and translations, are more effective.

3.1.4. Atkins and Varantola

The fact that lexicographers are seeking to develop new measuring methods is illustrated by the last experiment to date (Atkins and Varantola, 1997). Its aim was 'to monitor the dictionary look-up process in as natural a situation as possible'. (1997:2) One hundred and three people, (71 in Oxford and 32 in Tampere) were asked to carry out a translation task *to* or *from* their first language. This was Finnish for the 32 students

in Tampere. Those in Oxford were of mixed linguistic origin, 38 of them having English as their first language. Working in pairs, one of the subjects would use the dictionary and the other would write down what the first one did. The result of the experiment is a written account of no less than 1.000 dictionary look-ups.

The conclusions of the researchers are that: subjects prefer bilingual to monolingual dictionaries; 90% of the look-ups are related to *encoding* tasks; advanced students use the dictionary more than beginners; in the case of a translation into English as an L2, people first look up the word in a bilingual dictionary and afterwards in a monolingual one; finally, advanced learners were less easily satisfied with their dictionary than beginners. In spite of the comprehensiveness and the large number of subjects of this test, these results are not really innovative.

As the researchers themselves observe, one of the problems of this test was that the evaluation of the success rate of the look-ups was entirely left to the discretion of the subjects. They themselves had to decide whether their use of the dictionary had given the expected results. It would have been interesting to know in a more objective way if the dictionaries had supplied the correct information, and if the users were able to retrieve it. Without this knowledge it becomes difficult to agree with statements like: 'The fact that in 59% of the cases the dictionary users pronounced themselves satisfied with the results of their search is encouraging' (p.24). Mainly when, on the other hand, Atkins and Varantola claim that 'the failure rate of 40% (of the total number of look-ups) cannot be due only to some inadequacy on the part of the dictionaries involved:

inadequate strategies and unrealistic expectations on the part of the user must also contribute to this figure'. (p.27)

Another flaw of Atkins and Varantola's test is that, although the authors admit that 'the situation will change when new dictionaries are compiled for electronic access only', they do not take this into consideration. At one point they respond to an alleged lack of information on collocates: 'there is a physical limit to the amount of information which a bilingual dictionary can contain'.(p.31) The truth is that with present electronic storage capabilities, there is not.

Atkins and Varantola's test draws a few interesting conclusions and the accumulated material should be a gold mine for anyone tackling user's needs with qualitative research methods. Indeed, the fact that the analysis of the data was almost exclusively quantitative is the main criticism one can make of this otherwise very fine piece of work. The few qualitative analyses of specific look-ups are far more interesting than the overall, massive, quantitative conclusions.[4]

Conclusion

In the case of *questionnaires* it is the *learners*, rather than the *dictionaries*, who are being evaluated. Atkins et al. (1987:29) leave no doubt about this: 'The purpose of this research project is to discover how effective a learner the student of English as a foreign learner is when working with a bilingual and/or monolingual dictionary'. Not surprisingly, questionnaires typically conclude that learners do not make a proper use of

[4] On page 9, 10 and 11, the authors give an account of a number of look-ups which could lead to interesting conclusions and hints to improve the dictionaries used in the experiment. Maybe these conclusions were considered too 'small-scale' to be worthy of much attention. That is a pity.

the tools at their disposal. Their authors frequently insist on the importance of teaching 'dictionary skills' (Béjoint, 1981: passim; Nuccorini 1992, 1994; Meara and Nesi:1994).

In the case of *tests*, on the other hand, it is primarily the *dictionary* which is being evaluated. It is therefore interesting to see that over the course of time emphasis has shifted from questionnaires to tests. Emphasis is now more on adapting the tool to the user than the user to the tool. Tests tend to end up blaming the dictionaries for the bad results of the learners' look-ups and questioning the dictionaries' effectiveness. Surprisingly, however, these same tests often conclude that lexicographic innovations, or even dictionaries *tout-court*, make little or no difference (Bensoussan et al., 1984; Jain, 1981; Laufer, 1992, 1994; Nesi, 1996). Even so, one has the impression that the authors of this kind of test are also convinced that dictionary users are slightly naughty (Nuccorini, 1994), or adepts of the 'kidrule strategy' (Meara and Nesi, 1994) looking only at the first sense they come across.

3.1.5. Hillary Nesi's Research on Examples

In my opinion, Hillary Nesi's research on examples shows the limits of the exact scientific method as applied to dictionary research and points at the same time to a new way of tackling the problem. In 1996, Hillary Nesi (1996b) published an article in which she gave an account of an experiment she carried out with forty EFL learners to discover to what extent dictionary examples were helping them to produce sentences featuring words previously unknown to them.

> Two versions of the test were prepared, using entries taken from the second version of the *Longman Dictionary of Contemporary English* (LDOCE). In each version, example sentences and phrases for half the target words were removed. In version A examples for the first nine target words were removed, but examples for the last nine words were retained. In version B examples for the first nine target words were retained, but for the last nine words were removed'. (No page number, electronic version.)

The results of the test were surprising, not in the least for the researcher herself. Indeed,

> No significant difference in appropriacy was found between language produced with access to examples, and that produced without access to examples. (...) These results are important because they seem to challenge lexicographers' beliefs regarding the value of examples'. (Ibid.)

The fact is that these results challenge the beliefs of anyone who ever studied a foreign language, and invite us to have a critical look at the method used in this experiment. Although this test was conceived and conducted following apparently objective guidelines, it is a good example of how this kind of research can go wrong.

At several stages of her research, Nesi had to make choices. Choices involve opinions and a particular understanding of the facts, based on what one knows. The first choice concerned the subjects. Was it possible to know what they were like and if the sample was homogeneous? Although a specific test was carried out to have an idea of the learners' background in terms of previous vocabulary knowledge, this knowledge is only measurable in terms of number of words, and not in terms of the more general knowledge which supports the lexical knowledge. Surely, to know what 'shot' means is a different kind of knowledge for a medical student than for a language teacher. Furthermore, it is difficult to know what the examples in the LDOCE meant for each of the test students in particular. Let us take the sentence selected in the LDOCE to illustrate *perpetrate*: 'It was the managing director who perpetrated that frightful statue

in the reception area'. (The researcher herself has doubts on the appropriacy of this example.) Western learners of a certain age and experience will have no difficulty in identifying the irony in this example and it may even have some use for them. For other learners, however, this example will be meaningless. This is an extreme case, but it is difficult to say what happened with any of the test examples in the heads of the subjects. Consequently, it is difficult to say what the impact on their learning was.

Secondly, another choice which Nesi had to make was the selection of the words the subjects would be asked to use in their sentences. It eventually became: *enlighten, err, gravity, incorporate, intersect, perpetrate, retard, rudimentary, symptom, version, agitate, civic, clarify, collide, compute, controversy, interact* and *interlude.* Inadvertently[5], all of the 18 words are of Latin origin. This means these are all words which are typical of a particular kind of discourse (written, intellectual, non-fiction). In my experience, and for reasons I can only guess at, words of this kind are easier to learn than Anglo-Saxon and this for Western and non-western learners alike. If the subjects had themselves a Neo-Latin background on top of that, a circumstance not revealed in the article, to grasp the meaning of these words would be utterly simple. We could even consider the difference in impact of examples on learners coming from different language families: agglutinative, analytic, incorporating, or inflectional. The variables are countless. Sadly, when researchers do not direct their choices so as to make their hypothesis come true, the results are difficult to assess and sometimes apparently contradict common sense. In the research under discussion this was so much the case that the researcher herself cast doubts on the validity of her conclusions, claiming that:

[5] Or maybe because English native speakers tend to find words of Latin origin more difficult than others. I think most foreigners would disagree with this.

> Further investigations of the value of examples in learner's dictionary entries will need to develop a more sensitive method of measuring almost imperceptible developments in word knowledge resulting from dictionary consultation. (...) Qualitative rather than quantitative methods may prove more appropriate...' (1996: no page, electronic version.)

It was this last reference to qualitative instead of quantitative research which seemed to me of the utmost interest. It might mark a turning point in lexicographic research. I will come back on this in section 3.3.

3.2. Dictionaries and Computers

Hillary Nesi, again, wrote an account of three electronic learner's dictionaries, *The Longman Interactive Dictionary* (1993)), *The Electronic Oxford Wordpower Dictionary* (1994) and *Collins* COBUILD *on CD-ROM* (1995). She claims the following are the 'areas in which electronic dictionaries can excel':

> They can cross-reference within and between sources published separately in book form, they can provide direct links to other computer applications, they can enable 'fuzzy' and complex searches, and they can interact with users to help develop vocabulary and dictionary skills. (No page, electronic version.)

Until now, however, electronic dictionaries have exploited only a few of these possibilities and, as Sue Atkins stated, existing dictionaries 'may even come to you on a CD-ROM rather than in book form, but underneath these superficial modernisations lurks the same old dictionary'. (Atkins, 1996:515)

Sue Atkins' research project (1996) is to date the only one which makes an attempt to look at dictionaries bearing in mind computer capabilities. She claims that technological advances have made it possible to meet the needs of foreign language learners much better than before and that, although electronic dictionaries have been published in a number of languages, the potentialities of the computer have hardly been

exploited. Atkins sums up the most important features which an electronic dictionary could have, showing a tremendous insight into the question.

- hypertext functionality eliminating linear text restrictions and opening the way to new types of information by offering new ways of presenting it;
- no space constraints other than the need to avoid swamping the user;
- no distortion of the source language description by the needs of the target language;
- flexible compiling liberated from alphabetical order;
- alternative ways of presenting the information, as for example graphics;
- rapid access to large amounts of lexicographical evidence in corpora;
- large scale user customisation. (1996:527)

Unfortunately, the scope of Atkins' research is limited to bilingual dictionaries. She sums up a few of their defects which could be remedied in electronic versions and mentions, for instance, the high level of redundancy which makes it cumbersome to use bilingual dictionaries, particularly for *decoding*. Apart from that, she claims, bilingual dictionaries show gaps in the coverage of neologisms and polysemous words. Additionally, there is at times distortion of the source language part because of the target language: particular senses of a word are grouped only because they happen to be translated by the same word. Finally, current bilingual dictionaries lack information on collocations and synonyms. These defects, Atkins claims, can be corrected by means of a 'multilingual hypertext lexical resource' using databases in several languages. In this 'resource',

- the monolingual databases are real;
- links (including metalanguage and instructions) between database items are real;
- the dictionaries themselves are *virtual*. (p. 531)

Atkins admits there are practical and commercial obstacles to this project[6]. In my opinion, it has a few other flaws as well.

First, I am not sure if this new model, instead of starting from the needs of the dictionary user, does not start from an already existing product, the bilingual dictionary as it is today. This prevents us from tackling the problem at its root and building up the solution from there. Additionally, I had the impression that the theory of semantics used to design this new electronic dictionary model, Fillmore's 'frame theory', put a heavy burden on the dictionary right from the start. To have a theoretical understanding of the area one is dealing with is natural. To let this theory be the guiding principle for the compiling of a dictionary is a different thing. This is what one concludes after having a look at model entry *crawl* on the dictionary's homepage where anybody can have a look at it. It is of undeniable linguistic interest and shows clearly how the word relates to others. However, does the mind of the intended user operate this way?

Another criticism I have of Atkins' plan is that she limits her proposals to bilingual dictionaries, while there is no technical reason why the boundaries between different lexicographic genres (monolingual, bilingual, thesaurus, synonyms, etymology, etc.) should be maintained. There is even the question whether one of the starting points of a new electronic dictionary should not be the integration of the various genres into one. This does not mean putting all the different dictionaries on one CD-ROM, but integrating the various parts of which they each consist into a single tool. As Atkins herself established in her most recent survey (Atkins and Varantola, 1996), *encoders* use several dictionaries to obtain a single piece of information.

[6] Atkins' model is complex and ambitious and she doubts if an editor will take the risk to compile it. The sample on the

A new electronic dictionary should take advantage of the almost unlimited capacities of electronic storage and put to use all the different kinds of lexicographic and encyclopedic knowledge accumulated over the centuries. From an electronic point of view, several sources of information can be available at the same time. It is possible to match the different parts of a dictionary, indeed of several dictionaries, with the different steps of a learner's query, merging a number of search utilities regardless of what they were originally designed for. Some of this material will be ready for immediate incorporation into some new structure, other elements will have to be adapted.

3.3. The Need for New Kinds of Research

> There are two direct routes to more effective dictionary use: the first is to radically improve the dictionary; the second is to radically improve the users. If we are to do either of these things – and obviously we should try to do both– the *sine qua non* of any action is a very detailed knowledge of how people use dictionaries at present. (Atkins and Varantola, 1997:1)

Over the last twenty years, lexicographers have been studying the lack of correspondence between dictionaries and their users. The problem has been tackled in several ways, the main ones being questionnaires and tests. These have by now brought forth all the information they could possibly produce. At the same time, technological innovations make it feasible to disregard any restrictions placed on the imagination by lexicography's traditional hindrances: lack of space and linear ordering.

Methods such as questionnaires make an epistemological leap from *behaviour* to *needs*. If a survey concludes that *spelling* is why learners use a dictionary most

---

web site, however, still looks like a traditional dictionary (http://www. linguistics. berkeley .edu/hyperdico/).

frequently, this does not mean dictionaries should specialise in spelling, nor that it is the learners' most urgent need. Questionnaires and, for that matter, tests always work within the realm of what is possible. Needs exceed this realm and point to what is not yet existent. Paradoxically, in the history of mankind solutions often have preceded problems and the software industry in the last few years –as the invention of printing– has demonstrated this. In the case of questionnaires and tests, one investigates what is apparent. When not predictable, the results are often doubtful. Investigating the link between needs and habits is useful only in order to find out how learners at present tackle the problems which current dictionaries are expected to solve. They reveal what the needs are only to the extent that they are soluble by already existing means.

If we succeed in using computers adequately for lexicographic purposes this will render obsolete a large number of research publications on grammar codes, phonetic alphabet, number of examples, need for synonyms, need for encyclopedic information, place where learners look up multi-word items, classification of fixed expressions, and others. If this technology were fed with qualitative research results, a blueprint could be drawn up which could give rise to a major revolution in the world of lexicography.

In the following chapter, I give an idea of how qualitative research could be carried out. It is rather labour intensive, but it gives in my opinion a better overview of the issues under investigation. I chose *examples* as the topic of this research since I think this is one of the most important and least studied elements of foreign language lexicography.

CHAPTER 4 EXAMPLES: THE CORE OF FOREIGN LANGUAGE LEXICOGRAPHY

4.1. General Considerations

4.1.1. Overview of the Research on Examples

Before COBUILD started using its corpus to retrieve examples in 1987, little had been published on the subject. Up to then, the problem had been virtually ignored and it was presumably considered minor when compared to topics such as the definition of meaning, or translational equivalents. Choosing an example to illustrate the meaning of an item, or one of its meanings, is probably not as difficult as choosing one which has to show how an item can be used productively. And since until recently dictionaries were primarily used to understand texts, not to produce them, not too much effort had been spent on this question. Traditionally, examples were made up by the lexicographer and their length was kept to a minimum –as often in native speaker's dictionaries–, or linked very closely, and unnaturally, to semantically related items –as often in learner's dictionaries. The first option would give something like 'The sale drew large crowds.' (*draw* in the *Random House Webster's*), the second 'She hammered the nail into the wall'. (*hammer* in the *Oxford Wordpower*) In both cases, the user is offered an example of how the item *can* be used, without much information about how it *should* be used. This kind of concern became important only from the moment when dictionaries started helping learners to *produce* in a foreign language. It was COBUILD which first gave *encoding* learners authentic examples which they thought were more adapted to encoding needs. Whether this option was the most sensible one will be discussed later in

this chapter, but COBUILD deserves the credit for having put the entire examples question on the map.

Since a number of articles published on the subject have been discussed in Chapter 3, I will limit this overview to a few highlights. In 1992, Ahmad, Fulford and Rogers published an article titled 'The elaboration of special language terms: the role of contextual examples, representative samples and normative requirements'. Apart from the results of a questionnaire conducted among lexicographers on the perceived function of examples which will be discussed below, the authors also mention a series of criteria they claim should inform the choice of examples. Although the area of Ahmad, Fulford and Rogers is specifically terminology, these criteria are worth quoting.

> a. Avoid examples containing pronouns referring outside the example
> b. Avoid examples containing more than two other technical terms
> c. Avoid examples with a complex structure
> d. Avoid examples which are long
> e. Favour examples which are a complete sentence
> f. Favour examples where the term appears early on rather than late (Ahmad et al., 1992:146)

Also mentioned in this article are the criteria which lexicographers themselves felt to be important for the choice of examples: 'typicality; naturalness; length; usefulness of syntactic information provided by the example; semantic complexity of the example'. (142) Concrete examples in dictionaries indicate that these criteria are put into practice by lexicographers in a more sophisticated way than the straightforwardness of the theory suggests. The fact that in the lexicographers' answers no reference was made to the distinction *encoding/decoding* indicates that at that time the matter was not yet top of the agenda.

In 1993, Batia Laufer conducted an experiment at the University of Haifa on 'The effect of dictionary definitions and examples on the use and comprehension of new L2 words', articles which I have commented on in Chapter 3. As I mentioned, Laufer found that for *encoding* purposes, combined entries were better than a definition alone and a definition alone was better than an example alone. For *decoding* purposes, combined entries were also better than a definition or an example alone, but when used separately, definitions and examples were equally efficient. (140)

Since access to both a definition and an example increases the net amount of information it should not be surprising that the results of learners who had access to combined entries were better than of those who only had access to one of these elements, for *decoding* as well as for *encoding*. That in the case of *encoding* a definition alone would be better than an example alone contradicts the general belief, and one's own experience. Probably the subjects were still unsure as to the meaning of the item. Less astonishing but still surprising is the result that in the case of *decoding,* an example would be just as efficient as a definition. In the chapter on methodology I have mentioned my reservations regarding this kind of research. The results are, however, worth considering.

An internal COBUILD paper (Harvey and Yuill, 1992) contradicts Laufer's conclusions to a certain extent. This research was based on an 'introspective' questionnaire given to 211 informants after a production test. Since the research was commissioned by COBUILD, a great deal of attention was given to examples. Harvey and Yuill concluded that learners preferred to turn to the examples rather than to the definition in order to work out the meaning of a word (Harvey K. and Yuill D.,

1992:20). This leaves us wondering if the examples are particularly suited to explain meaning, or if the definitions are particularly bad. Here too, defects inherent in the research method are acknowledged by the research team. Indeed, some to the subjects reported they found a particular piece of information in places where this information is not being provided (p. 10). The research by Harvey and Yuill also defends the use of authentic examples, a theory questioned by Laufer. I will myself come back to this issue in the section dealing with authentic vs. made-up examples (4.4.).

In Hong-Kong, Amy Chi and Stella Yeung (forthcoming) carried out a survey in which they reached the same conclusions as Harvey and Yuill. Learners said they preferred to use the examples in an entry to understand the meaning of a word.

> In the case of using dictionaries for examples, 64% of the students said they had been taught in class how to use example(s) of a given word for reference in their own writing (Q.18) when in fact, examples came fourth in the ranking we set for Q.14. Moreover, when students were asked 'How do you usually decide which one best explains your words?' (Q.16) (appendix VIII), over half of the students said they would use the example(s) given to decide which explanation is most appropriate. (No page number, electronic version.)

Since the Hong-Kong survey consisted of a questionnaire, the results should be treated with some caution since they reflect only what the subjects think they do and not necessarily what they really do. However, these findings coincide with one's own intuition and experience. Since the mind is lazier than the body, one can expect a learner to prefer a concrete example to a definition which is always a micro exercise in philosophy.

In the review of criticism (Ch. 3), I discussed Hillary Nesi's research on examples in the *Longman Dictionary of Contemporary English* (LDOCE). The results of this experiment show how little examples have been examined and how surprising the

results can be when they are. One might question the overall validity of the results of Nesi's experiment, and she herself points out some of its possible flaws. However, I tend to agree with her when she claims that:

> The number of example sentences provided by LDOCE for each target word ranged from none (for COMPUTE) to five (for VERSION), but, surprisingly, no policy was apparent to account for this variation. (No page number, electronic version.)

In what follows, I will suggest that this apparent lack of policy is typical not only of the LDOCE.

4.1.2. The Nature of Examples

A number of rather different types of utterances are traditionally considered to be appropriate dictionary examples and a definition of their various types has been proposed by Hausmann (1979). He distinguishes between *free associations*, *collocations*, and *fixed expressions*. In what follows I will reformulate this distinction in terms of John Sinclair's theory (1991). I propose that examples be subdivided into three types: those that abide by the *open-choice principle*, *idioms*, and *fixed expressions*[7].

The open-choice principle. As regards the *open-choice principle* kind, an example such as *the* road *to Teruel* shows the target word *road* used in its *core meaning* ('the most frequent independent sense' (Sinclair 1991:113)). On each side of *road* a number of lexical items can be filled in: *the road to Teruel goes that way, it would be better to find the road to Teruel, there must be a road to Teruel*, etc. There are, of course, limitations on what kind of words can be slotted in on each side of the target word (*on, at, from, up, into, down, off, along, across*, to name only the prepositions), but it is still

[7] The terminology of this chapter is explained in the Glossary at the beginning of this thesis.

a large list. This is why this example can be considered as abiding by the *open-choice principle*.

The idiom principle. Even as a cautious lexicographer who s/he probably is, the person who was in charge of *Collins Spanish* entry for *road* took a few liberties with *to get a show on the road.* According to the COBUILD corpus, one can 'get this show on the road' or 'get the show on the road' or 'get someone's show on the road', but not 'get a show on the road' as *Collins Spanish* implicitly claims. Idioms are not as inflexible as fixed expressions, but their possibilities of variation are restricted. In the words of John Sinclair:

> The individual words which constitute idioms are not reliably meaningful in themselves, because the whole idiom is required to produce the meaning. Idioms overlap with collocations, because they both involve the selection of two or more words. At present, the line between them is not clear. In principle, we call co-occurrences idioms if we interpret the co-occurrence as giving a single unit of meaning. If we interpret the occurrence as the selection of two related words, each of which keeps some meaning of its own, we call it a collocation. (Website 'Corpus linguistics', http://clg1.bham.ac.uk/glossary.html)

As Sinclair convincingly states, people speak and write mainly by combining chunks of language, not words. In an idiom, a word acquires a new meaning and does not mean the same anymore as when it is used independently. So the meaning which people intuitively give to a word is a criterion for distinguishing parts of text created on the basis of the *open-choice principle* and parts created on the *idiom-principle*. When a word in a text cannot be taken in its core-meaning, then it is part of an idiom[8].

[8] The Brazilian translator Boris Schnaiderman told me one day that Nabokov was considered a traditional writer in Russia, contrary to the way in which people thought of him in the West. Likewise, when Mallarmé said: '*Mais, Degas, ce n'est point avec des idées que l'on fait des vers, c'est avec des mots*', he unknowingly referred to the *idiom principle.* Verbal artists choose words, rarely chunks of words, while ordinary people do not usually speak in words, but in chunks. When writing in English, Nabokov was thinking in words, due to the fact that his command of English was not perfect, and this was the best guarantee of originality.

Fixed expressions. Fixed expressions are a more familiar concept. They are traditionally classified among the examples, even if they allow no variation (unless the speaker wants to make a pun). In the case of 'one for the road', there is less possibility for variation than in the case of idioms. *Fixed expressions* do not allow any creativity on the part of the speaker. Furthermore, their meaning cannot be deduced from the sum of its elements. In 'one for the road', *one* cannot be normally replaced by *two* or any other number, unless jokingly, nor can one speak of *one 'on' the road* or *one 'off' the road.* The essential difference between *idioms* and *fixed expressions* may be this: the possibility or not of altering the order of the elements or of intercalating another element in between.

Conclusion. The different kinds of examples can be put on a cline with on one side *the road to Teruel*, on the other the fixed expression *one for the road* and in between the idiom *get the show on the road.* The more the item is of the fixed expression type, the less its overall meaning is deducible from the sum of the meaning of its separate elements. Fixed expressions and idioms are, to a variable extent, chosen by the speaker as wholes, chunks which act like large single words.

I call examples *in the strict sense* of the word only those utterances which do not show the target word in a fixed environment. In the case of *road* this is exemplified by utterances of the type 'the main road is very busy at times' or 'you must take compulsory basic training before you are allowed to ride on the road' (instances taken from *The Bank of English*). They show the application of the rule by which the word 'road' is governed in particular contexts.

Examples such as *one for the road*, on the other hand, are entities in their own right and they can be regarded as a single choice to which firstly apply the rules for 'one for the road', and only secondarily the ones for 'road'. The kind of example users are interested in depends on the activity they are engaged in. In the case of *decoding*, *fixed expressions* are the examples that are really useful and, to a lesser degree, *idioms*. When *encoding*, on the other hand, examples which show the constraints on the word in an open-choice environment will be more useful. (I will come back to this in section 5.1.2.)

4.1.3. The Function of Examples

The fact that examples, as Sinclair (see below for more details, 1987:XV) claims, often repeat what has been said in the definition might not be entirely mistaken. Surveys (Chi, forthcoming) (Harvey K. and Yuill D., 1992:20) have shown that in learner's dictionaries, users prefer to turn to the examples in order to work out the meaning of a hard word. This might be an indication that definitions are not working properly, not that examples should be used that way. However, the fleshing out of definitions might be one of the functions of examples. Drysdale (1987:215) summed up these functions as follows:

1. To supplement the information in a definition.
2. To show the entry word in context.
3. To distinguish one meaning from another.
4. To illustrate grammatical patterns.
5. To show other typical collocations.
6. To indicate appropriate registers or stylistic levels.

It is true that these criteria are rather vague. One might fail to see how to 'show the entry word in context' would not be some sort of superordinate for all the other

features. Similarly, there does not seem to be much of a difference between 'to supplement the information in a definition' and 'to distinguish one meaning from another'. In 1992, however, these criteria still agreed with those of the lexicographers themselves. In the previously mentioned survey conducted by Ahmad, K. et al. (1992), lexicographers declared that examples were 'to enable the user to distinguish between senses; to provide usage evidence; to show typical collocations; to amplify and clarify definitions; to show typical use of words'. (1992:142)

4.1.4. Examples for *Decoding* and Examples for *Encoding*

For the problem at hand –the choice of examples,– this is of great importance. A *decoding* learner is interested in the meaning of a word or lexical item and an example which illustrates the meaning might not show clearly its syntactic features or relevant collocates, which is the point of interest of the *encoding* learner. However, since this call was not heeded until recently, examples were chosen to satisfy native and non-native speakers involved in both *encoding* and *decoding*, with a clear emphasis on the latter. Even in the *Longman Language Activator*, no special policy for choosing examples for *encoding* is perceptible. It is no surprise that some forty years after Hornby's invention of the learner's dictionary (1948), some people were still complaining that the made-up examples in learner's dictionaries were part of the definitions and that the main purpose of examples was to 'clarify the explanations' (Sinclair in COBUILD I, 1987:XV). Surely the purpose was still *decoding*. However, by the time COBUILD I was published in 1987, millions of non-native speakers were using

English productively, and *encoding* dictionaries with examples to serve this purpose were sorely needed.

Conclusion

Research done so far tackles the problem from different angles. Some of it deals with the way actual dictionary examples work. Some of it is about the function of examples. Other researchers still have started from an analysis of what exists and finished with proposals. What I shall do is analyse the specific case of one word (*road*), examine the underlying logic of the examples in the main learner's dictionaries, and formulate a few suggestions for improvement. Finally, I shall start from what I and other researchers think are the main issues on which examples have to give information : Collocations and Syntax.

4.2. Examples in Learner's Dictionaries. The Problem of Collocation and Syntax.

In order to see what logic underlies the choice of examples in learner's dictionaries, I examined the entries for *road* in COBUILD I and II, *Longman*, *Oxford Advanced* and *Oxford Wordpower*. (The complete list of examples can be found in Appendix 1.) I chose the word *road* because of the wealth of examples I found under this entry. This made a comparison between the different dictionaries less haphazard.

Although there are differences between the dictionaries I examined, they all follow roughly the same pattern. COBUILD has more and longer examples. COBUILD and *Longman* tend to give full sentences, whereas the OALD gives only short strings of words. *Oxford Wordpower* has both long and short examples. The OALD has a higher

percentage of sayings and fixed expressions, while *Longman*, *Oxford Wordpower* and COBUILD concentrate on non-idiomatic uses.

4.2.1. Classification

In all five cases, the classifying principle of the examples is *meaning*, since the entries are subdivided according to the different senses which *road* can have. Within this classification there is precedence of the *literal* sense (i.e. piece of ground betwe*en two places*) over the *figurative* one (i.e. *way or course*). Fixed expressions –not examples in the strict sense of the word– come at the end of the entry. The first senses in the entry have the greatest number of examples. Indeed in all dictionaries, literal senses have more examples than figurative ones.

Within this general grouping there are further and less constant minor classifications. For instance, in COBUILD I's third subdivision and *Longman*'s first, the examples show a list of prepositions commonly used with *road* and one has the impression that, within a classification based on meaning, a syntactic organisation has suddenly taken over. This suggests there should be a link between this sense of *road* and the prepositions, which is not true. In other places the information contained in the examples highlights some sort of syntactic particularity ('If you get onto the ring *road* you'll avoid the town centre', *Oxford Wordpower*), a collocate ('a busy *road*', *Longman*), or a multi-word item ('he was killed in a *road* accident', *Longman*). It is impossible to avoid transmitting this kind of information while conveying meaning, but there are differences in emphasis. Any example transmits information on meaning, but not every example transmits adequate syntactical or collocational information.

I shall now look at the *road* examples from the point of view of their use first for a *decoding*, and subsequently for an *encoding* learner. I do not think personally that examples in a learner's dictionary are the most adequate decoding aids, but since learner's dictionaries aim at helping learners with that as well, and some learners say they use examples to work out the meaning, I will evaluate these dictionaries on this topic as well.

4.2.2. Decoding

The classification of examples I described is one to which we have become accustomed and which seems self-evident, but I am not sure if this should be so. When *decoding* – reading– one is concerned with the meaning of a lexical item. This is an easy task for a dictionary. The most logical thing to do is to look up the hard word in a bilingual dictionary and let the context correct any inaccuracies. However, this standpoint contradicts the presuppositions of those in favour of the use of learner's dictionaries against bilingual dictionaries. I shall not go into the validity of the arguments here; I will only point out that as a result a number of learners became accustomed to monolingual dictionaries. In addition, the research I mentioned above points out that learners tend to use the examples and not the definition to work out the meaning of a difficult word. To help learners work out the meaning of a word, examples should concentrate on clarifying it. As it is now, examples do this partially and almost inevitably because the word means what it means. They do not do this in the most effective way because a number of examples are supposed to show other features of the lexical item as well: collocates, prepositions, etc. These features inevitably complicate

the example since they make it more complex. I will analyse some of the examples for *road* more closely.

The first COBUILD II example, *There was very little traffic on the roads*, transmits the meaning of *road* efficiently since *traffic* situates the word in a precise co-text and *on* limits the kind of words by which *roads* is followed. In this sentence few other words can replace *roads,* and since no syntactic intricacies render the understanding of the example difficult, provided that is the intention, it is possible to infer a more or less correct meaning from the example. There are, of course, other problems which any lexicographer will be hard-pressed to solve and which derive from the flaw in conception common to all learner's dictionaries: the fact that *traffic* is a less frequent word than *road.* This problem is difficult to solve since few words are easier than *road.* In this case, the intention was probably, first of all, to locate *road* in an unequivocal *traffic.* As an illustration of the meaning, this example is one of the best ones, but it also highlights the problem in its entirety. It is difficult to transmit the meaning, particularly of very frequent words, through an example. The best one can attempt is to not suggest a meaning which is false. Jain (1981) did some interesting research on this point, showing that dictionary examples do often suggest a wrong meaning.

The rest of the examples in this entry, all clear, help to circumscribe the word well. As a way of comparing, we can look at the COBUILD I examples: *There is an antique shop at the top of my* ***road*** and *the quiet Edgbaston* ***road*** *where he had lived for some thirty years.* In both cases, if *road* were replaced by another word the sentences

would still make sense. This does not mean that these examples are not useful for other purposes, but they are not for *decoding*.

In *Longman*, the example *a busy road* does not give any conclusive information about the meaning of the word and the intention was obviously to show a significant collocate. *At the end of the road* does give information about the meaning and *We live just down the road* fulfils both the conditions of conveying the meaning adequately by using more frequent words than *road*, and transmitting a useful collocate *(live just down)*. This example therefore ought to be placed first.

In the present state of affairs, in which examples play a role in the elucidation of the meaning of a word and in which this is also a declared intention of the dictionaries, to choose examples without this goal in mind can lead to the incorrect replacement of the target word by a lexical item with a, for learners, roughly similar syntactic behaviour, but a different or opposite meaning. The following COBUILD II examples for *play down* are an illustration of this: *Western diplomats have played down the significance of the reports. He plays down rumours that he aims to become a Labour MP. Both London and Dublin are playing the matter down.* In these sentences, *play down* can be replaced by *take interest in*, or *reject*, or some other verb. This means that while a learner would surely extract a meaning from these examples, it could very well be the wrong meaning.

I do not know of any research which shows that learners get more information from a dictionary definition than from a translation combined with the context in which the word was found. Personally, I think bilingual dictionaries would be more suited for

decoding than monolingual dictionaries, if they had the same coverage. Since this is not the case, learners dealing with a more specialised kind of vocabulary will necessarily have to use a monolingual dictionary. Moreover, given current practices, and with teachers advising their students to use monolingual dictionaries instead of bilingual ones, the meaning transmitted by examples continues to be important. Either learners have to be advised to switch to bilingual dictionaries again, or lexicographers should take this point into consideration and choose examples which transmit the meaning of a word unequivocally. It is clear, given the characteristics of language, that this is an extremely difficult task.

4.2.3. Encoding

An *encoding* learner needs much more information than a *decoding* learner and the information is of a different nature. When *encoding* learners look up a word they already know its meaning to some extent. They may want to confirm this meaning, but their main interest is *usage*, a combination of *syntax* and *collocation.* This kind of information can be transmitted directly or indirectly: *directly*, by means of grammar formulas and lists of collocates, and *indirectly* by means of examples. Examples are the way in which most learners expect to find this information in the dictionary. Particularly in the case of grammar, an example is more significant than the cryptic rules we find in current dictionaries. However, the classification of the examples is anything but adapted to queries about usage.

In the doubtful case that learners look up *road* for *decoding* purposes in a monolingual dictionary and understand the meaning of the word by reading the

definition 'way between places' (OALD) they will not bother to see their view confirmed by the examples *the road to Bristol/Bristol road* (OALD) or *There was very little traffic on the roads* (COBUILD). These examples are aimed at *encoding* users who would like to know if one can say, for instance, *Bristol road* as well as *the road to Bristol*. However, as we saw in the previous section, information of a type which is more sought by *encoding* learners than by *decoding* learners has been classified entirely according to *decoding* needs, making it nearly impossible for *encoding* users to retrieve it. This can be seen in the first subdivision of the *Longman* entry which deals with the main literal sense of *road*.

1. a *busy* ***road***
2. *at the end of* the ***road***
3. We live *just down* the ***road.***
4. It takes three hours *by* ***road.***
5. Take the *main* ***road*** out of town and turn left at the first light.
6. He was killed in a ***road*** *accident.*
7. Kids of that age have no ***road*** *sense.*
8. A ***road*** *safety* campaign.

Figure 1 Examples for road *in* Longman.

The entry has obviously been compiled in a very orderly fashion. The first example gives us a collocate that can help to understand the word's meaning, allowing for the fact that *busy* is a less frequent word than *road*. Examples 2, 3 and 4 give syntactic information, in this case on prepositions. Example 5 gives the learner a collocate (*main*), and 6, 7 and 8 give collocates which amount to multi-word items. Who benefits from these examples? The multi-word items are presumably intended for *encoder*s and not for *decoder*s, since there is no explanation of what they mean. It may be that the lexicographers thought that their meaning could be derived from the meaning

of the words which compound the multi-word item, but that would be a miscalculation. More specifically, *road sense* seems rather difficult to grasp.

On the other hand, I am not sure if *encoder*s will benefit from the syntactic information. At most *encoding* learners will see their assumptions confirmed. They can suspect that one can say 'just down the road' and find the phrase in the dictionary, but this will never be more than a happy coincidence, since the list is anything but comprehensive.

What is the purpose of the multi-word items? It can only be to show that they exist. If an *encoder* has to translate *uma campanha pela segurança no trânsito*, s/he will suspect that *a road safety campaign* is the translation s/he is looking for. Unfortunately this is not a comprehensive list and one can ask why this particular one has been chosen. It could be frequency, but it remains difficult to deduce a policy from this selection.

The following are the examples for the first literal sense in *Oxford Wordpower*:

1. Is this the right road to Beckley?
2. Take the London road and turn right at the first roundabout.
3. Turn left off the main (= big, important) road.
4. major/minor roads
5. If you get onto the ring road you'll avoid the town centre.
6. road signs
7. a road junction.

Figure 2 Examples for the first literal sense of road *in* Oxford Wordpower.

These examples all illustrate, inevitably, the meaning of the word *road*, but do they do anything else? The first three examples all refer to the same activity: showing the way. They are probably useful, although one would expect them to be part of a dialogue in a course book, rather than examples in a dictionary. The examples show a

few significant basic collocates: 'right', 'take', 'main'. The latter is explained between brackets –(=big, important)– whereas the learner is left without any help as regards the phrase 'turn left off'. Example 4 gives collocational information useful for *encoding* in the not unlikely case that a dictionary user recognises *major/minor roads* as the item s/he was looking for. Examples 5, 6 and 7 seem to be intended to show a few multi-word items with *road* in them. One would tend to classify this as *decoding* information, but as before there is no indication as to what these multi-word items mean.

If we turn to the authentic language examples of COBUILD II, we see that the policy for the choice of examples is not fundamentally different. However, if we believe the COBUILD advertisement, the choice of examples reflects the corpus frequency. The following is the first section of the entry, illustrating the literal meaning of *road*:

1. There was very little traffic *on* the ***roads***.
2. We just *go straight up* the Bristol ***Road***.
3. He was *coming down* the ***road*** the same time as the girl was turning into the lane.
4. Buses carry 30 per cent of those *travelling by* ***road***.
5. You mustn't lay all the blame for ***road*** *accidents* on young people.

Figure 3 Examples illustrating the literal meaning of road *in COBUILD II.*

Five examples were chosen. They show the most common preposition for *road* 'on', and three significant collocates: 'go straight up', 'come down' and 'travel by'.

In the particular case of *encoding* learners, to have the possibility to recognise the information which they are looking for is fundamental. Unfortunately, this information has been scattered all over the entry in such a way that only patience and faith can help them find it. In the COBUILD entry for *road* this task has been made easier because there is only one section dedicated to the literal use of the word, and the examples are all

grouped. To find the multi-word item '*road* accident' at this same location is strange, yet the reasoning must have been that an accident takes place on a 'literal *road*'.

Conclusion

As things stand now, it would be better for learner's dictionaries to focus on helping learners with *encoding* and not with *decoding*. Strictly speaking, *decoding* learners do not need examples. A good bilingual dictionary solves their problems. If currently they do consult examples for information on meaning, this might indicate the doubtful usefulness of definitions. A definition is necessarily more abstract than an example and it has to be fleshed out with nouns and verbs to become understandable.

The consequence of a classification which attempts to fulfil the needs of both *decoding* and *encoding* audiences is that neither of them will easily find what they are looking for. *Decoding* learners do not always find examples which clarify the meaning, and the sequence of the examples does not take into consideration the needs of *encoding* learners. These do not always find a sufficient amount of collocates or syntactic information and have to wade through information unsuited to their needs.

4.2.4. Requirements for examples

In 1981, Béjoint claimed: 'the best dictionary of *encoding* is one that provides the most detailed guidance on syntax and collocates'. (Béjoint: 1981:210) In what follows I will analyse to what extent four learner's dictionaries, the *Cambridge International Dictionary of English*, COBUILD I and II, and the *Longman Dictionary of Contemporary English,* deal with this issue.

4.2.4.1. Syntax

Strictly speaking, the difficulty involved in using a lexical item in a foreign language, and consequently the number of examples needed, is proportional to the difference which exists between this word's behaviour and that of its equivalent(s) in the learners' L1. Spanish *gustar* is a case in point. English-speaking students may experience difficulties using *gustar*, a pronominal verb which makes the item(s) one *likes* become the subject of the verb (I like books / *me gustan los libros*). These same students have few problems using its antonym *detestar,* because in this case the syntactic construction is the same as in English (I hate books / *detesto los libros*). Brazilian students have the same difficulties with *gustar*, for the same reason as the English, but in turn they find it easy to distinguish between *ser* and *estar,* which is difficult for the English.

It is not the task of dictionaries to teach grammar points in a comprehensive way, but the dictionary is the place where learners look for information on syntactic constraints. Grammar rules in dictionaries should therefore be both specific and general. They should be general in the sense that there are difficulties common to all nouns, verbs, adverbs, etc.; they should be specific because every word has its own syntactic particularities when compared to its equivalent(s) in another language. In order to choose truly useful examples, dictionaries should ideally pay attention to the specific problems arising from the contrasting of two languages: the one it is supposed to elucidate and the language of the audience. Speakers of different languages have each different problems and need different examples. In a Spanish-English dictionary, English speakers will need a huge number of examples showing the use of *ser* and *estar*, of *gustar* and so on. In terms of learner's dictionaries, only *Cambridge* makes a modest

use of contrastive information for a number of words which their research pointed out to be false friends.

Since dictionaries do not usually give discursive information about syntax, there are two ways of conveying this: codes and examples. Although the first method, as used in COBUILD's 'extra column', is more direct and saves space, it might not be as effective as to put in simply more examples. In the case of COBUILD, the information in the 'extra column' repeats the information given in the examples under the form of codes, but this is not very clear nor attractive to the common user.

Generally speaking, the information a learner needs may be more readily supplied by a series of examples rather than by an abstract and abbreviated rule. In addition, a syntactic construction is rarely neutral with regard to collocations. If the number of these collocates is small, it should be easy to list them all, and the list of collocates would also give the list of possible syntactic constructions. If they were too numerous, an appropriate grouping would reveal the regularities. In what follows I analyse the case of *take a stand* and *indulge* from the point of view of syntactic information as conveyed through examples in a few learner's dictionaries.

Take a stand

To take a stand is a frequent expression which is difficult to learn, mainly because it can be used on its own, with different prepositions (*on*, *against*, *for*), or followed by a number of adverbial phrases. Most dictionaries do not give any indication on these uses, and when they do, as in the COBUILD Dictionary on CD-ROM based on COBUILD I, they mention formulas such as PHR : VB. – INFLECTS which must seem rather enigmatic

to most learners. A look at the COBUILD Collocations CD yields a wealth of examples displaying a variety of syntactic structures likely to be used with *take a stand.* They even show secondary syntactic information, such as the fact that the *infinitive* form is by far the most common one; that *against* –and not *on*, as the other dictionaries suggest– is the most frequent preposition. The following table shows examples retrieved from the COBUILD Collocations CD.

it's time to *take* a *stand*
on the movement to *take* a resolute *stand* against Iraq
felt it was important to *take* a *stand* against Souter
Law Lords yesterday failed to *take* a *stand* against that moral decline
by failing to *take* a strong *stand* against the students
to *take* such a *stand* attacks the credibility of the FASB
Joanna decided to *take* a *stand* herself
believe that he will *take* a tougher *stand* in negotiations with the Chinese
I wouldn't expect him to *take* a *stand* like that against Jews or blacks
government finally decided to *take* a *stand* not because it believed
to *take* a more principled *stand*
the group insisted he *take* a *stand* to overturn Amendment 2
doesn't know whether to *take* a *stand*

Figure 4 Examples for take a stand *retrieved from the COBUILD Collocations CD*

The examples above were not carefully chosen since they were retrieved from a corpus and therefore listed in a relatively random fashion. Still, the information they provide is much richer and much more accessible to a learner than an abstract rule or, worse, its abbreviated form. In an ideal situation, a learner would be able to retrieve, and sort at his/her leisure, a large number of examples which would make the use of grammar codes superfluous. This would also guarantee that information which would go unnoticed to the eye of the lexicographer or would not be capable of being grasped in codes, would reach the learner. This is, for that matter, one of the arguments of those who defend authentic examples, the fact that it is possible to transmit information of which not even the lexicographer is aware.

Indulge

Indulge illustrates another aspect of the problem of conveying syntactic information. It has at least two possible constructions, and cannot be translated by one single verb with the same syntactic particularities as in English, in any of the languages with which I am familiar. Furthermore, from the moment one is sure *indulge* is the word one wants, one has to start the struggle to get the syntax right. One can be fairly sure that *indulge* is the verb which best translates *s'adonner à*, taking into account the possible pragmatic implications such as the pejorative aspect of the word, irony and slight formality. Still one can have doubts about how to use this verb with such a difficult syntactic pattern. Is it *indulge in* or *indulge with* or *by*? Is it *indulge oneself in* or *indulge oneself* without any preposition? To what extent do examples help the learner with this dilemma?

COBUILD. These are the COBUILD II examples for *indulge*:

Only rarely will she *indulge* in a glass of wine. He returned to Britain so that he could *indulge* his passion for football. You can *indulge* yourself without spending a fortune. He did not agree with *indulging* children.

Figure 5 COBUILD II examples for indulge.

COBUILD mentions two senses for *indulge*[9]. Two of the examples illustrate two possible constructions corresponding to the meaning 'allow yourself to have', and two illustrate the sense 'spoil'. The constructions which convey the meaning 'allow yourself to have' are: *indulge* in + a noun; *indulge* + a noun. Two examples is probably what a native speaker finds more than sufficient, but for a learner, unless there is an equivalent in the learner's first language with the same syntactic behaviour, three examples are not

[9] '1. If you **indulge** in something or if you **indulge** yourself, you allow yourself to have or do something that you know you will enjoy.

2. If you **indulge** someone, you let them have or do what they want, even if this is not good for them.'

enough. From the entries one can tell there was no awareness that *indulge* is a difficult verb, which is the reason why it is treated in exactly the same way as most other verbs. However, even without trying to translate it, the lexicographers should have been aware of the difficulty of this verb for a learner because of the number of its possible constructions. In the case of the COBUILD entry, the problem is that a general rule of thumb –one example per construction– was applied to a particular case, instead of tailoring it to each specific verb. The chosen examples are excellent, but way too few to allow a learner to use the word productively.

Longman. *Longman*'s examples are less satisfactory than COBUILD's:

Most of us were too busy to *indulge* in heavy lunchtime drinking. Eva had never been one to *indulge* in self pity. I haven't had strawberries and cream for a long time, so I'm really going to *indulge* myself. Ray has enough money to *indulge* his taste for expensive wines. His mother pampered and spoiled him, *indulging* his every whim.

Figure 6 Longman*'s examples* for *indulge.*

Firstly, to use the verb four times out of five in the infinitive form gives the learner the erroneous impression that 'to + *indulge*' is the verb's normal construction. There is an indication in the examples that the construction with *in* can apply to concrete as well as abstract activities (*drinking; self-pity*), which is indeed a question a learner is likely to ask. However, one single example is not sufficient to retrieve this information correctly; nothing in the examples will prevent a learner from saying **Most of us were too busy to indulge ourselves in heavy lunchtime drinking* if this information is not given in some other way. Here and in other dictionaries, the compilers were either unaware of the difficulty of the item or assumed that the learners had sufficient 'dictionary skills' to retrieve the information by themselves. This suggests once again that one of the drawbacks of existing dictionaries is that they are in most cases compiled

exclusively by native speakers who are not always aware of the difficulty of some of the features of their language.

Cambridge. For *indulge*, *Cambridge* gives the following examples:

The soccer fans *indulged* their patriotism, waving flags and singing songs. With his friend's family he was able to *indulge* his **passion** for the outdoors, especially skiing. I love champagne but it's not often I can *indulge* my**self**. The children indulged me **with** breakfast in bed. This was a deliberate decision by the company to *indulge* **in** a little nostalgia. She was furious with her boss and *indulged* **in** rapturous **fantasies of** revenge.

Figure 7 Cambridge examples for *indulge*.

Cambridge gives a few more examples than the other dictionaries: two with the noun-construction, two with *in*, one pronominal and one with a pronoun. This last one, however, features *with* in bold, giving the learner the impression that this is a typical construction at the same level as *indulge in*. Furthermore, 'indulge their patriotism' is a slightly ironic use not likely to be grasped by the average learner. In addition, the 'company' indulging in 'a little nostalgia' is not what one would call a 'typical use', and 'rapturous fantasies' may be a little farfetched for a second language learner.

Conclusion

As a rule, dictionaries lack examples which highlight syntactic constraints. Instead, and perhaps only to save space, they use codes which are supposed to indicate how words are used. I doubt if these codes, perfectly justifiable in a paper dictionary, should be maintained in electronic dictionary. I can see no reason for it. Moreover, the above analysis reveals a deficient identification of what the difficult features of the different lexical items are. One has the impression that a uniform guideline for the number and type of examples is applied regardless of the difficulty of a specific item. At least a few guidelines should be taken into account. For instance, a learner needs fewer examples to

start using a *noun* than a *verb*. A flexible examples policy should be applied, preferably adapting the number and kind of examples to the specific word.

Finally, if learner's dictionaries are to help learners with *encoding*, the selection and sequence of examples should obey a logic of possible syntactic constructions and not of senses, even if syntax and meaning are inevitably intertwined, as research by Gill Francis and her team (1996) has demonstrated. It is true that *encoders* want to transmit a message and that this message *is* meaning. However, there is a stage in the *encoding* process when dictionary users already know what word or lexical item fits the encoding of this meaning, and the problem becomes purely formal.

In these cases, I think a classification according to syntactic patterns is more useful. Frequently advanced learners are fairly sure of what lexical item they should use in their situation because their passive knowledge of the language is much greater than their active competence. It is my experience that very often the problem is purely formal. A classification according to these syntactic patterns may therefore be more appropriate. Since the dictionary of the future will probably be electronic, users will at any rate have the option of classifying the examples according to either meaning or syntax.

4.2.4.2. Collocates.

Together with information on syntax, collocations may be the main reason why *encoder*s use dictionaries. In this section I will show the results of some research I did on the subject in both monolingual and bilingual dictionaries.

It is surprising how little attention the issue receives. Dictionaries give quite a bit

of information on exceptional uses of lexical items but very little on common usage. Fixed expressions are generally well represented, whereas little attention is given to everyday use, as if exceptional use were more difficult. For example, the OALD lists 14 typical expressions with *cat* (*it rains cats and dogs*, *let the cat out of the bag*, etc.), but gives no indication on collocates such as *feed the cat*, *let the cat out/in*, *stroke the cat*, etc. The reason is the same as before: learner's dictionaries were fundamentally designed to serve *decoding*, not *encoding*, even if their introductions say the opposite. Fixed expressions are useful for *decoders*, much less for *encoders*. Besides, it seems to be easy for native speakers to think of fixed expressions when asked for them. Normal uses are so much a part of everyday life that they escape one's awareness and can only be retrieved with the help of a corpus.

Face, fact, failed, fail, and *fade*

In some research I did on collocates, I compared the presence of collocates in COBUILD II and the *Cambridge International Dictionary of English* for the randomly chosen items *face*, *fact*, *failed*, *fail*, and *fade* (full details in Appendix 2). Although COBUILD itself produced the Collocations CD from where I took my information, it scored only slightly better than *Cambridge*. Of the 66 COBUILD examples for *face*, only 4 contain a collocate; in *Cambridge*'s 12 examples there are no collocates at all. One wonders why COBUILD, which signalled its awareness of the importance of collocations by producing a Collocations CD and whose theoretician John Sinclair wrote extensively on the subject, dedicated so little attention to this feature. In one case *–fact–*, *Cambridge* even scored relatively better than COBUILD, even if *Cambridge* had overall many fewer examples. This makes us suspect that the hits were not the consequence of a policy but

of the lexicographer's intuition. In the case of *fail*, *Cambridge* manages to show nine significant collocates in nine examples. In the overall evaluation COBUILD scores better only because it features a greater number of examples.

Criteria

In another piece of research I looked at the word *criteria* with which I myself had problems. According to the COBUILD Collocations CD, the most common collocate for *criteria* is *meet*. Of *Collins German, Collins Portuguese, Collins Russian, Collins Spanish, Collins-Robert, Collins-Sansoni, Hazon-Garzanti, New Proceed, Oxford Spanish, Oxford-Duden, Oxford-Hachette, Taishukan,* all major bilingual dictionaries, along with *Collins* (monolingual), *Oxford Wordpower, Longman, Cambridge* and COBUILD I and II, only the last one and *Oxford-Hachette* had an example with *meet*. These same dictionaries did not have any examples with other high-frequency collocates such as *set* or *apply*. This is once again surprising since it seems such an obviously useful piece of information, and easy to provide. There are only a limited number of actions to which *criteria* can be subjected: *set* them, *meet* them, *use* or *apply* them or *fail* to do these things. If dictionaries were really designed with the *encoding* user in mind, it would have been relatively easy to include information about how these actions are lexicalised.

Proposal

According to the COBUILD Collocations CD-ROM, the most frequent verbs that collocate with *proposal* are: *reject, put forward, make, support, accept* and *consider*. In Appendix

3, I list the examples for *proposal* in five monolingual learner's dictionaries and nine bilingual ones.

OALD. The second edition of the OALD (1963) paid little attention to collocational information and concentrated on the preposition required for *proposal.* The new encyclopedic edition (1992) shows no systematic change in this policy. There are more examples, one with *put forward,* but the focus remains on the use of the preposition. It is, however, remarkable that, from one edition to the other, there is a tendency to increase the length of examples. Again, this might indicate a shift in the function of examples from illustrating the meaning to showing how the item functions within a text.

COBUILD. Also between COBUILD I and II an evolution is perceptible. Although the example-like definition used *put forward,* COBUILD I examples did not contain a single typical collocate. COBUILD II, however, has three: *put forward, reject* and *accept,* which is quite a good result. An obstacle to more collocational information may have been the COBUILD 'authentic example' policy and the need, therefore, to find authentic examples that are at the same time 'typical' to concentrate the relevant information.

Cambridge. Cambridge successfully concentrates collocational information in *Congress has rejected the latest economic proposal* ***put forward*** *by the president* (the underlining is mine, bold is original). *Reject* is as much a collocate of *proposal* as *put forward* but only the latter was highlighted. This might indicate that the presence of this collocate was coincidental rather than the result of a policy. Unfortunately, there is no indication of the fact that one can '*refuse*' *a marriage proposal* but not, for instance, *a peace proposal.* On the whole it seems difficult to discern a rationale behind the choice

of examples. There are a few collocates, a few prepositions, and a few typical constructions present, but none of these areas has been given extensive treatment.

Conclusion

There are indications that lexicographers are becoming increasingly aware of the need for dictionaries to become tools for *encoding* and not only for *decoding*. However, this switch is still not apparent in the section which is the main aid to *encoding*: the examples. My analysis suggests that there is some concern with syntax, but much less with collocates. There seems to be at present no genuine policy regarding these issues. The main reason why this is not the case is that dictionaries do not choose between *encoding* and *decoding*. It is only when *decoders* are unable to deduce the meaning of a lexical item from the definition that they will turn to the examples. Otherwise, examples should concentrate on giving information on syntax and collocation.

4.3. Examples in Bilingual Dictionaries. The Problem of Classification.

For foreign language learners, the bilingual dictionary is still the main reference tool, particularly in the case of *encoding*. Because of this, one expects these dictionaries to pay special attention to examples and to follow a particular policy in this respect. However, this is not always the case and the choice of examples seems often rather random. Once again the lack of distinction between *decoding* and *encoding* is the main impairing factor. In what follows I will illustrate what I think is the distinction between *encoding* and *decoding* examples, this time with a focus on bilingual dictionaries. In a subsequent section I will analyse the choice and ordering of examples in one specific case, the *Collins Spanish Dictionary*.

***Decoding* and *encoding*.** Current bilingual dictionaries attempt to meet the needs of both *decoding* and *encoding* learners at the same time. A Spanish-English dictionary is supposed to help Spanish speakers with the understanding and production of English while at the same time assisting English speakers with the understanding and production of Spanish. In fact the needs in terms of examples for *decoding* and *encoding* are very different.

When *decoding*, there is no need for examples in the strict sense of the word, as I defined this term in section 4.1.2. Even if a source word can be translated by several different words in the target language, the context in which the word was found will elucidate most doubts. For instance, in a sentence like 'Ni la solitude ni les épreuves n'avaient pu venir à bout de cet homme dont l'énergie n'avait d'égale que

l'insensibilité', (Gaucher R. 1965 *Les Terroristes*, Albin, Paris, p.32) any French-English dictionary will tell us that *épreuve* can mean *test*, *print*, and *ordeal*. Given the context, however small, there can be no doubt that the correct translation is 'ordeal'. An example in the dictionary would have given us even more certainty, but would probably be superfluous.

In the case of *encoding*, on the other hand, examples are of the utmost importance, since they help with the choice of the item and show how it is used. Particularly in the case of beginning learners, the first function is of great importance. Since the translation of a word depends on the context, learners will choose a translation possibility either guided by a label, or by comparing the source sentence to the example sentences. In the case of *encoding*, thus, the needs of learners for examples will be more for examples in the strict sense of the word. Only advanced learners will want to translate fixed expressions. I have monitored myself for two years in my look-ups in Japanese, of which I am a beginning learner, and in Italian for which I consider myself intermediate. In both cases I looked up high frequency words, with a large range of meanings and therefore more likely to be ruled by the open-choice principle. In English, on the other hand, a language in which I had to express more complex ideas and use more infrequent words, my look-ups concerned collocations and syntax as this information was provided by examples. My quest for examples, I may say, has been tireless.

The need for examples of *decoding* and *encoding* dictionary users are opposite. *Decoder*s are exclusively interested in examples in the broad sense of the word (see 4.1.2.), whereas *encoder*s need more common uses. Since traditional bilingual dictionaries cater for both kinds of audiences at the same time, and this in two different

languages, it is difficult to meet these needs in a satisfactory way. This is what I will show in the following analysis of how examples are handled in the entry for *road* in the excellent, but more traditional *Collins Spanish-English Dictionary* and, much more briefly, in the equally outstanding, but more modern *Oxford-Hachette French-English Dictionary.*

4.3.1. Analysis of the Examples for *Road* in Collins Spanish-English

The *Collins Spanish Dictionary* is a typical example of the more traditional kind of dictionary. Initially a monumental single person piece of work, the latest editions have benefited from the support of a more extended group of people. Since this dictionary is a rather typical example of its kind, and since more anecdotal evidence suggests the same is true of other dictionaries, the criticisms set out here are therefore applicable to most bilingual dictionaries. The dictionary does not mention the use of a corpus.

In what follows I look at the examples for the entry *road* and try to find out why they were selected, or made-up, to discover what is the rationale behind the examples policy in an average traditional bilingual dictionary. I look at what their sequencing criteria are and who their intended audience is.

In an ideal dictionary, the selection and ordering of examples should be guided by what functions these examples are supposed to perform. If the main purpose of the dictionary is to help with *decoding*, the fixed expressions should be listed first, or at least clearly separated from the rest. If the purpose is *encoding*, preference should be given to examples showing the word first in open-choice contexts, then in contexts where a collocation was required, and finally in fixed expressions. Multi-word items, if

not listed as separate entries, should close the entry.

The original hypothesis of this analysis was that, due to the fact that bilingual dictionaries had to meet the needs of two different audiences, with possibly two different aims, the lexicographers had no other choice than to make a random selection of examples without considering what these examples were supposed to be helpful for. In my view, this appeared to be true.

Secondarily, the analysis of the examples revealed that this was an untenable position and, in reality, there was a privileged audience. However, since for commercial reasons this kind of bias cannot be admitted, the needs of this privileged audience could not be met as well as if the lexicographers had had full play. A necessary heterogeneous product was the outcome.

Order of the Examples

Since it was immediately clear that there was no clear-cut classification in terms of common uses, idioms and fixed expressions, I tried a number of alternative classifications, starting with the one distinguishing *literal* and *figurative*. This classification is the one which has traditionally been used. However, here too, this criterion was not applied very openly.

Road is a very frequent word and its entry contains a large number of examples. It is used literally as well as figuratively and there is a Spanish equivalent *camino*, whose semantic range by and large covers that of *road.* This makes a comparison with the Spanish part of the dictionary fairer. *Collins Spanish-English* enumerates 38 examples

in the broad sense of the word. Below are the examples in the order in which they appear in the dictionary.

1. *roads* (naut.), *road* narrows, *road* up, the *road* to Teruel, at the 23rd kilometre on the Valencia *road*, the *road* to success, one for the *road*, across the *road*, she lives across the *road* from us, by *road*, my car is off the *road*, to be on the *road*, my car is on the *road* again, he's on the *road* to recovery, we're on the *road* to disaster, the dog was wandering on the *road*, to get (or take) a show on the *road*, to take to the *road*, to be on the right *road*, to get out of the *road*, our relationship has reached the end of the *road*, to hold the *road*, to take the *road*.

2. *road* accident, *road* construction, *road* haulage, *road* haulier, *road* hump, *road* junction, *road* racer, *road* safety, *road* sense, *road* tax, *road* test, *road* trial, *road* traffic, *road* transport, *road* vehicle.

Figure 8 List of the examples for road *in* Collins Spanish English Dictionary.

Although they are all listed under the same entry, these examples are very different in character. Group 2 shows multi-word items whose only reason for being there is their spelling. Somewhat surprisingly, *roads* as a nautical term heads the list. If we keep only the examples listed under number 1, minus *roads*, we are left with 22 examples. A first attempt at classification according to the 'open-choice vs. idiom principle' shows this was not the criterion used. The list starts with two fixed expressions (*road* narrows, *road* up), continues with three examples of open-choice (the *road* to Teruel, at the 23rd kilometre on the Valencia *road*), an idiom (the *road* to success), goes on with a fixed expression (one for the *road*), gives another example of open-choice (across the *road*, she lives across the *road* from us), a typical preposition (by *road*), etc.

Literal/figurative

Since the distinction literal/figurative is used by the dictionary itself and explicitly added to two examples ('to be on the right *road* and 'to get out of the *road*'), I checked if this was the rationale behind the selection of the examples. From my experience as a

language teacher I know this is one kind of subdivision into which learners expect the examples in a dictionary to be classified. Whereas the literal meaning of a word travels easily from one language to another if the same extra-linguistic reality exists in both languages, the way in which these same words are used to express concepts may well be very different. I listed the examples for *road* according to this criterion.

Literal	**Figurative**
the *road* to Teruel at the 23rd kilometre on the Valencia *road* across the *road* she lives across the *road* from us the dog was wandering on the *road* to be on the right *road* (in one sense) to be on the *road* (in one sense)	*road* narrows road up the *road* to success by road my car is off the *road* to be on the *road* my car is on the *road* again he's on the *road* to recovery we're on the *road* to disaster to get (or take) a show on the *road* to be on the right *road* (in one sense) (also fig) to get out of the *road* (fig) one for the *road* our relationship has reached the end of the *road* to take to the *road* to hold the *road* to take the *road* (to X para ir a X)

Figure 9 Classification of the examples according to their literal or figurative use (the indications found in the dictionary are in parentheses and italicised).

Even allowing for a few doubtful cases, the comparison with the actual order in the dictionary makes it clear that this distinction was not the guiding principle for the classification of the examples. Although not always satisfactory, an ordering in literal and figurative uses has the advantage of facilitating the access to the information. If applied, it works as a shortcut to the information. A more sophisticated classification might even take account of the fact that a number of figurative uses of words are common to different languages.

Literally Translatable/Not Literally Translatable

Since we are dealing with a bilingual dictionary, the compilers necessarily have a contrastive knowledge of the two languages, I thought they could have brought this to use in another kind of helpful classification of the examples: those that are literally translatable and those that are not.

The following listing undertakes this classification by comparing each example to its translation. I considered 'not literally translatable' those examples that would not be acceptable in Spanish in the case of a word for word translation, either because of syntactic intricacies ('the road to Teruel/la carretera *de* Teruel') or because they would be unintelligible.

Literally translatable	**Not literally translatable**
at the 23rd kilometre on the Valencia *road.*	*road* narrows.
the *road* to success.	*road* up.
across the *road.*	the *road* to Teruel.
she lives across the *road* from us.	one for the *road.*
by *road.*	my car is off the *road.*
to be on the *road.*	to be on the *road.*
my car is on the *road* again.	to get (or take) a show on the *road.*
he's on the *road* to recovery.	to take to the *road.*
we're on the *road* to disaster.	to take the *road.*
the dog was wandering on the *road.*	
to be on the right *road.*	
to get out of the *road.*	
our relationship has reached the end of the *road.*	
to hold the *road*	

Figure 10 Division of the examples according to their translatability.

This list cannot be recognised in the original ordering either. Since learners obviously do not know beforehand if what they are looking for can be translated literally or not, the literal/figurative distinction may be seen as a classification level prior to the literally/not-literal translatable one. It would not necessarily speed up the search, but it

would have a pedagogical advantage. It does not seem however, that the compilers of the *Collins Spanish* had this classification in mind.

The Audience

I finally looked at the examples for *road* from the point of view of the audience. I classified them in two categories: those that were useful for an English audience, and those that were useful for both English and Spanish audiences. The examples useful for the English would be those presenting some kind of difficulty for them. These same examples, however, would be immediately understood by Spanish speakers. Any Spanish speaker understanding *at, the, 23rd, kilometre, on*, and *road* understands *at the 23rd kilometre on the Valencia road.* This example is therefore intended for an English speaking audience. *Road up* and *one for the road*, on the other hand, are examples intended for a Spanish and an English audience alike. It is useful for Spanish speaking learners because even if they know what *road* and *up* mean this is not sufficient to understand *road up*. The item has to be translated as a whole. For English speakers, on the other hand, this example is useful as well since it shows them that *road up* cannot be translated by *camino arriba*. This gives the following list which I will discuss below.

English speakers (*encoding* into Spanish)	**Examples useful for both** (*encoding* into Spanish or *decoding* from English)
the *road* to Teruel *road* narrows at the 23rd kilometre on the Valencia *road* the *road* to success across the *road* she lives across the *road* from us by road my car is off the *road* to be on the *road* my car is on the *road* again he's on the *road* to recovery we're on the *road* to disaster the dog was wandering on the *road* to be on the right *road* to get out of the *road* our relationship has reached the end of the *road*	road up one for the *road* to get (or take) a show on the *road* to take to the *road*

Figure 11 Classification of the road *examples according to the audience (English-Spanish side).*

What conclusions can be drawn from this list? Again, the order of the examples does not coincide with the original list and the audience was clearly not the guiding principle. The lexicographers either tried to satisfy both audiences simultaneously, or had no specific audience in mind and chose their examples randomly. However, the list of examples intended for English speakers is much longer than the one for both English and Spanish speakers taken together. A look at one of the most likely translations for *road*, *camino,* makes one realise the degree of this bias. Classifying the examples according to this criterion confirms a tendency to favour the English needs. Since this makes the dictionary more effective for these users, it can only be regretted that this tendency was not implemented more thoroughly by cutting out the examples useless to them (and that potential Spanish buyers were not informed of this bias on the cover). The following figure classifies the examples for *camino* in the same way as the examples for *road* were classified in figure 12.

Examples useful for Spanish speakers (*encoding* into English)	Examples useful for both (*encoding* into English and *decoding* from Spanish)
camino de acceso	camino sin firme
camino de entrada	camino francés
camino forestal	camino de herradura
camino de Santiago	camino de ingresos
camino de Damasco	camino real
camino de peaje	camino de sirga
camino de tierra	camino trillado
el camino a seguir	tener el camino trillado
el camino de La Paz	camino vecinal
es el camino del desastre	Caminos, Canales y Puertos
el camino de en medio	vamos camino de la muerte
camino de Lima	a medio camino
en el camino	de camino
después de 3 horas de camino	tienen otro niño de camino
nos quedan 20 kms de camino	está en camino de desaparecer
es mucho camino	traer a uno por buen camino
¿cuanto camino hay de aquí a San José?	abrirse camino
por (el) buen camino	allanar el camino
¿vamos por buen camino?	echar camino adelante
errar el camino	ir por su camino
todos los caminos van a Roma	partir el camino con uno
llevar a uno por mal camino	quedarse en el camino
ponerse en camino	camino de mesa

Figure 12 Classification of the camino *examples according to the audience (English-Spanish side)*

The examples which I consider useful for (*encoding*) Spanish speakers are those which would offer some difficulty if they had to translate them into English. These examples include multi-word items such as *camino de acceso,* which a Spanish speaker might be drawn to translate literally, and sentences like *después de 3 horas de camino,* which cannot be translated as *after 3 hours of road.*

These same examples are of little use to English speakers. If one knows what *camino, acceso, después* and *hora* means, then the phrase *después de 3 horas de camino* is easy to understand.

As regards the examples useful for both audiences, *tienen otro niño de camino* will prove difficult for a *decoding* English speaker, as well as for an *encoding* Spanish speaker. It is striking that the list of examples useful for speakers of both languages is

longer on the Spanish-English side of the dictionary, confirming the bias encountered on the English-Spanish side.

I will save the reader of more detail, but if one compares the *Collins* entry to the corresponding entry in the *Oxford English-Spanish*, one notices some evolution. No audience –nor particular activity– has been chosen to classify the examples, but they have been ordered roughly into common uses and fixed expressions[10]. In practice this comes down to one subdivision into examples useful for *encoding* Spanish speaking users and one for both audiences. Unfortunately, since this has not been totally carried through. The choice of examples still seems a little fortuitous and many of the collocates of *road* which the COBUILD COLLOCATIONS CD-ROM shows to be the most common ones (*run down the road, continue down the road, head down the road, walk along the road*) have been left out.

4.3.2. Analysis of the Examples for *Road* in Oxford-Hachette

The *Oxford-Hachette* has been acclaimed as the most advanced of bilingual dictionaries. Its features confirm the upcoming trend of more *encoding*-oriented dictionaries even if this was not clearly announced in the dictionary itself. The entry for *road* lists a large number of examples classified according to their possible translations in French (mainly *rue* and *voie*). This semantic subdivision amounts to the traditional subdivision of *literal* and *figurative* since *route* and *rue* are used literally, and *voie* figuratively most of the time. Within this first subdivision, there is another subdivision according to the

[10] **Common uses:** is this the road to Boston?; the Cambridge road; five miles down the road; a factory just down the road (from here); there's a baker's over *o* across the road; the people from over the road; the house is set back a mile or so from the road

preposition required. At the end, the entry lists three fixed expressions (any road (up); let's get this show on the road!; one for the road). In fact there are a few more scattered over the rest of the entry, but still their number is reduced compared to the open-choice examples. This, together with the type of classification, characterises this dictionary as preponderantly aiming at an *encoding* audience. Classifying the examples according to a semantic criterion implies indeed that the user has an idea of the meaning in mind, which means s/he is *encoding*. This *encoding* trend is new, particularly in the case of bilingual dictionaries.

In spite of these innovations, the classification of the examples is still a problem which in the electronic version has been greatly helped by means of a simple search-facility. One can search for a word which usually accompanies the word looked up and get to the example immediately. When translating *mettre la main sur quelque chose,* one goes to the entry for *main* and searches for *mettre* which will instantaneously take one to the fixed expression. One can lament, however, that this is not a solution to the classification problem on the basis of a reflection on the needs of the users and the way these users proceed when using a dictionary. Indeed, the example loses in this case its function of 'exemplifying' since it represents only itself. Not only will the search facility not come up with the example if the morphological form of the search string does not coincide literally with the form in the example –it will not recognise *mettre* if the example uses *j'ai mis*– it will certainly not recognise analogous structures. It will not indicate the example *avoir une brûlure à la main* if the user searches the examples using an alternative such as *blessure, tatouage*, etc. The necessity for a sensible classification

---

Fixed expressions: a major/minor road; a dirt road; it's good road all the way now; by road; my car's off the road; to take to the road; road closed; road narrows; to have one for the road; all roads lead to Rome

of examples continues therefore to be on the agenda, even in the age of computerised dictionaries.

Conclusion

Except in the case of pedagogical dictionaries, dictionaries are not books for reading but tools. This means they should be made as efficient as possible and a sensible classification of the examples is one way of doing this. In his story on Chinese Encyclopedias, Borges shows how classification is relative and depends on what the classified information is supposed to be for. Efficiency implies adjusting means and purposes. As I have already suggested a number of times before, the fact that dictionaries do not separate *decoding* and *encoding* makes this difficult to implement.

I could discover no apparent logic governing the classification of the examples for *road* in *Collins English-Spanish*. They are not grouped according to figurative uses, nor according to their possibility of literal translation, nor to their usefulness for one or the other audience. The only order is maybe an alphabetical one following the prepositions. This, however, is not certain. The lack of organisation as a consequence of a lack of definition of the audience and its needs has a direct impact on the way users can profit from their dictionary.

First, the look up process is slowed down since the answer to the learner's question might be found anywhere, and nothing can be skipped. In my experience, this puts off more than one learner from using a dictionary altogether.

Secondly, there is no guarantee that users will find what they are looking for. No feature has been exploited in full. If an English speaking audience had been the declared audience, more examples of common usage should have been included and the list of fixed expressions should have been comprehensive.

Other possibilities which examples offer, such as giving contrastive information, were not exploited. It would have made sense to contrast examples in which *road* is translated as *camino* and when it is translated as *carretera.* There is an evolution under way of which the *Oxford-Hachette,* the *Collins-Robert* and, to a minor degree, the *Oxford Spanish Dictionary* are testimonies. However, the problem of classification remains intact and will become ever more pressing with the increasing use of corpora, capable of supplying an unlimited number of examples, and becoming a basic tool.

4.4. COBUILD AND THE AUTHENTIC VS. MADE-UP EXAMPLES QUESTION

Introduction

The authentic vs. made-up debate is linked primarily to the publishing of the first COBUILD dictionary in 1987 when authentic language was first used directly as a means to supply examples for dictionary entries. The mere fact of using authentic language was not in itself a novelty. In the field of native speaker's dictionaries, Dr. Johnson used authentic language in his *Dictionary of the English Language* (1755) and present-day lexicographers use 'citation files' with authentic language to compile their dictionaries. However, there are a few differences between traditional lexicographers and COBUILD.

First of all, before COBUILD 'authentic' language had never been tried out on learners. The difficulties in terms of vocabulary and syntax were considered too great. The first edition of the COBUILD dictionary was very much criticised by people like Hausmann and Gorbahn (1989) and, to a lesser degree by Fillmore (1989). In reality, the level of learners of English as a foreign language has probably improved due to the globalisation of all sectors of human activity and the dominant role English plays in it as a language. COBUILD offered a first response to this trend.

The second difference is that, until recently, citation files were used as a source of inspiration rather than to provide on-hand examples. COBUILD takes its examples directly from the corpus and puts them into the dictionary with little or no modification.

Thirdly, Dr. Johnson, the 'canonical critic proper' as Harold Bloom calls him (1994:183), would never have dreamt of using sources other than 'canonical' authors. That is why today his dictionary can still be read as a dictionary of quotations. COBUILD, for its part, uses a corpus which includes literary authors, but consists basically of non-fiction, newspapers, magazines, and speech. From Johnson's dictionary which had the aim to fix the standard of a language, COBUILD has moved on to a more modern linguistic conception of describing the language instead of being strictly normative.

4.4.1. The Theory Behind the Defence of 'Authentic Language'

The concept of 'authentic language' is not as plain as it may seem. According to those who defend the use of 'authentic language' in pedagogic materials, specifically in dictionaries, it is not sufficient for an utterance to be produced by a native speaker to

have a right for it to be called 'authentic'. This is even less so if this utterance was produced to illustrate a grammar item or to show the meaning of a word. To use John Sinclair's (1988) terminology, metalinguistic utterances of this kind can be 'grammatical', but they are not 'natural', and therefore not 'authentic'.

In this way, the COBUILD undertaking ended up criticising indirectly the other learner's dictionaries for making up samples of language that had not actually occurred and whose main purpose was to clarify meaning. According to the COBUILD team, this practice gives an incorrect idea of the language and thus misinforms the learner. Fox, a senior member of the COBUILD team, claims:

> If a word typically occurs in a sentence which is grammatically complex or alongside vocabulary items that are infrequent, it would be misleading of a dictionary to present that word in a very simple clause or sentence with easy vocabulary. (Fox 1987:138)

According to John Sinclair, the theoretician behind COBUILD, each sentence carries with it the characteristics of the text from which it was extracted. He calls this phenomenon 'encapsulation'. In Sinclair's opinion,

> a text is represented at any moment of interpretation by a single sentence... each new sentence encapsulates the previous one by an act of reference. (1993:7)

As a consequence, made-up examples are deceptive since the lexical or grammatical choices made in them do not depend on any text whatsoever. They can be 'grammatical', but they are not 'natural' and it is not enough for learners to be able to produce sentences that are grammatically well-formed. They must also be recognised as 'natural' by native speakers (Fox 1987:139). According to the COBUILD team, traditional lexicographers, being native speakers, make up examples that are grammatically acceptable, but because their intention is not to communicate anything,

other than information on a lexical item, this information is, strictly speaking, incorrect. Therefore, in these made-up examples, there is a chance of the words carrying negative connotations to be used positively; or for the collocates to be wrong, or the syntactic construction to be correct but not usual. In other words, a made-up example has no validity as a model. Sinclair states:

> (made-up examples) have no independent authority or reason for their existence, and they are constructed to refine the explanations and in many cases to clarify the explanations. They give no reliable guide to composition in English and would be very misleading if applied to that task. (...) Usage cannot be invented, it can only be recorded. (Sinclair 1987:XV)

Most people will not doubt the exactitude of most of the theoretical premises on which the choice for authentic examples was based. Even so, much criticism has been heard, not so much of these theoretical premises, as of their practical implications.

4.4.2. Criticisms of COBUILD and the Authentic Examples Policy

At the time of its first edition, COBUILD was the only dictionary to use an electronically-accessed corpus of sizeable proportions. Before this, the technical means were hardly available and only recently has computer technology made the search for and retrieval of examples in large data bases a technically easy task. However, while there is this technical side to the problem, it is not the heart of the matter. Probably the main reason why traditional lexicography would not use authentic language was its lexical and syntactic difficulties. It was also thought there would be too much interference from an absent co-text, whereas examples were supposed to teach a particular feature, singled out by the lexicographer and concocted with this feature in mind. The results of the

questionnaire (Ahmad et al., 1992) distributed among lexicographers, which I discussed above, confirms this.

In practice, learner's dictionary publishers had not been aware of a general change in attitude of the public and its needs. COBUILD was the first attempt at renewing foreign language lexicography. Criticisms followed and people like Della Summers (1996) (Longman) popularised a distinction between corpus-based and corpus-bound examples aimed at stressing that lexicographers should find in the corpus their inspiration, check frequencies and collocations, but not be bound to it. In reality, after COBUILD, all dictionaries started emphasising they used a corpus.

The debate on authentic examples can be subdivided into two sections: the problem of corpora, and the didactic implications,

Corpora

The idea that a corpus is a valuable aid for the compilation of dictionaries has been generally accepted and at present all learner's dictionaries stress the fact that they use one. It is probably because of Sinclair that lexicographers have acquired a sound distrust of their own intuitions, and this applies particularly to making up examples. When comparing contemporary examples to those of a few decades ago, one notices that an effort has been made to make them sound less contrived. Although there are still disagreements as to the management of a corpus or its representativeness, all lexicographers now agree that the non-corpus based dictionary belongs to the past.

Other parts of Sinclair's theory however, have not been as widely accepted. The authentic/made-up distinction has been empirically questioned. In an experiment conducted by Batia Laufer in 1992, she concludes that native speakers are unable to distinguish between authentic and made-up examples. I myself did a similar experiment with native speakers of Portuguese which confirmed Laufer's results.

This does not necessarily mean there is no difference between authentic and made-up language, it does mean that this difference is not immediately apparent. One might object that made-up examples only contain what the lexicographer has put in it whereas authentic examples carry information of which the lexicographer is unaware. This, however, is not self-evident, it is hard to prove, and it does not mean it is better for the learner. If authentic language has an unconscious salutary effect, this has not been proven. The idea of 'encapsulation' sounds attractive, but if native speakers mistake made-up examples for authentic ones and vice-versa, this part of Sinclair's theory does not really convince. And even if one makes allowances for this unproved distinction, the question remains whether corpus samples should become dictionary examples without any previous adaptation, since dictionary examples have a didactic function.

The Didactic Point of View

COBUILD's argument claiming that learners should learn *natural* and not *grammatical* language, is controversial. One can accept, even without any empirical proof, that there is a distinction between the two kinds of languages. If *natural* language is more valuable because it conveys intended and unintended information, it is still the question whether this unintended information is not conveyed just as well by a sentence of which one

aspect was intended. In other words, a lexicographer could concoct a sentence to demonstrate a particular aspect of a word –e.g. the different structures of *indulge*– and unknowingly also transmit other information about this word.

Secondly, one wonders if natural language is more suited for teaching than grammatical language. There are several stages in learning a foreign language, each characterised by a particular kind of language. Even very advanced L2 speakers use a language which is different from a native speaker's. Grammatical language is only one step away from natural language and many advanced learners will think themselves fortunate if they ever get near to it. It could be argued that grammatical language can best be taught with sentences made up by a native speaker with a specific purpose. In fact, they allow a lexicographer to draw the attention of the dictionary user to specific features. Undoubtedly, this would work at its best if lexicographers had a contrastive knowledge of the audience's language which is not usually the case. But even without this knowledge, grammatical analysis should lead a lexicographer to an awareness of what are the features which the learner's attention should be drawn to. Authentic language examples are much harder to control in terms of syntactic and lexical intricacies. Besides, the fact that a sentence is unnatural does not mean that it is not instructive. To put it in an extreme way, 'colourless green ideas sleep furiously' may well be meaningless to native and non-native speakers alike, but it teaches a basic English grammar: NP-VP-AP. An unnatural sentence such as 'this is a pencil' can make a grammar point more clearly than an authentic example. This does not mean it *has* to be made-up and unnatural. It means that it very well *can.* It all depends on what is to be taught.

In an above mentioned article, Laufer (1992) reports on an experiment which tested the efficiency of authentic examples as opposed to made-up in comprehension and production tasks. This led her to conclude that:

> Lexicographer's examples are more helpful in comprehension of new words than the authentic ones. In production of the new word, lexicographer's examples are also more helpful but not significantly so. (1992:75)

And:

> The findings suggest that lexicographer-made examples are pedagogically more beneficial than the authentic ones (further studies would be useful to substantiate this claim). If this is so, can they be considered as unacceptable on the grounds of lack of naturalness? (1992:76).

As is the case with a number of experiments conducted in the area of language learning, the results of this survey depended on a range of more or less disputable decisions: Laufer had to choose a set of examples, then a number of subjects of which the previous knowledge is difficult to measure, and eventually she had to write a test which conceivably could have been written in various other ways. All this leads us to take these results *cum grano salis.*

Furthermore, as Jean Binon pointed out to me in a personal communication, the lexical density of authentic examples is often so high that understanding them becomes a problem in itself. If one looks at it from a cultural point of view, references to proper names and dates of events are rapidly outdated and can make a sentence incomprehensible. In the second edition of their dictionary, the COBUILD team took special care to avoid this kind of reference. But to eliminate cultural information entirely is difficult, particularly if you are a member of this culture and unaware of what requires a special knowledge of this country. Language is culture, of course, and it should be

transmitted to the learner, but not everything has to be learnt at the same time, and dictionary examples may not be the best place to do this.

Finally, in terms of the didactic implications of the use of authentic language, there is the psychological factor which is rarely or never commented on. In conversations with learners it appeared to me that the feeling of being in direct touch with 'authentic English', as the COBUILD cover says, was an asset. Since learning the language remains a basically psychological matter, this factor should be taken into consideration as well.

Behind an innovative project such as COBUILD one sometimes has the feeling that a view on foreign language teaching and learning is missing. This is, by the way, something COBUILD has in common with all other learner's dictionaries, except the *Longman Language Activator*. No one nowadays sustains that it is not valuable for an analysis of language to examine a corpus, the question is which is the most efficient way to transmit the knowledge that is obtained via this analysis. One thing is to analyse the language and see that, although 'blatant' means 'obvious', apparently no one speaks of a 'blatant choice'; another thing is to teach this to learners. If a collection of utterances taken from a corpus reveals a particular usage pattern for a particular lexical item, this knowledge is certainly worth transmitting to the learner. It is, nevertheless, the question whether this knowledge can be conveyed more clearly using the raw material which led to this knowledge.

Conclusion

In the field of lexicography, the use of authentic language was a radical innovation. As

tends to be the case with innovations, it had to be defended. The result was a predictable radicalisation of the COBUILD team and a renewed insistence on 'orthodoxy'.

Meanwhile, competitors have adopted an intermediate position which is, in my opinion, likely to become the standard: interrogate the corpus for evidence, and use adapted authentic, or authentic examples in the dictionary. In the following section, I will discuss to what extent speaking of 'authentic language' is a fallacy.

4.5. CONCLUSION

The area of examples in learner's and bilingual dictionaries has still some way to go and a few problem areas have been identified. There is a relationship between the choice of examples and the task one assigns to dictionaries for foreign language learners in general. If the emphasis is on *decoding*, the task is relatively simple. If it is on *encoding*, it is necessary to describe what the needs of the *encoding* learner are. When *encoder*s turn to a learner's dictionary, they are less likely to have a problem with meaning than with syntax and collocates.

In my experience, this information can best be presented in the form of examples. Until now, either no choice has been made in favour of *decoding* or *encoding*, or there is only a patchy awareness of how these goals can be fulfilled. However, the tools are at hand. Most dictionaries now have a corpus and the COBUILD Collocations CD is freely available on the market. The information only has to be made accessible to the learner in a systematic way, and integrated with the other elements of the dictionary.

In doing this, care has to be taken to deal with each word as an entity, since every word is characterised by particular difficulties. This is true from the viewpoint of a native speaker, and it is even more true when one puts the lexical item in contrast with the learner's first language. Some dictionaries such as *Cambridge International Dictionary of English* have gone some way in modelling their work according to the first language of their audience. It is clear, however, that these are first attempts and that the research underpinning these projects is still in its infancy. In spite of the criticisms, the future will give *Cambridge* its due.

Since English is becoming a most necessary second language with a high number of potential dictionary buyers, it is probable that in the near future more target-conscious *encoding* dictionaries will be put on the market with examples chosen accordingly. Whether these examples will be authentic or made-up seems to be of less concern than was initially thought. Probably a mixed form of modified authentic examples will be the eventual outcome of the authentic vs. made-up debate. However, since electronic dictionaries are most likely to be the future of lexicography and space will no longer be a problem, it is reasonable to think that in the future learners will have at their disposal both made-up and authentic examples integrated into the same tool.

CHAPTER 5 AN INTEGRATED APPROACH TO FOREIGN LANGUAGE LEXICOGRAPHY

In the following pages I will set out my views on what I see as the foreign language dictionary of the future. The main characteristic of this new dictionary is that it clearly distinguishes between a *decoding* and an *encoding* part. This suggestion is in itself not new, but it is to the best of my knowledge the first time that the consequences of such a separation have been practically thought through. Elaborating on this will take up most of the present chapter.

Although I will not go into details as regards the software aspect of the question, I take it for granted that the dictionary of the future will be electronic. Anyone who uses electronic dictionaries on a regular basis does not need to be persuaded of the advantages which computers offer in this area. Taking into account the chronic need for space in traditional dictionaries, unlimited storage capacities are the most obvious advantage of a computer, but speed and random access are equally important. I shall not discuss aspects such as customising possibilities, the compiling of a personalised dictionary, innovative search features, hyperlinks and others, which will all considerably improve the life of language learners. Since these innovations are basically technical, I consider them to fall outside the scope of my research.

Although it should be no coincidence that both a new proposal and the means to implement it emerge at the same time, what I put forward is not dictated by what computers can do, but is based on an analysis of the learner's needs. Computers just happen to meet these needs better.

The reader will also observe that this new dictionary does not presuppose a large amount of work at the micro-structure level. I am not proposing a different kind of definition, nor a different kind of translational equivalent. I basically propose to break up in parts what has been achieved all along the history of lexicography and restructure this material in a different way. Although one of the main characteristics of computers in the field of reference works is to allow a user random access to the information, an important part of my proposal consists in defining the order in which a search is carried out and what kind of tools are necessary to make this search successful.

Encoding vs. decoding

Asking for separate *encoding* and *decoding* dictionaries comes down to asking that the *encoding* parts of dictionaries be emancipated from the *decoding* tool which current dictionaries still basically are. As I suggested in previous chapters, often the *encoding* information is present in the dictionary, but is submerged by *decoding* information, which hampers the look-up process. On the other hand, there is no doubt that the more dictionaries attempt to meet *encoders*' needs, the more they burden the *decoders* with unnecessary information.

The first request for a different set of dictionaries for *decoding* and *encoding* dates back to the thirties when Lev Sjtsjerba:

> launched the idea that two types of dictionaries have to be compiled –one type for users who translate from a foreign language into their mother tongue, and the other for users who translate from their mother tongue into a foreign language. Therefore, according to Sjtsjerba, for a particular pair of languages it was necessary to have four dictionaries: A➔B and B➔A for users with the mother tongue A, and A➔B and B➔A for users with the mother tongue B. (Berkov, 1996:547)

This idea of 'Deux langues, quatre dictionnaires' (Bogaards, 1990) was afterwards periodically repeated by a number of researchers, without practical consequences. It is rather easy to think of a few commercial reasons why Sjtsjerba's proposal was never put into practice. On the other hand, Sjtsjerba's proposal can be regarded as the demand of an *encoding* user and at the time he made it, dictionaries were still predominantly used for *decoding.* In 1983, Cowie outlined the problem tentatively as follows:

> Some foreign learners may be more concerned with interpretation than production –they may use the dictionary primarily for quick retrieval of individual items or meanings when reading– but no general dictionary for the advanced student can afford to neglect one need in favour of the other. (Cowie, 1983:136)

The essence of the problem had by then long been understood. Apparently, however, it seems like there would still not be a majority in favour of this kind of proposal.

> I wish to suggest as a general principle that the interpretative function places a high premium on ease of access and thus on the strict alphabetical ordering of entries. The productive function, on the other hand, places a high premium on the clustering of derivatives, compounds, idioms, phrasal verbs, and so on around the simple lexemes (or particular meanings of those lexemes to which they are related.) (Cowie, Ibid.: 141)

But however obvious these claims for separating *encoding* and *decoding* seem to many, to others this seems a rather superfluous demand. In 1983, Zgusta was not convinced of the necessity of separating *encoding* and *decoding*:

> The reason for this overlapping (of *decoding* and *encoding* dictionaries) is clear: the statistically 'normal user' does not wish to buy several dictionaries of the same language, and therefore many dictionaries are designed to serve more than one purpose. Observably, one of the purposes taken care of in such a multi-purpose dictionary usually enjoys a degree of preference; nevertheless, the chance of being useful to more sets of users and therefore appealing to a broader public (i.e. more buyers) proves to be attractive to many editors and to most publishers. (Zgusta, 1983)

Of course, Zgusta wrote this fifteen years ago, when the separation of *decoding* and *encoding* dictionaries would mean an enormous pile of paper. The prospect of having to print four dictionaries where before a single one would do, cannot have appealed to most commercial publishers. Apparently, time was not ripe for dramatic changes in dictionary conception. But with the advent of the computer era needs would change, and so would the means to meet them.

Judging from my personal experience as a learner and as a language teacher, I am convinced that the main transformation dictionaries have to undergo is the separation of *encoding* and *decoding* . The information needed to compile a *decoding* dictionary can of course be put to use in an *encoding* dictionary and both can be part of one and the same electronic dictionary.

In what follows, I will deal with the *decoding* and the *encoding* part of the dictionary separately, following the sequence which learners follow when they look up a word. Since an electronic dictionary can be randomly accessible, this sequence is not a necessary ingredient of the dictionary I propose. It would only be like a path put at the disposal of the users if they wished to follow it.

As the reader will observe, the part of this next section dealing with *decoding* is much shorter than the part dealing with *encoding*. I indeed consider *decoding* to be a lesser problem than *encoding*.

5.1. *DECODING*

Decoding can be understood in two ways: *decoding* to *understand* and *decoding* to *translate*. In both cases one goes from a language one knows less well to a language one knows better. This means, in most cases, going from a foreign language to the mother tongue. In the first case, a dictionary user is reading a text, or looking up a word s/he heard, merely in order to understand a message. This situation will by far be the most common one and corresponds to the primary tasks of a dictionary.

Translation, on the other hand, involves a process of *encoding* which follows the process of *decoding*. It is because of this *encoding* part, and not the *decoding part*, that *translation* gave rise to a separate discipline of 'translation studies' with its considerations of a linguistic as well as a broadly cultural and literary kind. A number of distinguished authors (Jakobson, 1959; Mounin, 1963; Catford, 1965; Steiner, 1971; Basnett-McGuire, 1991) show translation is a complex process, so complex that a few authors have come to the conclusion that it is virtually impossible. However, as Mounin convincingly states:

> Si l'on accepte les thèses courantes sur la structure des lexiques, des morphologies et des syntaxes, on aboutit à professer que la traduction devrait être impossible. Mais les traducteurs existent, ils produisent, on se sert utilement de leurs productions. On pourrait presque dire que la traduction constitue le scandale de la linguistique contemporaine. (Mounin, 1963:8)

However, being aware of the complex nature of the translation process, I think this distinction between understanding and translation can be disregarded for the specific purpose of my research, and I will not take it into consideration here. It seems to me that the demands put on dictionaries by translation –in the *decoding* and subsequent

encoding sense of the word– are of a kind native speaker's dictionaries should tackle, even if in the future nothing will prevent elaborate native speaker's synonym dictionaries and thesauruses from being included on one CD-ROM, together with a foreign language dictionary, thus merging native speaker's and foreign language lexicography.

Decoding is at any rate easier than *encoding.* Particularly for advanced learners, most of the existing reference tools will be adequate, even if open to improvement. This does not mean *decoding* is easy. It is, nonetheless, on the whole a smoother process than *encoding* because the learner is expected to have more resources in the language s/he knows best.

Decoding always consists of some form of translation, be it interlingual or intralingual (Jakobson, 1959), and can be assisted either by a bilingual or by a monolingual dictionary. *Advanced* learners in particular can choose between a bilingual, a native speaker's, a synonym, a learner's, or a bilingualised dictionary, at least in the case of a few languages. As a number of surveys show (Tomaszczyk, 1979; Baxter, 1980; Bensoussan et al., 1984) and common sense dictates, bilingual dictionaries are in most cases the obvious choice. If advanced learners often prefer a monolingual reference work, this is probably due to the fact that current monolingual dictionaries still cover a broader area of the lexis than bilingual dictionaries, which does not necessarily mean that an intralingual translation is better than an interlingual one. But whatever kind of dictionary advanced learners prefer, the existing tools are usually sufficient.

Even if the dictionary confronts them with more than one translation or definition, the context will help them out.

The situation of *beginning* learners is different. In what follows I will analyse beginners' problems assuming that they use almost exclusively bilingual dictionaries and that this unavoidable use of bilingual dictionaries characterises them as beginners. In the figure below I list different kinds of items beginners may have trouble looking up in a dictionary.

Example	**Type**
1. *found/find* *tabemasu/taberu* (to eat *in Japanese*) *omoshirokunai/omoshiroi* (not funny/funny, *in Japanese*) *jemandem/jemand* (*someone*, dative and nominative *in German*)	morphological problem
2. *road block* *swimming bath* *swimming pool* *swimming costume*	multi-word items
4. *at first glance* *at great length* *in zeven haasten* (in a hurry, *in Dutch*) *de uma cajadada só* (to kill two birds with one stone, *in Portuguese*)	fixed expressions
5. *wie laatst lacht best lacht* (he who laughs last laughs longest, *in Dutch*) *muita esmola o santo desconfia* (too much effort makes one suspicious, *in Portuguese*)	proverbs
3. *in aanmerking komen voor* (qualify for, *in Dutch*) *dar uma esperadinha* (to wait, *in Portuguese*)	idioms
6. *lead* *break*	polysemy

Figure 13 Decoding *problems of beginning learners.*

Assuming that the information which learners are looking for exists in the dictionary, their failure to retrieve it can be traced back to three causes. First, the dictionary form is different from the looked up form. Second, the word is part of a *multi-word item* not recognised as such. Third, the item is a case of *polysemy*.

In what follows I will deal with each of these problems separately. I will tackle the problem of different word forms and multi-word items together in one section.

5.1.1. Morphology and Multi-Word Items

The problem of learners being unable to find the right entry because they do not know the dictionary form of the item they are looking for is the easiest problem to solve. It is sufficient to include all forms in the dictionary, as is already a common practice in a number of English dictionaries for foreigners, at least for verbs. This may not be possible for all types of languages, but it is for all languages I have some knowledge of, including an agglutinating one like Japanese. Including all word forms may result into large lists, but a computer can generate them and there is no storage problem on a CD-ROM. It would be useful to include plurals, verb forms and others, depending on the language.

The problem of *multi-word items* is somewhat harder to tackle. However, in the case of long recognised multi-word items such as *swimming bath, swimming pool, swimming costume*, and the like, the solution is relatively simple. In an electronic dictionary, looking up *swimming*, *bath*, *pool* or *costume* should show the learner all possible combinations. The capability of electronic dictionaries to present the information on successive screens prevents the user from feeling overwhelmed and put

off by the sheer excess of information. When looking up *swimming*, the words *bath*, *pool* and *costume* could pop up and would trigger a reaction from the learners if the same word(s) were part of their context. For instance, if not aware of the fact that *levée de terre* was one single item, learners looking up *levée* would be confronted with *terre,* which would automatically draw their attention to the multi-word item.

Thirdly, *fixed expressions* can also be considered multi-word items at the same level as *proverbs*. *At first glance*, *at great length*, *in zeven haasten* and *de uma cajadada só* do not allow any variation in their sequence and function as single words. In a traditional dictionary, users had to look up these items under one of the words the expression consisted of. Because of lack of space they were not included under all of the entries involved. Since some of these entries would end up being very long, there was always a danger that the learner would miss the item out. In an electronic dictionary, fixed expressions do not have to be classified anywhere in particular and it is possible to access them through any of their component words. This has already been implemented in the COBUILD Dictionary on CD-ROM and other electronic dictionaries such as *Euroglot*. Again, the information is not presented to the learner all at once, but on successive screens, which allows the dictionary to point out the existence of the fixed expressions.

The third variety of multi-word items consists of *idioms* in the Sinclairian sense of the word which I will discuss next. Finally, I will deal with the problem of polysemy.

5.1.2. Idioms and Collocations

Sinclair defines *idioms* as 'semi-preconstructed phrases that constitute single choices,

even though they might appear to be analysable into segments' (Sinclair 1991:110). Under this heading, Sinclair classifies a number of items which, for lexicographic reasons, I prefer to categorise in two groups.

The first group consists of rigid constructs which I consider *multi-word items*. In *of course*, example given by Sinclair, two words are intimately linked forming virtually one single lexical item translated by one word in many languages (*claro*, *natuurlijk*, *natürlich*), in two words (*bien sûr*) in others. I consider I have dealt with them under the previous heading.

The second kind of idioms is the one Sinclair exemplifies through phrases like *set one's eyes on* and *it's not in his nature to*. These idioms allow *lexical*, *syntactical* and *word order* variation, which makes it much harder for a dictionary to deal with them (Sinclair 1991:112, 113). However, dictionaries should, and for the same reasons. It is indeed typical of idioms that in most cases –depending on the target language–a word for word translation does not make sense. Even in Dutch, a language closely related to English, a word for word translation of *make up your mind, maak op uw verstand*, will only lead very clever Dutch speakers to understand this as *beslissen*. In French, the literal translation is total gibberish (*faire sur ton esprit*), just as in Portuguese *(fazer sobre tua mente).*

Dictionary compilers have been aware of this problem and a number of idioms are commonly listed under one form or another in some entry. Others, like *it's not in his nature to,* have never been even recognised as such. The fact that even native speakers are often not aware of the fixedness of these chunks of words, and their changeable

nature, makes it difficult to solve this problem in a dictionary. In addition, the challenge for the lexicographer is not so much to translate these idioms as to make learners aware of the fact that they *are* dealing with an idiom. There is a possibility to draw their attention to it when there is a word in the idiom which the learner does not understand or when the literal translation does not make any sense.

In the first case, a dictionary should simply put all possible idioms in which a word can take part before the learner's eyes. In the second case, the dictionary should allow users to input all the words they suspect form one unit.

One reason why this problem is difficult to solve in traditional dictionaries is that these dictionaries must necessarily choose one single classification order, which makes cross-referencing difficult. In a traditional dictionary, idioms can be classified either under the word they start with; under the word which is considered most important; or according to the semantic field. Not all of these orders can be used at the same time. Traditional dictionaries often have the right information, but they have limited ways of making it accessible.

Since an idiom is a mutable construction, it is a challenge for the lexicographer to pin it down. In the case of *make up your mind* there are all the possible morphological variations of the verb and of the possessive pronoun, the variations in word order, plus the intercalation of other optional items which can result in phrases like 'make your *own* mind up'. However, if to make a list of all these possible variations is not an easy task, a corpus makes this feasible. This presupposes there is a list of all possible idioms and, according to Sinclair, this project is, under way. At least in English.

One way of making the *decoding* of idioms easier is by drawing the attention of the dictionary user to the most frequent combinations in which a word can appear. Even for highly frequent words such as *take* this is quickly done with the help of a simple tool like the COBUILD Collocations CD-ROM:

(take) part in; (take) place; (take) off; (take) legal action; (take) care of; (take) more than; (take) some time; (take) some kind of; (take) advantage of; (take) full advantage of; (take) a (very) long time; (take) (too) long; (take) any chances; (take) a (quick, hard, detailed, fresh, long, hard) look; (take) account of; (take) into account; (take) years, months, weeks, days.

Figure 14 Collocates of take *on the COBUILD Collocations CD-ROM.*

This list is probably not complete for reasons which can be possible flaws in the construction of the corpus, failure to recognise mechanically the morphological variations of the idiom, or simply limitations of the CD-ROM as a medium. Even so, it is a start and looking up the collocates is easily done. Surprisingly, I do not know of any dictionary where this has been systematically undertaken, not even in dictionaries which profess to work with a corpus. Undoubtedly the listing of all possible idioms would take up much space in a paper dictionary, but not on a CD-ROM. In addition, the same information could be used in several different places at once. Taking advantage of the sequencing possibilities of the computer, looking up *take* would bring up other possible collocates: *part; place; off; legal action; care of; time; kind; advantage; long; chances; look; account; years; months; weeks; days*. When looking up *part*, *take* would also be one of the possible directions in which to proceed, together with *time; most; play; any; important; because; large; life; world; first; also; become; country; work.* Even ignoring the fact that the word they look up is part of a larger item, a glance at the possible collocates would quickly make learners aware of this.

To speed up *decoding*, the likelihood that a word will combine with another word should be ranked. In the case of *place*, the most likely combination is *take*. If a word can form idioms, there is a fair chance it will form more than one and in this list of combinations one collocate will be more frequent than another one. The current programmes designed for searching a corpus automatically detect this sequence.

5.1.3. Polysemy

Words can be polysemous within a language itself, or only when looked at from the point of view of a foreign language. In the first case, which I would call *internal polysemy* and in which the various senses of the word will frequently be metonymically related to each other, a word refers to two or more different entities (*hammer* as a tool; as a part of a gun; as a device on a piano). The second kind of polysemy only appears in translation and I will refer to it as *external polysemy*. Native speakers of French do not normally know that *quelque chose* can be understood in two different ways and can be translated in English by *anything* or *something*. Another example is *esperar* in Portuguese, which can be translated as *expect, hope*, or *wait*. The average language user is not aware of this kind.

It is the first kind of polysemy which is troublesome for *decoding,* whereas the second kind is a problem for *encoding*. From the point of view of the foreign learner, most words are polysemous, since high-frequency words have easily more than one possible translation, particularly verbs. When looking up a word in a bilingual dictionary, the user will frequently find that more than one translational equivalent is

suggested. This is, *mutatis mutandis*, also true for definitions in a monolingual dictionary.

Cases of polysemy are elucidated by taking into account the *context*. Choosing between the various meanings of a word will be easy for advanced learners, since they understand the context from which they retrieved the word. It is more difficult for beginning learners. It would therefore be a solution if a 'word' could be amplified to some sort of larger lexical item containing the context of whatever nature it may be. This is what a dictionary should endeavour to do.

The *Random House Webster's College Dictionary* lists 62 different meanings for *lead*. A number of them mean roughly the same, but quite a few have to be translated by different words, depending on the language. According to my calculations there would be at least 30 different translations for *lead* in French, Portuguese and Spanish (in Spanish: *plomo; balazos; sonda; escandallo; excusas para no trabajar; mina; lápiz; emplomado; cabeza; primer lugar; ejemplo; iniciativa; pista; correa; traílla; cable; papel principal; protagonista; solista; principal; artículo principal; introducción; mano; llevar; guiar; conducir*; etc.) The right translation depends in some cases on the immediate context of *lead*, in others on its broader context, in still others on whether *lead* is a *verb* or a *noun*. I hope to make this statement clear below, when I deal with the specific case of *lead*.

The question of how to describe a word in terms of context –how to build up a lexical item by broadening the boundaries of a word– is closely related to how to look up a word in a corpus. When one knows how to look up a specific meaning of a

polysemous word in a corpus, one has described it as a lexical item. If I know what are the specific characteristics of a word I have to mention to retrieve only those instances in which it has a meaning X, and not Y, I know everything that is necessary to disambiguate it.

However, current programs to analyse a corpus, such as *Look-Up* at COBUILD, are only partially satisfactory since they limit searches to a nine-word-span. They are not able, for obvious reasons, to pick out the pertinent words from the rest of the text which the word was extracted from and which, in some cases, determine its meaning. Words which determine the sense of a polysemous word may not be syntactically related to it. If one takes the case of *lead* and parses it out by translating it into Portuguese, one comes to the following figure:

I, you, she, he, we, they **lead**	liderar, encabeçar
I, you, she, he, we, they **lead** + life	levar uma vida
lead + to	levar
lead + story	artigo principal
lead + [cachorros]	corrente
lead + swing	inventar desculpas
of lead	de chumbo
's lead	liderança
point **lead**	vantagem
to lead	guiar
lead +[química] [exploração de minas]	chumbo
will **lead**	levará/levarão
would **lead**	levaria
[auxiliar] + **lead**	levar
the **lead**	a liderança

Figure 15 Translations of lead *according to co-text and context.*

The figure above gives an overview of how *lead* is translated into Portuguese, depending on the co-text and the context. It is a partial description and it will be clear from this example that to describe the whole lexis of a language in this way would be a

huge task. It might, however, be worthwhile. This figure shows how easy it is to help a beginning learner determine the meaning of a word by means of one other word or an indication of the general topic of the text. This might seem a bit laborious to an advanced speaker, but beginning learners are more willing to put some effort in their search since they cannot cheerfully skip words and go on understanding the text.

When looking at a corpus it is astonishing which elements, and how few elements, characterise a specific meaning of a word. These elements of course, do not have to belong to the 'nine-word span'. In the case of *lead*, all combinations with *'s* automatically rule out any other translation if not *liderança*. Similarly the presence of *of* indicates the cases in which *lead* means *chumbo*[11]. This is also the correct translation if the topic of the text is *mines* or *chemistry* or a related area. This analysis of *lead*, which should of course have a more user-friendly aspect in a software dictionary, deals with the question in a pragmatic and not in a scholarly way. Hence the motley aspect of the figure. The purpose is to help the user out by whatever means this may be. The way in which the data are presented here is not the way in which they should appear on a screen. The intention is only to give an example of how a polysemous lexical item can be disambiguated by adding succinct information, possibly of very different kinds: syntactic, semantic, morphological. It is indeed a mixture of formal information (co-text), meaning information (context) and grammatical information (e. g. 'auxiliary').

A possible way of presenting this information could be in the form of a grid, comparable to the 'picture' of a word in a programme such as COBUILD's *Look-Up*.

[11] Looking at it from another angle, *of lead* and *'s lead* can be considered multi-word items just like, e.g., *railway-track*.

Users would click on a word in the grid which they would recognise from their context and the translation would pop up.

	Would	**lead**	to	
	the		singer	
	point		story	
	will			
	would			
				into
	's			
	auxiliary			
	personal pronoun			
exploração de minas	*música*	*jornalismo*	*política*	
química	*cachorros*			

Figure 16 Example grid of the entry lead *in an electronic dictionary.*

The case of *break* is slightly different from *lead. Break* is even more polysemous than *lead.* It is also more devoid of meaning, which means its sense is less dependent on the context and almost exclusively on the co-text. Apart from more traditional types of co-text such as prepositions or objects in the case of *break* as a verb, less expected words like *came* indicate in what sense *break* is being used. The grid below is only an indication of what an entry could look like and is by no means comprehensive. A great number of occurrences of *break,* in all its forms, should be translated in the target language and in each case it should be established what words determine a particular translation.

break *descanso* (substantivo) ou *quebrar* (verbo)				
	a/the	descanso, pausa		
	take a	descansar, fazer uma pausa		
		From	interromper para descansar	
	+ into + car house home building	Assaltar		
	+ away	Desprender-se, libertar-se		
	big +	Grande oportunidade		
	+ bones	Quebrar		
	+ off	Separar-se, desistir, interromper		
		+ diplomatic relations negotiations dialogue	cortar	
	+ out	Rebentar, brotar		
		+ fighting disturbances epidemic fights conflict war row	começar	
	+ free	Libertar-se		
		+ from	libertar-se de	
		+ of	libertar-se de	
	+ glass	Quebrar		
	+ up	partir-se, dissolver-se		
			+ marriage	fracassar
			+ protests demonstrations	dissolver
			we, they +	separar-se
			+ with	separar-se de
			the break-up of	dissolução
	+ came	oportunidade		
	+ ranks	sair de forma		
	+ record	quebrar um recorde		
	+ even	sair sem perder ou ganhar		
	+ rules	violar as regras		
	+ law	violar a lei		

Figure 17 Example grid of the entry break *in an electronic dictionary.*

The grid above is based greatly on an analysis of the 'collocations tree' of *break* as taken from the *Bank of English*. This collocation tree looks like this:

collocate	frequency of collocates in 1000 sentences in which *break* occurs	statistical significance of the collocate (over 2 is considered significant)
down	91	8.823179
up	71	6.322514
out	67	6.189736
after	43	4.810485
record	22	4.364512
into	38	4.231458
a	239	4.157161
in	201	3.959329
his	66	3.914527
rules	16	3.815851
when	41	3.701246
silence	14	3.652114
the	525	3.641007
off	24	3.573706
had	50	3.177460
away	16	3.145142
leg	10	3.025647
tax	12	3.003254
through	19	2.990364
free	13	2.977855
marriage	10	2.956224
lunch	9	2.845882
war	14	2.813841
before	19	2.808706
bones	8	2.769681
glass	8	2.603870
over	23	2.587004
law	10	2.583217
during	12	2.485047
big	11	2.396905
talks	8	2.381263
try	9	2.379298
tie	6	2.331762
since	12	2.259616
laws	6	2.250825
taboo	5	2.222910
deadlock	5	2.217709
trying	8	2.180300
taking	8	2.177877
ranks	5	2.158782
news	9	2.151499
Heart	7	2.125252
Fighting	6	2.122338
Diplomatic	5	2.094875
Ago	9	2.087023

Figure 18 Collocation tree for break.

Clearly, a number of these collocates have to be eliminated since they do not influence the translation of *break*. Words like *during, since, before* or *ago* show that *break* is often used with an indication of time, but this is not enough to know what *break* in these cases means. Most polysemous words deserve an in-depth investigation into their behaviour, and it is not sufficient to include the words which appear in the collocations frequency chart. Indeed, some words which co-occur with high frequency polysemous words are themselves not extremely frequent and do not even appear in the collocations chart. But when these words appear within a nine-word span of the polysemous word, they determine the meaning. This is the case of *came*, which does not feature in the list of most frequent collocations for *break*, but when it appears within the nine word span, it is very likely it obliges *break* to mean *opportunity*: 'But the *break* for Southampton came in the 74th minute when Le Tissier sent former Arsenal reserve Neil Heaney off through the middle'.

Inevitably, representing the hypertext reality of a computer screen in two dimensions is rather demanding on the imagination of the reader. The columns in the grid of figure 17 should be displayed in separate windows or in boxes popping up at the click of a mouse. The way they would appear could depend on the user's personal taste. Apart from a translation, users would have the possibility of calling up a definition in the target language or a set of examples. However, in the case of *decoding* this would probably not be necessary. In the grid above it is understood that *decoder*s can stop their search whenever they think they have had sufficient information. Typing in *break* would give the user *quebrar* and *descanso*. If this did not make sense, the user could click on one of the words which they recognise in the context, e.g., *off*. Not yet satisfied with the

translation, the user could go on clicking on other words which would be part of her/his text and end up finding *começar*. As can be noticed for *break*, the translation of the word does not depend on the broader context, but only on the co-text. The translation of very polysemous words tends to depend on the words found in the traditional nine-word span. *Decoder*s do not know what are the words that change the core meaning of the hard word, and these must be suggested to them. Learners should not be asked to enter more words beyond the word they are looking up, they should be prompted to look for suggested words in the surroundings of their hard word if the core translation does not make sense.

The users themselves will decide on how much of all this will be visible on the first screen. They may think it more practical to have just the prompting words or concepts first and the mouse to click on them.

5.1.4 Including the Context. Two Experiments.

The task of the *decoding* dictionary is to give the user the right meaning of the hard word by whatever means, and as quickly as possible. In the preceding section I concentrated on the *co-text*. In the case of some words, nevertheless, the meaning of a word depends on the broader *context*. The translation of *hammer* can be deduced from its co-text in some cases, but it may be easier to proceed via the context: *martelo* if the text is on *metallurgy*, *percussor* if on *guns*, *martelo* again if on *pianos*. If the dictionary could establish what the general context of the hard word is, it would be able to suggest the most probable translation to the user right away. Looking up the meaning of a word would be greatly facilitated if the search area were not so broad. As things stand now,

one starts from a book which contains a fair part of the whole lexis of a language to find one word. But one can either make the pond smaller or the fish bigger. What follows is an account of two experiments aiming at trying to make the fish look bigger.

First experiment

In the first experiment, I asked a group of English students learning Spanish to read a number of short newspaper articles from the Catalan newspaper *El Mundo*. Their task was not to point out all the words they found hard to understand, but only those which they would look up in a normal reading situation, one in which they wanted to understand the message. The articles were on different subjects, ranging from international politics to computing. In addition, one student took on the first Chapter of *Don Quixote.* This gave a number of lists for every article with the words the students thought they would need to look up.

When I realised that these lists contained the information that was *new*, I was rather pleasantly surprised. Had the students failed to look up these words, they would certainly have understood what the topic was, but not what it was worthwhile writing about.

An telling example of this is the following article on the increasing number of thefts of computers. I put the questions to which the unknown words are an answer in the second column.

Hard word	Question it is an answer to
Ordenadores (3 times)	*What is being stolen?*
datos (de las compañías de seguros)	*Where does the information come from?*
Alrededor (de 208.000 ordenadores robados)	*How many?*
Fundas y bolsas	*How do the thieves go about it?*
suelen	
dueño (del ordenador)	*How do the thieves go about it?*
pertenencias	*What is being stolen?*
piezas metálicas	*How do the thieves go about it?*
se apodera (el ladrón)	*How do the thieves go about it?*
formar cola (en la línea del detector)	*How do the thieves go about it?*
sustraído (por el ladrón)	*How do the thieves go about it?*
funda	*How do the thieves go about?*

Figure 19 Unknown words in: 'El robo de ordenadores portátiles aumentó de forma significativa' (Article approximately 180 words long).

As can be deduced from this figure, most of the unknown words are an answer to the question 'How do thieves go about stealing a computer?'. This is the newsworthy part of the article, not the fact that there are ever more robberies, which is what everybody would expect. Seven of the twelve unknown words are directly related to this topic. This means that, allowing for some information obtained by induction from the context, this part of the article would be largely missed out by the learners if they did not have access to a dictionary.

In the case of the learner who took on the first chapter of *Don Quixote* (1900 words), the unknown words (app. 2.5 % of the entire text) are like a summary of the chapter and indeed of the book as a whole. Without them, the story is devoid of interest. This is the list of words which the student who read the first chapter would have to look up.

Hard word	What it refers to
astillero	*knighthood*
adarga	*knighthood*
rocín	*knighthood*
galgo corredor	*knighthood*
olla	*habits of Don Quixote*
vaca	*habits of Don Quixote*
carnero	*habits of Don Quixote*
salpicón	*habits of Don Quixote*
quebrantos	*habits of Don Quixote*
lentejas	*habits of Don Quixote*
palomino de añadidura	*habits of Don Quixote*
velarte	*habits of Don Quixote*
vellorí	*habits of Don Quixote*
ama	*who lives with Don Quixote*
frisaba	*appearance of Don Quixote*
recta	*appearance of Don Quixote*
enjuto de rostro	*appearance of Don Quixote*
desatino	*knighthood*
requiebros	*knighthood*
desafío	*knighthood*
fermosura	*knighthood*
desvelábase	*lack of sleep*
heridas	*knighthood*
rostro	*appearance*
darle fin al pie de la letra	*books*
melindroso	*knighthood*
no le iba en zaga	*knighthood*
turbio	*lack of sleep which typifies Don Quixote*
pendencias	*knighthood*
OTHERS: acabar; ensillaba; podadera; ratos; enflaquece; sacara; inacabable, competencia, muy acomodada situación.	

Figure 20 Words experienced as hard by the test subject in the first chapter of Don Quixote.

Basically this first chapter is about who Don Quixote is, what he looks like, what his habits are, who he lives with and his obsession with books and knighthood. Some of the unknown words in this list are of course not ascribable to any special feature of the *Quixote,* but the nice thing is that so many are. There are maybe no four words that characterise *Don Quixote* better than *pendencias* (fights), *desafíos* (challenges), *heridas* (wounds) and *requiebros* (flatteries addressed to women), all in the hard word list.

These words are extremely frequent in the *Quixote*, but are relatively rare in the language in general.

The conclusions to be drawn from this experiment are, first of all, that the difficult words of a text are responsible for the important part of its content. Second, that since the hard words are the less frequent ones, taking out the more frequent words leaves us with a summary of a text. This may seem obvious, since otherwise there would exist a limited number of texts, but it is something dictionary design could take into account. If the gist of a text resides in the words which are less frequent in the lexis of a language as a whole, the look-ups of a learner give an idea of this topic. This means that in the case of polysemous words, electronic dictionaries could guide the learner more directly to the meaning of the word if they had at least a rough idea of what the text was about. In the case of a polysemous word such as *drive*, it is unlikely the meaning should be anything else but *dirigir/unten suru/besturen/lenken/manejar/conduire* if the text is on *cars*. As a consequence, the learner need not know what *he drives me mad* means, unless the author of the text wanted to make a pun. This can also be illustrated by the word *key*.

Key can be translated in French by *clé; remontoir; clavette; touche; secret; légende; liste; solutions; corrigé; tonalité; capital; essentiel.* Depending on the situation, a learner will retrieve the right translation by means of a *collocation* –e.g. if *key* is followed by *issue*, preceded by *master, house, car*–, or a *morphological feature* since there is a fair chance that *keys* refers to *clefs.* However, the topic of the text might also help. If the text is on *music*, *key* should be understood as *tonalité*, if on clocks by *remontoir,* if on language textbooks, *corrigé*. A *label* could solve this problem, but what

if the reader has only a vague idea of the topic of the text, or if the label is not easy to tell? In this case, the list of hard words would give the (electronic) dictionary an idea of what the topic is and lead the learner to the right translation of a polysemous word.

Second experiment

To check this hypothesis on frequent and infrequent words, I took a series of articles published in the *Herald Tribune* and filtered out the 600 most frequent words, according to the *Bank of English*. This left me with skeletons of the articles which still had enough function words to be roughly understandable. (I reached the number of 600 words empirically.)

The experiment with the *Herald Tribune* articles confirmed my hypothesis that the information of a text lies in its infrequent words and that these are very frequent within the text itself. I will limit myself to one example, an article on *Desert Storm*[12]. In the figure below I list the words which were retained after taking out the 600 most frequent words. The original article consisted of 846 words, the sifted one of 638. I do of course mention each word or multi-word item only once.

Smart Arms In Gulf War Are Found Overrated

Pentagon's Reliance On High-Tech War Questioned in Review / Gulf / Pentagon dramatically oversold / effectiveness / expensive - tech aircraft / missiles, / thorough independent study / date / Pentagon / principal contractors claims / precision / impressive / weapons / Stealth fighter jet / Tomahawk land-attack missile / laser-guided "smart bombs" /overstated, misleading, inconsistent / data / unverifiable/ study / non-partisan / Accounting concluded /costly "smart" weapons systems / necessarily perform / Gulf / -fashioned / cheaper "dumb" ones. / wisdom / plans / depend increasingly / weapons / extend / art / tens / billions / dollars.

Figure 21 Infrequent words in Herald Tribune *article on* Desert Storm.

[12] The full text can be found in Appendix 4.

This experiment supports the hypothesis that the new information of a text lies in the less frequent words. It also suggests, as is natural, that this new information does not lie in the function words which are all included in the list of the 600 most frequent words[13]. Furthermore, it is noteworthy that many of the *collocations* have not been affected by leaving out these 600 most frequent words (*depend increasingly, costly 'smart' weapons systems, laser-guided 'smart bombs', Stealth fighter jet, principal contractors claims, thorough independent study*). In other words, the structure of the text as a construction of chunks was left greatly intact.

Finally, a secondary result of this procedure is that we are now in possession of an important portion of the semantic field of the topic of the article, *modern warfare,* including the *collocations*. Here is a list of these words extracted from the same article on *Desert Storm*.

Effectiveness, expensive, high-tech aircraft, missiles, precision weapons, Stealth Fighter jet, Tomahawk land-attack missile, laser guided 'smart bombs', costly 'smart' weapons systems, perform, cheaper 'dumb' ones, plans, depend increasingly, Congress, secret, conducted, Operation Desert Storm, Defence Department, commanders, intelligence, pilots, planners, battlefield, unclassified summary, secret, scheduled, figures, videotapes, shaft, extraordinary accuracy, Tomahawk missiles launched, precision-guided munitions, aircraft, carry, assess, effectiveness, bombing campaign, improved smart weapons, weaponry, arsenal, ways, locate, destroy targets, overwhelmed, deployed nearly 1,000 combat aircraft, unleashed nearly, tons, bombs, dropped, superior technology, presumed target, tank, truck, destroyed, sensors, laser, electro-optical, infrared systems, clouds, rain, fog, smoke, humidity, sleek F-117 Stealth fighter jet, highly touted ability, target evading detection, necessarily outperform older, cheaper aircraft, claimed, 80 percent success, bombing runs, fighter, inappropriate, aircraft, effectiveness demonstrated, higher- aircraft, generally, capable, lower- aircraft, bomb tonnage dropped, guided munitions, accounted, munitions, campaign data, validate, purported efficiency, guided munitions, target, smart bombs built, guided munitions, praise, Pentagon weaponry, Patriot missile, nearly perfect, shooting, missiles, aimed.

Figure 22 Semantic field of 'modern warfare' extracted from a Herald Tribune *article on 'Desert Storm'.*

[13] Literature teachers may claim that very important information lies enclosed in these function words and in the way they are combined, and even in a newspaper article these very frequent words inform us about the hidden agenda. But in this case we are dealing with problems which precede these considerations. Undoubtedly, very frequent words are the first thing one learns in a foreign language, then come the less frequent ones, and eventually one has to start all over again and learn the true significance of the most frequent words again. However, the problems I am trying to deal with regard the content of messages communicating states of affairs.

Naturally, this list should be completed and adapted manually, but even this rather crude catalogue shows the possibilities of the technique, enabling one to draw maps of semantic fields, taking authentic texts as a starting point. These semantic fields, together with their translations, provide us with a syntagmatic –in some sense 'onomasiological'– map of the language instead of a more traditional one, in which a word is related to its synonyms and not to the words it is likely to occur with in a normal text[14]. In a text containing the words *aircraft, bombing, guns*, the meaning of *drop* would probably be related to *bombs* and thus the French translation would be *larguer* and not *baisser*; in Portuguese, *largar* and not *abaixar*. Similarly, *rounds* will mean *cartucho* in Portuguese and not *redondos*. As soon as a user had looked up a handful of words, the electronic dictionary would be able to identify the semantic field and direct all further searches accordingly. In the case of *polysemy*, this means that the dictionary programme would go first to the meaning most likely to appear in a text on *gardening, football,* etc. Fields would not have to be formally delineated. It would be sufficient to link words or items and their translation: if in the same text an item **x** is proven to have a translation **x1** and an item **y** a translation **y1**, than the translation of **z** is likely to be **z1** as opposed to **z2**, **z3** etc. [15]

[14] It is this kind of consideration which induces some lexicographers to apply this classification to the whole dictionary. As I mentioned in the section on pedagogical dictionaries, this classification has the advantage of being more rational, but less practical, at least in the case of paper dictionaries. As Jean Binon puts it: 'De nombreux arguments militent en faveur d'un classement onomasiologique lorsqu'il s'agit d'*encoder*, de produire, de passer à la mise en discours. En effet pour le décodage la consultation d'un dictionnaire a lieu au gré des lectures et des occurrences. Lors de la production en revanche on part d'une intention de communication, d'une idée, on aborde un thème. On ne parle, on n'écrit pas alphabétiquement. Il faut donc que le lexique soit organisé de façon sémantique, que l'on associe par exemple 'offre' et 'demande', même si ces deux mots se trouvent assez éloignés dans l'ordre alphabétique. Il va cependant sans dire qu'il faut prévoir un index alphabétique avec la traduction des mots, des renvois pour faciliter l'accès'. (Binon, 1995, no page number, electronic version)

[15] Just one practical application of this. Journals and other magazines are more and more being published on the Net. This means texts come to an increasing extent to the reader in machine readable form. An electronic dictionary could sift the 600 most common words out of an article, deduce from what is left from what the topic of the article is, and sort the order of the senses the words, likely to feature in texts on

In an electronic dictionary, all the lexical items can be classified alphabetically with their various meanings one after the other (lead, *music*; *boxing*; *nautical*; etc.), but also within these bigger groupings themselves, together with the words with which they normally occur. In current dictionaries, the various meanings of a word are listed together, but it is highly unlikely that the same word will be used in more than one sense in the same text. It is when the text ends that another sense can be chosen for the same word *(lead* refers to the *mineral* in an economic survey, and *to play first violin* in an article on music). In an electronic dictionary a polysemous word can be stored in both classification systems, in the alphabetical as well as in the semantic grouping that gives it its sense.

So far the *decoding* part of the dictionary. In what follows I will outline what I think the *encoding* part of a dictionary should be like.

---

this topic, can have in order to put the most likely alternative first.

5.2. ENCODING

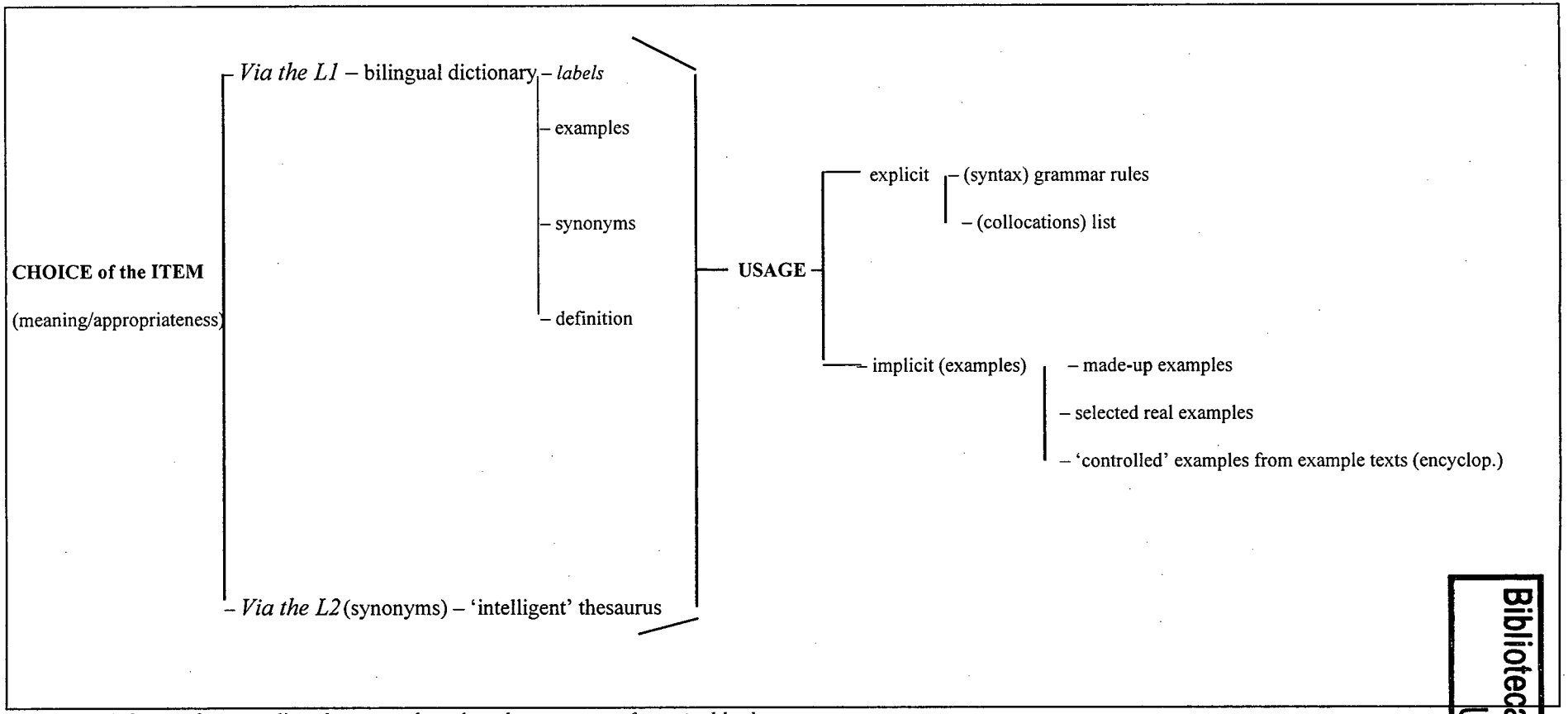

Figure 23 Scheme of an encoding *dictionary based on the sequence of a typical look-up.*

Introduction

As was the case for *decoding*, in tackling the problem of *encoding* it is necessary to make a distinction between *beginning* and *advanced* learners. From an *encoding* point of view, a distinction is traditionally made between beginners, intermediate and advanced.[16] For the compiling of a dictionary, however, I would suggest that there is no need for that much differentiation. It is sufficient to distinguish between beginners and more advanced learners. In the first part of this chapter, I will tackle a few problems which I think are specific for beginners. In the second part, I will tackle the needs of advanced learners. When *encoding,* one can have the following kinds of problems:

1. one does not have any idea of what the word is in the target language;
2. one knows the word superficially and would recognise it, but cannot recall it actively;
3. one knows the word but is unsure of its collocates or its syntax.

All three problems will occasionally trouble any foreign language learner of whatever level. However, the first kind of difficulty will be more characteristic of a beginner's level and for the design of a dictionary it will have specific consequences. The second and third kind of situation are more typical of an advanced level and will be dealt with separately.

[16] Although this distinction is well established, it has been poorly described and is generally taken for granted. According to my own experience as a language teacher and learner, this distinction exists. At a beginners level, learners are concerned with getting simple messages across on everyday matters. They are aware of the limitations placed on their communication capacities by their lack of knowledge and their queries aim at diminishing this lack. At an intermediate level, learners can express anything they want by means of congruent language, to use Halliday's nomenclature. They only use grammatical metaphors sporadically. Seen from a different angle and in accordance with Sinclair's terminology, intermediate learners are able to express themselves well grammatically, but not naturally. Advanced learners aim at a native like fluency and they use idioms, grammatical metaphors and pragmatically appropriate language.

5.2.1. Encoding for Beginners

I will illustrate the problems of beginning language learners by means of two examples in Japanese. Japanese dictionaries are specially suited for this kind of demonstration since any Westerner who has a command of only European languages is a true beginner in Japanese. Lexicographers working on Japanese dictionaries for foreigners clearly conceive their dictionaries with true beginners in mind. The Japanese writing system leaves no doubt as to the intended audience of each dictionary. It is therefore easy to choose precisely those designed for learners of Japanese, and not of English. Indeed, dictionaries using *kanji*, Chinese characters, are intended for Japanese or very advanced learners, whereas those written in *romaji*, the romanised alphabet, are designed for a non-Japanese audience. As for dictionaries written in *hiragana*, they can be considered to be half way between *kanji* and *romaji*, but are in practice only a way of simplifying the localisation of words for Japanese, since the *Introductions* to these dictionaries are written in Japanese only.

First test sentence

In order to make my point, I chose to translate a typical beginner's sentence: *The flowers are on the table*. Beginning learners of Japanese will at least know that there are no articles in Japanese, no plurals, no verb inflections which depend on the personal pronoun, that the word order is the inverse of the English, and that the theme of the sentence is indicated by the particle *wa*. This gives the following basic pattern to be translated: *Flower-wa table on be*.

Martin's Concise Japanese Dictionary is a small dictionary with an English-Japanese and a Japanese-English part. It uses *kanji, hiragana, katakana,* and roman characters. Using this dictionary allows us to translate our sentence into: *Hana wa teburu* **on** *desu.* The dictionary entry for *on* is:

on ... de ...で, *(located)* ... ni ... に; *(atop)* ... no ue (de/ni) ... の上(で/に)

Needless to say this is not an entry which is particularly helpful for our sentence, since there is no indication of the differences between these different possible translations of *on*. Then who is this information intended for? The *romaji* part was necessarily intended for a non-Japanese audience. The *kanji* character for *ue*, on the other hand, can only be intended for Japanese. For these users, however, the *romaji* information is useless. If this information is meant for Japanese learners *decoding* an English text and not knowing what *on* meant, then the dictionary should give these three possible translations in *kanji* or *hiragana*, not in a Western writing system.

If this information is instead intended for English learners of Japanese, then they must be capable of distinguishing between these three possibilities for *on*, which is not exactly beginners' knowledge. The dictionary tells them there are three possibilities and only makes them aware of the fact that they are indeed incapable of telling the difference. Not a pedagogically very refreshing experience.

If, on the other hand, one is advanced enough to know how to use these different *on*'s in Japanese, one will most probably remember them without a dictionary, and so this information is not for advanced learners either. If one feels unsure about the usage

on [ɑ:n] *prep* **1** (indicating position) ...(の上)に〔で〕...(no uḗ) ni〔de〕
on the wall 壁に kabḗ ni
it's on the table テーブル(の上)にあります tḗburu (no uḗ) nì arímasù
on the left 左に hidári nì
the house is on the main road 家は幹線道路に面しています iḗ wà kańsendṑro ni méń shite imásù
2 (indicating means, method, condition etc) ...で ...dè
on foot (go, be) 歩いて arúíte
on the train/plane (go) 電車〔飛行機〕で déńsha〔hikṓkì〕de; (be) 電車〔飛行機〕に乗って déńsha〔hikṓkì〕ni notté
on the telephone/radio/television 電話〔ラジオ，テレビ〕で déńwa〔rájìo, térèbi〕de
she's on the telephone 彼女は電話に出ています〔電話中です〕 kánòjo wa déńwa ni détè imasu〔deńwachū desù〕
I heard it on the radio/saw him on television 私はラジオで聞きました〔テレビで彼を見ました〕 watákushi wà rájìo de kikímashìta〔térèbi de kárè wo mimáshìta〕
to be on drugs 麻薬をやっている mayáku wò yatté irù
to be on holiday 休暇中である kyṹkachū de arù
to be away on business 商用で出掛けている shṓyō dè dekákete irù
3 (referring to time) ...に ...ni
on Friday 金曜日に kiń-yṓbì ni
on Fridays 金曜日に kiń-yṓbi ni, 毎週金曜日に maíshū kiń-yṓbi ni, 金曜日毎に kiń-yṓbi gótò ni
on June 20th 6月20日に rokúgatsu hatsúka ni
on Friday, June 20th 6月20日金曜日に rokúgàtsu hatsúka kiń-yṓbi ni
a week on Friday 来週の金曜日に raíshū nò kiń-yṓbi ni
on arrival he went straight to his hotel 到着すると彼は真っ直ぐにホテルへ行きました tṓchaku suru tò kárè wa massúgù ni hṓtèru e ikímashìta
on seeing this これを見ると koré wò mírù to
4 (about, concerning) ...について ...ni tsúíte, ...に関して ...ni káń shite
information on train services 列車に関する情報 resshá nì kań surù jṓhō
a book on physics 物理の本 bútsùri no hóń
♦*adv* **1** (referring to dress) 身につけて mi nì tsukète
to have one's coat on コートを着ている kṓto wo kité irù
what's she got on? 彼女は何を着ていますか kánòjo wa nánì wo kité imasù ká
she put her boots/gloves/hat on 彼女はブーツを履いた〔手袋をはめた，帽子をかぶった〕 kánòjo wa bṹtsu wo haíta〔tebúkuro wò haméta, bṓshì wo kabútta〕
2 (referring to covering): ***screw the lid on tightly*** ふたをしっかり締めて下さい futá wò shikkárì shímète kudásaì
3 (further, continuously) 続けて tsuzúkètè
to walk/drive/go on 歩き〔車で走り，行き〕続ける arúkì〔kurúma dè hashíri, ikí〕tsuzukèru
to read on 読み続ける yomítsuzukèru
♦*adj* **1** (functioning, in operation: machine) 動いている ugóite irù; (: radio, TV, light) ついている tsúíte iru; (: faucet) 水が出ている mizú gà deté irù; (: brakes) かかっている kakátte irù; (: meeting) 続いている tsuzúite irù
is the meeting still on? (in progress) まだ会議中ですか mádà kaígichū desù ká; (not cancelled) 会議は予定通りにやるんですか káìgi wa yotéi dṓri ni yarún desù ká
there's a good film on at the cinema 映画館で今いい映画をやっています eígakàn de ímà íì eíga wò yatté imasù
2: ***that's not on!*** (*inf*: of behavior) それはいけません soré wà ikémaseń

of each of them, only examples or a grammatical explanation can help. None of these exist in the entry.

I suspect this is another example of the *emblematic* information all dictionaries include to a certain extent. Nobody can say that the information is not included. Nobody, on the other hand, can tell what it is good for.

Collins Shubun tackles the problem of function words like *on* in a different way. In the concrete case of *on*, a whole section is dedicated to its various senses with examples to help beginners with their choice. Three hundred 'keywords', function words and other very frequent words like 'have' and 'be', were in this way given an 'extended treatment'. In the case of our test sentence, *The flowers are on the table*, *Collins Shubun* gives the example *It's on the table* (teburu ni arimasu) which immediately reminds us of our test sentence. The translation is given both in *romaji* and in Japanese characters, although it is not entirely clear what the *kanji* characters are for. Since the example sentences do not include fixed expressions or idioms, they were obviously intended for non-Japanese who, having to look up basic

open 1. opens it akema͡su (akeru, akete) 開けます(開ける，開けて), *(opens it up)* hirakima͡su (hira͡ku, hira͡ite) 開きます(開く，開いて); *(begins it)* hajimema͡su (hajimeru, hajimete) 始めます(始める，始めて) **2. it opens** akima͡su (aku, aite) 開きます(開く，開いて); *(it begins)* hajimarima͡su (hajimaru, hajimatte) 始まります(始まる，始まって), *(a place, an event)* o͡pun shima(゛)su (suru, shite) オープンします(する，して), kaijō shima͡su (suru, shite) 開場します(する，して), *(a shop/business)* kaiten shima͡su (suru, shite) 開店します(する，して)

information of this kind, will not be able to decipher the *kanji* anyway. One could argue that they make the user accustomed to these characters and add to the Japanese flavour. More likely, however, they are to convince Japanese learners of English that this dictionary was also intended for them

Going back to our test-sentence, thanks to *Collins-Shubun*, the first tentative translation can be corrected into *Hana wa teburu ni arimasu.* However, learners could still doubt if *arimasu* is in this case the right verb. This is where the romanised *Martin's* Japanese-English part of the dictionary comes in handy. Since it mentions 'it is (located)' as one of the possible translations of *arimasu*, it confirms the information retrieved from *Collins Shubun.*

Second test sentence

Another Japanese example illustrating the dictionary problems of beginners is the sentence *Is the bank open? Martin's* provides the translation for *bank*: *ginko*. *Open* is more of a problem.

open [ou'pən] *adj* (not shut: window, door, mouth etc) 開いた aíta; (: shop, museum etc) 営業中の eígyōchū no, 開いている aíte iru; (unobstructed: road) 開通している kaítsū shite iru; (: view) 開けた hiráketa; (not enclosed: land) 囲いのない kakói no nai; (*fig*: frank: person, manner, face) 率直な sótchoku na; (unrestricted: meeting, debate, championship) 公開の kókai no
♦*vt* 開ける akéru, 開く hiráku
♦*vi* (flower, eyes, door, shop) 開く akú, 開く hiráku; (book, debate etc: commence) 始まる hajímaru
in the open (air) 野外に yagái ni
an open car オープンカー ōpūnkā

Martin's entry for *open* has two sub-sections, one headed by the expression *opens it*, the other by *it opens*, a way of avoiding the use of *transitive* and *intransitive*. A further subdivision leads to: *(a shop/business)* kaiten shimasu (suru, shite).

This entry induces learners to believe that the fact that something *is open* should be expressed by a verb and a complement and not by an adjective, since the dictionary does not show this possibility.

According to *Collins Shubun*, however, *open* can be translated by an adjective as well as by a transitive or an intransitive verb. In the case of *shops* and *museums*, the suitable form, says the dictionary, is *eigyochu no* or *aite iru.* No indication is given as to the difference between these two possibilities, which makes the information anything but helpful. Since what we want to ask is not if *the bank opens* but if it *is* open, one can only venture to follow *Collins Shubun* which results in: *Ginko wa eigyochu no desu ka* or *Ginko wa aite imasu ka.*

In order to confirm this, we go back to the Japanese-English part of *Martin's* which tells us that *eigyo* means *(running a) business*! This makes one suspect that *eigyochu no* does not apply to our case. *Aite*, on the other hand, appears to mean *coming open* and *iru* is the dictionary form of *imasu* (*is* or *stays*). This gives a new translation: *Ginko wa aite imasu ka* which is indeed the correct translation.

All this shows a typical beginner's problem: uncertainty as to what has to be looked up. This is not so much of a difficulty if one stays within the same family of languages, but it might be very intricate indeed if one is accustomed to, say, inflectional languages, and studies an agglutinative language for the first time. The main problem I had with translating *Is the bank open?* came down to knowing in what way *open* was lexicalised in Japanese. This shows clearly the fundamental importance of simple examples for the beginning learner.

Function words

A first conclusion that can be drawn from these experiences is that a beginning learner has to go back and forth a number of times to retrieve the information necessary to solve, all things considered, a rather straightforward problem. At *Martin's,* not enough research was done as to how the dictionary was practically going to be used.

Secondly, if we look at these experimental look-ups, it appears that all beginners' problems have some relationship with grammar: grammar at a macro level –the *verb* system in Japanese–, or at a micro level –function words. In terms of micro level, *Collins Shubun* is admirable in its treatment of 300 keywords. This is particularly useful for beginning learners who will seldom *recognise* the target item and will often learn it through a dictionary. In the case of *as*, the examples are as follows (in the dictionary these examples are followed by their translation in Japanese, both in *kanji* and *romaji*).

1. (referring to time) *as the years went by; he came in as I was leaving; as from tomorrow*
2. (in comparisons) *as big as; as much/many as; as much money/ many books as; as soon as*
3. (since, because) *as you can't come, I'll go without you; he left early as he had to be home by 10*
4. (referring to manner) *do as you wish; as she said*
5. (concerning) *as for/to that*
6. *he looked as if he was ill*
7. (in the capacity of) *he works as a driver; as chairman of the company, he...; he gave it to me as a present*

Figure 24 As *in* Collins Shubun English-Japanese Dictionary.

The examples contain only a few fixed expressions, showing that this dictionary actually targets English speaking beginners. There is indeed not much chance that a Japanese learner will look up (to *decode*) a sentence like *he came in as I was leaving,* since it is perfectly understandable if one knows what *he, came, in, as, I, was,* and

leaving mean. The example phrases are almost exclusively of the open-choice type and allow dictionary users to fill in the pattern with the words of their own sentence. The adding of *labels* should in this particular case of *as* be considered superfluous, since they are unavoidably more abstract than the examples, which are in themselves clear enough.

The separate grammar part in *Collins Shubun* at the end of the volume is clear, but this being a paper dictionary, the information was not included at the appropriate spot. This dictionary could have been an ideal beginners' paper dictionary had the lexicographers been allowed to concentrate on the English speaking audience. The present dictionary contains 30.000 headwords, far too many for (*encoding*) learners of Japanese and not enough for the (*decoding*) Japanese learners of English. This results in a lack of examples and poor treatment of very frequent and difficult to use words such as *think*. Unfortunately, *Collins Shubun* aimed at being 'useful to the language student and the native speaker alike' and as a result a compromise had to be made. This had as a result that an excellent concept was carried out only partially.

Conclusion

Beginning learners have modest aims when *encoding* into a language they have just started to study. This means they have modest demands in terms of content words and a pressing need for information on function words. When choosing between several alternatives for a word, they will proceed by comparing their sentence in their L1 to the ones listed under the form of examples. Examples, particularly for function words, are in this kind of dictionary of the utmost importance since an *encoding* learner does not

always have a clear idea of what is the item to be looked up.

The analysis of two Japanese dictionaries for beginners suggests that even in cases in which it is obvious that a choice should be made as to the audience, commercial considerations get in the way. In an attempt to satisfy both *encoders* and *decoders*, clever lexicographic projects are imperfectly carried out.

5.2.2. Encoding for Advanced Learners

Ideally, advanced learners have no morphological problems, want to use grammatical metaphors, the correct collocates and, from a pragmatic point of view, the appropriate expressions. In addition, the last stage in language learning is characterised not only by the use of the appropriate word, but also by the use of the appropriate sequence of words, i.e. *idioms* and *fixed expressions*. Advanced learners aim at a native-like command of the foreign language and the question is what a dictionary should be like in order to help them with this.

Basically, the look-up process can be subdivided into the *choice* of the item and looking up information on its *usage*. These stages will normally follow each other chronologically and in my argument I assume they do. However, the possibility should not be excluded that a learner accesses the dictionary at any of these stages. In other words, learners may skip the *choice* stage if they are only interested in *usage*.

Since I start from the presupposition that the dictionary of the future will be computer-based, random access to any of its component parts is taken for granted. Since I have not found much literature on this subject, what follows draws basically on my

own experience as a dictionary user and the experience of fellow dictionary users whom I have observed and interrogated over the last few years. I myself wrote down a number of my own look-ups in as many languages as I could, in some way, handle. These include Dutch, French, Spanish, Portuguese, English, Italian, German, Japanese, Latin, modern and ancient Greek, Danish. These are languages in which I am a native speaker, advanced, intermediate, false beginner, absolute beginner, or almost ignorant, according to the case.

5.2.2.1. Choice of the Item

The process of choosing the appropriate equivalent in the target language will usually start by looking up the item in a bilingual dictionary, i.e., go via the source language, and this will be the subject of the first section of this chapter. Users who have already an idea of the item have the possibility of proceeding via the target language and refining their choice. I will deal with this second possibility in the second section of this chapter.

5.2.2.1.1. Via the Source Language

When proceeding via the source language, i.e., through a *bilingual* dictionary, the odds are that the learner will be confronted with more than one option. The ways to distinguish between these various options are basically four: *synonyms, labels, examples* and *definitions*. Although all four have advantages and disadvantages, at least one is suited for each particular kind of item. In some cases a learner will be helped out by a *synonym* in the source language, in others, by an *example*. To be effective, the dictionary should propose the most adequate option first, even if the lexicographers choose to offer more than one.

The moment dictionary users have localised their item in the dictionary, there is a stage of *decoding* within the operation of *encoding*, a stage for which are valid most of the observations I formulated for *decoding*. The difference here is that *encoding* is, as a whole, a much more precise activity. If the word to *encode* is *avaliar* and the dictionary suggests *appraise, evaluate, value* (Taylor, *Portuguese-English Dictionary*. 1963. Record, Rio de Janeiro), *encoders* will have to *decode* these translations to see which one fits their context. Since in a dictionary the particular context of the user is of course not included –while *decoding* users have their own text–, what the dictionary has to do is to provide one, of whatever kind, in order to enable users to make their choice. This is what *examples* and *definitions* and, in a broader sense, *synonyms* and *labels* are for. They all attempt to connect the unknown word with something the learner knows.

The information for the dictionary user can be of two kinds: *abstract* and *concrete*. The two extremes of the spectrum are occupied by *labels* and by *examples*. Although not always, *labels* are often more abstract than the looked-up word when they give information on the semantic field to which the word belongs. *Labels* themselves can be of various degrees of abstraction. Examples, on the other hand, give concrete information by providing users with a type of context which they will or will not recognise as the one they had in mind.

In order to distinguish between different lexical items from an *encoding* point of view, the distinction between *function* and *content* words is again of some utility. Generally speaking, the meaning of a content word tends to depend on the context, while the meaning of a function word depends on the co-text. In English, particular function words, such as pronouns, are generally quite straightforward. Others, such as

prepositions, are highly complex and can be translated in a variety of ways, depending on the words which immediately follow.

In dictionary terms, this means that the correct translation of a function word can be decided on by putting it in an *example* followed by its translation. The translation of a content word, on the other hand, is decided on by referring to a semantic field, or a *synonym* in the source language, in a *label*. In *Collins Shubun*, the function word *at* has a different translation depending on whether the co-text is: *at the top, at school, to look at something, at 4 o'clock*, etc. The *labels* which in this case accompany the examples (referring to *position; time; rates*, etc.) are superfluous. As for a content word like French *énergie*, it can be translated in several different ways in English, but the *synonyms* used in *Collins-Robert (force physique; fermeté, ressort moral)* go a long way towards disambiguating the possible translation options.

Clearly, each item should be examined in itself but, as a general rule, sense disambiguation of function words can be attempted fruitfully by means of *examples*, and of content words by means of *labels*. In particular cases, a *definition* can help by stressing specific features of a word as opposed to its *synonyms*. However, this is at present rarely done. *Synonyms* in the source language are often useful, but since current dictionaries have four possible audiences, two *encoding* ones and two *decoding* ones, it is not always clear in which language the *synonyms* should figure in the dictionary.

Another way of looking at the problem is by means of the *frequency* of the item. Allowing for a number of exceptions, one can say that the more frequent a word is, the greater the chances are that it can be translated in many different ways[17]. In addition, the

[17] This is not always the case, as *I* and *you*, onomatopoeias (*cliquetis, rattle*) and others show.

number of possible translations of a word depends of course on the target language. Linguistically, and culturally, related languages have a greater likelihood of being partially isomorphic than others.

Labels

I will call a *label* anything between brackets giving an indication of the *meaning* or the *pragmatic appropriateness* of a word. Although extremely common, not much has been written on *labels*[18] and I know of no publication which considers the matter for the specific case of foreign language lexicography. *Labels* of the *meaning* type situate the item in a semantic field: directly (*maritime, agricultural, medical*), or by means of another word which refers to it. This last modality –as in *régler* (mécanisme), *regulate*, where the *label* refers to *machinery*– has some characteristics of a disambiguation by means of an example. It is an example of the type learners, as it were, have to 'assemble' themselves by using the *label* word in a made-up sentence. *Labels* of the *pragmatic* type indicate the word's register by naming it: *slang, formal, archaic*, etc. and are rather straightforward.

Labels can be of various kinds: the already mentioned reference to a semantic field; a verb which can be used with the target noun; a noun which can be used with the target verb; a noun to which the target adjective can apply. The *Oxford Spanish Dictionary* uses, among others, the following *labels*: (naut.); (astrol.); (aviat.). These *labels* are very effective at disambiguating the translation of a *content word.* The same technique would be useless if applied to a *function word.* In the case of function word *at*, the same dictionary separates the various senses by means of the following *labels*:

[18] Hartmann 1981, 1983; and Osselton, 1996 discuss related areas.

indicating location, position; indicating direction; indicating time; indicating state; occupied with; with measurements, numbers, rates, etc.; with superlative; because of; concerning. However, given the fact that *at* is a function word, one has the impression the labels here only function as a way of classifying the high number of examples in the entry. It might be that in this and similar cases, classification by means of *labels* has no other purpose than to classify in some way or another, since most users will probably go straight to the examples. As I will maintain when I deal with *examples* as instruments for disambiguating translation options, I think that in cases like *at* examples are more helpful than *labels.*

In the *Oxford Spanish Dictionary*, the two possible translations of *construir (build* and *construct)* were distinguished by means of the following *labels*:

a ‹ edificio/barco/puente › to build
b ‹ figura geométrica › to construct
c ‹ frases/oraciones › to construct
d ‹ sociedad/mundo › to build; construir un nuevo mundo *to build a new world*

Figure 25 Construir *in the Oxford Spanish Dictionary.*

In this case, the user was directed to four alternatives by means of an indication of the semantic field. Seen from a grammatical point of view, the *labels* here are the direct objects of the translational equivalents in question, half way towards an example. One can *construir un edificio, construir una figura geométrica*, etc.

These are a few of the many other kinds of *labels* found in the *Oxford Spanish Dictionary,* and which are familiar to most people who have ever dealt with dictionaries.

> (Teatr) (Ven) (AmE) (BrE) (Ven fam) (Ven arg) (Méx) (Dep) (AmS vulg) (AmE sl) (BrE sl) (colloq) (crit). (para el pelo) (de un cangrejo) (en costura) (para la ropa)

Some of these *labels* refer to a geographical area, some to what the word applies to, and some indicate the level of formality. They orient the user not only to the right translation in terms of *meaning*, but also of *appropriateness*. In this case, *labels* which are intended to help with *decoding* have been mixed with labels intended for *encoding*. Indications such as 'de un árbol', 'del queso', *arg.* (argot), are intended for *encoding* Spanish speakers, which is the reason why they are in Spanish. Strangely enough, the indications *sl* (slang), *colloq* (colloquial), and *crit* (criticised usage) are in English, although English speakers of the kind who use dictionaries should know if the English word is *slang* or *formal*. This is an example of one type of incoherence to which the mixing up of audiences inevitably leads.

Concrete and Abstract

As I stated, *labels* can work in two directions, *concrete* and *abstract*. Whenever possible, concretising should be preferred. I will illustrate this *ex absurdo* by means of an example of the contrary.

In the entry for *oreille* in *Collins-Robert* the *labels* are, among others: *anatomie; ouïe; comme organe de communication.* They are almost the same as in *Oxford-Hachette*: *anatomie, ouïe, personne.* Unfortunately, these *labels* look somewhat like a scholastic exercise. One should ask oneself if users in search of a translation for *oreille* will start asking themselves if in their case *oreille* refers to something *anatomical, a faculty of hearing*, or an *organ of communication.* Any disambiguation which asks from the user a supplementary effort of abstraction contradicts the dictionary's aim of being a mere tool, as practical as possible. When users are in search of an item, they look for

something they recognise which is connected to what they want to express, mostly a word related to the field they are dealing with. A practical approach which gives the user supplementary information, of any kind, should therefore prevail and publishers should let lexicographers to be guided more by Hermann Rorschach than by Saint Thomas Aquinas. The entry for *demander* (in its transitive sense) in the *Oxford-Hachette* gives an example of a pragmatic approach. Any reference which could 'ring a bell' has been used:

1. (solliciter) [conseil, argent, aide, permission]; (dans une offre d'emploi)
2.(enjoindre)
3. (souhaiter)
4. (interroger sur)
5. (faire venir) [médecin, prêtre]; (dans son bureau) (au téléphone)
6. (nécessiter) [travail, tâche] [effort, attention, qualification]; [plante, animal]. [attention]; [sujet, texte].
7. droit [tribunal] [peine, expertise]; [personne] [divorce]; [dommages-intérêts];

Figure 26 Demander *in* Oxford-Hachette.

Here the user finds: *synonyms*, which I will expand on in the next section; the *object* of what can be *asked for* (conseil, médecin, etc.); and the *circumstances* (*tribunal*). Remarkably, the type of *labels* has been adapted to the item in question. The lexicographers were not guided by any kind of orthodoxy, obliging them to apply the same rule to each and every word[19]. I do not think orthodoxy agrees well with the diversity of lexis.

Synonyms

Another way of differentiating the possible translations of an item is through *synonyms*

[19] This kind of approach should inform all dictionary making and should go as far as adapting the treatment of a word to the language it is translated to. Here, as in all other aspects of dictionary making, contrastive information should be integrated into the dictionary.

in the source language. Since *synonyms* used for differentiating senses have to be superordinates, and if the word is itself abstract and a verb, this may require a considerable effort of abstraction from the user.

The entry for *demander* in *Oxford-Hachette* lists the following *synonyms*: *solliciter; enjoindre; souhaiter; interroger sur; faire venir; nécessiter*. The limits between these distinctions are in some cases tenuous, and dictionary users will be hard pressed to decide in which category their intended meaning falls.

Synonyms are quick and efficient ways of distinguishing senses in cases in which these senses are as clearly distinct as in *table: piece of furniture; list*. They are difficult to use in the case of subtle differences. The question whether synonyms have some utility should be asked in each case, instead of applying them to every entry as a rule.

This brings us back to the phenomenon of internal and external polysemy. In cases in which native speakers are aware that a word has more than one meaning –*table (and chairs)* vs. *(time) table*– translation disambiguation by means of *synonyms* will be successful. This will not be so easy in the case of external polysemy –*quelque chose* as *anything* or *something*– in which native speakers are generally not aware of the different senses of the word.

Examples

Examples can direct learners towards a particular translation by putting the word in a co-text which they can compare to theirs. Since I spent a few pages analysing the problem of examples in Chapter 4, I will here limit myself to highlighting the main features of their disambiguating function. As I argued before, in the case of beginning

*encoder*s, examples are vital and should consist of frequent phrases, or sentences with frequent words. In the case of *advanced* learners, however, the situation is different.

In the following list I give a random sample of words with their frequency (according to COBUILD 2). Between brackets are the number of senses listed for each in the dictionary.

no diamond[20]	impudent (1)	dispersion (1)	ellipse (1)	shirk (1)	seducer (1)
1 diamond	terse (1)	mod (1)	raven (2)	pristine (1)	orchard (1)
2 diamonds	ratio (1)	carve (5)	covering (1)	justified (2)	polite (2)
3 diamonds	justify (1)	massive (2)	neutral (9)	pipe (7)	prior (4)
4 diamonds	finger (16)	indicate (6)	liberal (4)	movie (2)	print (20)
5 diamonds	form (20)	half (16)	go (49)	nothing (18)	test (13)

Figure 27 Illustrative table of a few word frequencies indicated in number of diamonds, and (between brackets) number of senses.

As this list suggests, internal polysemy does not necessarily mean external polysemy, but it is a strong indication that a word has more than one translation. To distinguish between the various translations of *form* in Portuguese (20 senses in COBUILD2; 53 possible translations in Houaiss A. *Dicionário Inglês-Português*), examples would be very effective since the meaning of *form* will depend to a large extent on the co-text. However, in the case of *raven* (two possible translations: 'corvo' and 'cabelo preto'), examples might help but will be far less effective than a *label*. Here again the link between audience and content of a dictionary emerges clearly: beginners need the translation of frequent words differentiated by means of examples; advanced learners need infrequent words differentiated by means of *labels* or *synonyms*[21].

[20] In COBUILD2 'diamonds' indicate the frequency of a word; the more 'diamonds' the more frequent the word.

[21] This of course regards the differentiation of translation options and does not mean advanced learners do not need examples. I stated this clearly in Chapter 4.

The capacity of examples to distinguish between different meanings of words is directly proportional to their frequency. Indeed, dependency on the co-text to indicate the meaning of the word is directly proportional to its possibility of appearing in an idiom or a fixed expression. The more frequent a word is, the more it tends to rely on the co-text to acquire meaning and the greater is the probability that an example will be able to give an idea of its meaning. The less frequent a word is, the more it has a meaning independently from the co-text, and even the context. The more the meaning of a word depends on the co-text, the more examples will help to distinguish different options.

Definitions

A fourth possibility of distinguishing between various translation options is by means of a *definition*. This is the wordiest way of proceeding and should therefore be reserved for particularly intricate cases such as words for which there exists only one equivalent in the source language. Spanish *cambiar* can be translated by *alter* or *change*; *sustituir* can be translated by *substitute* and *replace*. In these cases the difference has to be *explained* and cannot only be hinted at with a *label* or a *synonym* in the source language[22]. Japanese college dictionaries have been doing this for many years and the *Cambridge Word Selectors* does something similar.

In practice, advanced learners get used to looking up a word in a bilingual dictionary and then looking for additional information in a monolingual dictionary. This would be more effective if both were in the same volume and if the definition followed the translation. If this definition is in the target language, there is a possibility that it will

[22] The difference between *substitute* and *replace* does not coincide with *sustituir* and *reemplazar*.

not be entirely understood by the learner. In addition, a definition in the language of the *encoder* could possibly be more helpful in cases of subtle differences and where the distinction made in the target language does not exist in the source language. *Proceed Japanese-English Dictionary* explains the differences between *declare, proclaim* and *pronounce*, between *job* and *post* and many others in Japanese and on every page. What enables them to do this with so much conviction is that they know who their audience is and what their problems with very specific items are[23].

The *Longman Language Activator* is a monolingual kind of dictionary, but one can imagine a bilingual dictionary using this type of definition in order to distinguish between translations along the lines of the *Portuguese-English Bridge Bilingual*, although not that radical[24]. The *Longman Language Activator* tackles the task of defining each alternative quite well even if it does not always avoid unhelpful abstractions. In the case of *improve*, for instance, the *Activator* makes the following introductory distinction.

---

[23] *Cambridge International Dictionary* has been much criticised for the way in which it introduced contrastive information in its first edition. However, it is predictable that other learner's dictionaries will start adopting similar strategies in the future. It is different to learn English if you are Brazilian or Malay (Nesi, 1994) and dictionaries will have to adapt to this.

[24] The *Portuguese-English Bridge Bilingual* defining style implies the use of bot English and Portuguese, as in:

> **lug, lugs, lugging, lugged**, VB com OBJ Se você **lug** um objeto pesado, você o carrega com dificuldade; uso informal. *She lugged the suitcase out into the hallway.*

IMPROVE
1. to become better in standard or quality
2. ways of saying that a situation or someone's life becomes better, more enjoyable, less difficult etc.
3. to make something better by making changes, working harder etc.
4. to keep improving something that is already good until you make it perfect
5. to make a situation, quality, or someone's life better and more interesting, more enjoyable etc.
6. a change or addition that improves something

Figure 28 Introductory distinction for the entry improve *in the* Longman Language Activator.

To each different section –numbered 1 to 6 in the figure– corresponds a number of alternatives. This means that there is a number of *synonyms* suggested for each of these different senses. It is unlikely that learners will choose their meaning of *improve* from this list. On the other hand, the definitions of the *synonyms* which render the idea of *improvement* are as clear as humanly possible. However, the task is hugely difficult since the differences between *make something better, fine-tune, brush up* and *clean up* (among others) are evident for native speakers, but difficult to master for learners. One of the problems is that the different alternatives are defined in relationship to the hyponym, not to the other *synonyms*. *Streamline* is defined as 'to improve a system or process by…', *upgrade* is defined as 'to improve something such as machinery …', but *streamline* is not contrasted with *upgrade*.

A bilingual dictionary would have the advantage that the *synonyms* to be distinguished would be fewer. If a Brazilian learner wanted a translation for *aprimorar* s/he would have to choose between *refine* and *perfect* and would not have to start from the hyponym *improve*. The *Activator* perpetuates a tradition of English dictionaries for

speakers of any language which has been very useful until recently, but which I think will split up into minor areas with dictionaries more focused on specific languages.

Summing up, definitions are useful to distinguish *synonyms (modernise/upgrade)*, not if different translation options are the result of polysemy (*lead* (metal) and *lead* (guide)). In this case, in which the equivalents in the target language are completely different, much effort would be saved with a simple translation. For the lexicographer and for the learner. Since definitions are more demanding on the learner's powers of abstraction, they should only be used to disambiguate *synonyms* in cases well suited to them.

5.2.2.1.2. Via the Target Language

More advanced learners may prefer to obtain their information directly via the target language. In this case, they will necessarily start from some kind of target language input. This can be of various kinds: a superordinate; elements of a definition; number of syllables; sounds; gestures ('something you do with your fingers to show approval'); or a combination of these elements. This way of searching will be particularly useful for advanced learners who have difficulty remembering an item, or want to turn their passive knowledge into active knowledge. These learners have often already a good idea of what they want to say, even in the target language, but they might want to express themselves in a more formal way, or by using an idiomatic expression, a metaphor or a fixed expression. Furthermore, learners may have an idea of the definition of the word they are looking for, without knowing the word itself. For a number of advanced learners the ideal dictionary may be an improved kind of thesaurus.

A thesaurus presupposes that there is already a word in the mind of the learner, but that this word is inadequate from some point of view. There can be several reasons for this: register; congruent vs. metaphorical; lack of precision; regional variant, and others.

Some of the publications now on the market which call themselves a 'thesaurus' could just as well be classified as *synonym* dictionaries and some, like the *Concise Oxford Thesaurus* indeed bear as a subtitle *A Dictionary of Synonyms*. In what follows I will consider a *thesaurus* any book which lists *synonyms* and gives at the same time supplementary information. This information can be in the form of an example, a *label,* or a non-alphabetical classification.

Finding a Word Starting from its Definition

This is the situation in which learners have an idea of the components of the definition of the lexical item they are looking for. A learner should be allowed to input 'generous' and 'fight' and be presented with the word *magnanimous*, or 'give' and 'abundant' and get *lavish*.

One might want to express a movement of the lips, the definition of which would imply words like 'forward' and 'discontent'. In this case, it would be useful if a learner could search the definitions with the help of these words to end up choosing *pout*. Clearly this presupposes that there is some kind of simple, congruent definition for the word like the one *Oxford Wordpower* gives for *pout*:

> to push your lips, or your bottom lip, forward to show that you are not pleased about something.

The electronic *Collins English Dictionary and Thesaurus* and *Random House Webster's* have a feature which allows users to search the definitions for specific words. The difficulty is, however, that the learner has to use exactly the same words as those that were used in the definition. Taking as a basis the *Oxford Wordpower* definition, this means one has to use the phrase *not pleased* instead of *discontented.* In a computer based dictionary, however, this could rather easily be remedied since a *synonym* dictionary can be used to reduce the user's wishes to a particular definition. On the other hand, a more analytical and even playful method would be to start from *lips*, go to *movement* then to *forward*, then to *irritation* or *discontent* or *not pleased.* Additionally, the definition stored in the dictionary/thesaurus should be of a kind learners themselves would be able to formulate, i.e., not include *sullen* or other infrequent words.

Next is another example of a problem which a thesaurus could solve. Suppose that a learner, inspired by the French expression 'être d'une grande utilité' and the Dutch 'een grote hulp zijn', wants to express the ill-formulated idea: 'the encyclopedias *are of great help*'. There must be a way of saying this in English along the lines of:

— are of great help
— are of great use
— are of great utility
— are a great help
— are a big help

Learners should be allowed to look up 'help' + 'utility' for the dictionary to give them the right expression under several forms: formal, informal, metaphor, vulgar, plain, etc.

Using the same principle, a learner could find the right *collocate*. 'Does one really say *draw* the curtains?' could be a typical learner's question. It would be useful to be able to click on the word *curtain*, then to choose between *adjective, verb*, etc. and be given a list of possibilities.

Varying the Register

It is for learners, even advanced, extremely difficult to be sure if the register of the word they are about to use is right. Foreigners using informal expressions in formal contexts and vice-versa are a just source of innocent jokes in intellectual circles. It would be a typical task for a Thesaurus to eliminate this kind of confusion. Learners often know the formal, or neutral, word for what they want to say, but would like to sound more vernacular. Specifically in English, a learner would often prefer to use a phrasal verb rather than the Latin equivalent. However, before doing so, one has to be entirely sure of the implications of the alternative option. In what follows I will look at how native speaker thesauruses tackle this question.

In the case of *man, Collins Thesaurus* offers the following alternatives: 'bloke (Brit. inf.), chap (Inf.), gentleman, guy (Inf.), male'. If the learner has only temporarily forgotten the correct item, this list will surely help. For most learners, however, the precise difference between a *bloke* and a *chap* will escape them if the only supplementary information they get is the indication *Informal.*

Once again, it is unclear whose problem this indication of *Informal* is supposed to solve. Typically, a thesaurus user is an *encoding* user and s/he has to be very sure that the word s/he is going to use is the right word. If native speakers are targeted here, they

might not think at first of one particular alternative and find it in the thesaurus. In this case the indications *Brit.* and *inf.* are superfluous since a native speaker knows if they are *informal* or typically *British*. For a non-native speaker in search of a more informal way of referring to a *man* –for instance, translating *cara* in Portuguese– the choice between *guy, bloke* and *chap* remains hazardous if no supplementary information is supplied. This supplementary information will have to be looked up in another dictionary. One could contrast the *Collins Thesaurus* with *Roget's*, published over and over again, directed exclusively at native speakers and well adapted to their specific needs: not to teach the word, only to recall it starting from the idea.

The *Random Webster's Thesaurus* is meant for native speakers but is helpful for L2 learners because it is a thesaurus that separates the senses by means of an example. The result is a very clear overview of all the possibilities. The *Thesaurus* does not contain any register indications sufficiently explicit to help L2 non-native speakers out. Still, something can be learned from it.

man n.

1. *Man cannot live by bread alone*: mankind, the human race, men and women, human beings, humankind, people, humanity, Homo sapiens.

2. *Every man must follow his own beliefs:* individual, person, human being, human, living being, living soul, soul, one; anyone, somebody, someone.

3. *The average man is taller than the average woman:* male, masculine person; gentleman, chap, fellow; Slang guy, gent.

4. *The minister pronounced them man and wife*: married man, husband, spouse; Informal hubby.

5. *Hire a man to take care of the garden:* handyman, workman, hired hand, hand, laborer, day laborer; employee, worker; manservant, male servant, boy, waiter, footman, butler, male retainer; assistant, helper, right-hand man; male follower, subject, liegeman, henchman. v.

6. *The crew was ordered to man the lifeboats*: attend, staff, take up one's position in, take one's place at, get to one's post; supply with hands, furnish with men, people; equip, fit out, outfit; garrison.

Figure 29 Man *in* Webster's Thesaurus.

This kind of example –common to a number of dictionaries– does not give the meaning of the word but only helps to disambiguate its various senses. The fact is that these examples do this more effectively than a definition would, but they show at any rate that native speakers are the intended audience and the indication *slang* for *guy* and *gent* is, at the same time, superfluous and too vague for L2 learners.

A thesaurus in which this problem was recognised and successfully tackled is the *Schulsynonymik der deutschen Sprache* by Rudolf Meldau. The design of this dictionary shows the author had a clear idea of its audience's needs. 300 entries cover between 1500 and 1600 *synonyms*. There is a good *explanation* of what the register of each word is in those cases where such an explanation is appropriate. Apart from this, there is a wealth of examples which helps to situate the word in its register context. Under *Mann* one finds *Herr, Mannsbild, Kerl.* The information for *Mannsbild*, for instance, is:

fig.; Umgangssprache, meist abschätzig, selten bewundernd.

There are 6 examples for *Mannsbild*, 34 for *Mann*, a comparable number for *Herr* and *Kerl* which makes this Thesaurus exemplary as a learner's thesaurus.

Native speakers use thesauruses to recognise words, learners use them to learn new words as well. To be really practical, a superordinate should be surrounded by boxes on which to click, each of them mentioning one of its different lexical realisations. One would start from *man* and ask for something less formal, or more scientific, vulgar, etc. and examples would complement the information.

man	→ formal	→common use		→ *gentleman*
		→ scientific, human beings		→ *male*
		→ as a type		→ *male*
	→ informal	→ American	→ *guy*	
		→ British	→ modern	→ *bloke* (esp. working class)
			→ modern	→ *guy*
			→ critical	→ *fellow*
			→ slightly old-fashioned	→ *chap* (middle and upper-class)

Figure 30 Man *in an L2 learner-oriented thesaurus.*

The difference between a native speaker's thesaurus and a learner's one resides in the amount of information necessary to distinguish between several options. One example to illustrate an option, even if well chosen and translated, is often not enough.

Advanced learners are able to evaluate the tenor of a word by its context provided there is a huge number of instances, and only some sort of corpus can provide this. Here the use of authentic material seems indicated since it is precisely in the area of register that made-up examples are weakest. Made-up examples are certainly justified in the

case of beginners. They cannot give a correct idea of the use of a word for advanced learners.

Changing from Congruent to Metaphoric

The main task of an advanced learner's dictionary is to help learners 'complicate' their messages. Making their messages less congruent is one way of doing this. Native discourse is totally pervaded by metaphors. Beginning learners have limited ways of expressing themselves and tend to adapt their message to what they are able to say.

Advanced learners, on the other hand, have a message and study its most adequate wording. The area of *metaphors* is a virtually unexplored one in foreign language lexicography. Although metaphors are language specific, languages often use metaphors for the same phenomena and there are certain areas which are more susceptible of being *encoded* by means of a metaphor than others. 'Not to see anything' is an example of a universal metaphorisation (*I can't see a thing; no veo ni gota, je n'y vois goutte, ik zie geen fluit, não vejo bulufas,* etc.); 'not to care about something' (*I don't give a damn; me importa tres pimientos; je m'en fous; het kan mij geen barst schelen; não estou nem aí*). There are certain areas for which one instinctively feels there must be a metaphorical expression and thus a more natural way of expression. In a learner's thesaurus one should be allowed to input the congruent expression and get the metaphorical one. There are, to the best of my knowledge, no dictionaries available which do this kind of translation, although the elements to compile one are already contained in the many native speaker's thesauruses currently on the market. *Webster's Thesaurus* dedicates special attention to metaphorical expressions, including what Halliday (1994, Ch. 10, *passim*) would call *grammatical metaphors*, but one has to pick

them out one by one. Here is the entry for *think*:

> reason, reflect, cogitate, deliberate, *turn over in the mind, mull over*, ponder, contemplate, meditate, ruminate, *have in mind, make the subject of one's thought*, dwell on, brood, keep in mind, remember, recall, recollect, *use one's wits, rack one's brain.* (emphasis mine)

Along with metonymies, advanced learners feel the need to use grammatical metaphors as defined by Halliday. He calls expressions like 'use one's mind', 'apply the mind', 'take a walk', 'dar uma esperadinha', 'aan de beurt zijn' ('to be one's turn') *grammatical metaphors*. They are essential for anybody aiming at a native-like fluency since they are the normal native speaker's way of talking. In the case of grammatical metaphors, unlike metaphors in the strict sense of the word, it is not evident that all languages should have them for the same activities. One 'has a headache', 'a des maux de tête', 'heeft hoofdpijn', 'hat Kopfschmerzen', etc., but in Japanese the more common way to say it is 'atama ga itai desu', 'my head hurts', although the metaphorical way exists. Likewise, I do not know of any language apart from Portuguese where you can 'give a (little or normal size of) waiting' *(esperadinha, esperada),* 'holding' *(segurada, seguradinha)*, 'talking' *(falada, faladinha).* However, these expressions are essential for anyone aiming at native-like proficiency in Brazilian Portuguese and a learner's thesaurus should give the user easy access to this information.

Fixed Expressions

The use of fixed expressions is another way of making the language of a non-native speaker less congruent. The problem with finding fixed expressions in a monolingual dictionary is that, although the ideas which can be expressed in this way often coincide, their lexicalisation is different. There is no reason why *think hard* should mention

eyebrows as in *quemarse las cejas,* or *neurones* as in *queimar neurônios*. Nevertheless, there is a fair chance that a language has a fixed expression for *think hard* as opposed to the activity of, e.g., *park one's car*.

Similes, a kind of fixed expression, are rather well documented (Mikhail, 1994; Stoett, 1901, and others) because they are so obviously typical of each language and are recognised as such by its speakers. In my experience, people are aware of the fact that you do not necessarily think of *pitch* as the epitome of *black* nor of a *cricket* as the quintessence of *liveliness*. The English do, and say *black as pitch* and *lively as a cricket*. The ideal thesaurus should allow a learner to input a few words and the thesaurus would deliver the fixed expression.

This procedure is not very different from finding a word through its definition, but I know of no electronic dictionary which has this feature even in an embryonic form. It presupposes listing the fixed expressions of a language and defining them in terms of words which would be used to call them up. Current idiom dictionaries –as fixed expressions are also called– list the idioms according to their components, not their meaning, which makes these dictionaries less useful for learners. To understand what go *through the mill* means, a learner will consult a dictionary, probably under the entry *mill*, and all the bilingual dictionaries I consulted mentioned the expression. However, what is lacking is a dictionary which leads you from *go through difficult times* to *go through the mill*, or through a *sticky patch*; *defeat* should lead to *meet your Waterloo*; *disagree* should lead to *be at odds with.* The fact that such an evidently useful and easily realisable tool is not yet on the market is indicative of the lack of research into the needs of advanced learners. Particularly in the case of English as a foreign language this is

rather surprising.

Formulas

There is a long-standing tradition of appendixes to dictionaries dealing with another kind of fixed expressions which could be classified as *formulas* and are a kind of idiom in the Sinclairian sense of the word. Previously, dictionaries would limit themselves to *formulas* used in letters, but this section has been expanded lately to fixed expressions based on an analysis of the functions of language. *Collins-Robert* lists *Suggestions, Advice, Offers, Requests,* alongside more conventional items like *Correspondence* and *Announcements*. No doubt this is very valuable information and of the kind learners aiming at a native-like fluency should try to master. However, the problem here is the interface with the user. The list of functions and the way they are worded as they appear in *Collins-Robert* is the most complete one I know and is very valuable. However, which learner will think first of the statement they want to produce as an expression of the 'mechanics of argument', then pass on to the section 'moderating a statement' to find '*Sans vouloir critiquer cette façon de procéder, il semble cependant qu'...*'?

This is a recurring problem in lexicography: I would call it the 'triangle problem'. It consists in this: the dictionary user has in mind an expression at a certain level of abstraction. In order to retrieve the information in the dictionary s/he has to elevate that level, only to come back afterwards, in the other language, to the level of abstraction s/he started from. For instance, a learner has in mind the Portuguese verb *papear*. In order to find the right translation (*chat*), this learner has to raise the level of abstraction

to, say, *communicate verbally in an informal way*, to end up with *chat*, which is the right equivalent and, consequently, of the same level of abstraction.

This is a mental operation which normal learners, while *encoding*, are reluctant to perform, unless there is no other way of doing it or skipping it. Instead of this, a more realistic option could be to include these *formulas* in the body of the dictionary where the input of two or more words would lead the dictionary to suggest the rest, exactly as is the case now with electronic dictionaries with a browse-feature. The problem could be dealt with in the way Sinclairian *idioms* could be dealt with, since this is what these *formulas* are. *There is a strong possibility that*, for instance, given in *Collins German Dictionary* as a way of *Argumente abschwächen* (soften arguments), allows very few variations, if any. Starting from the source language, there is no reason why the formula could not be included under the entry for *think*.

5.2.2.2. Usage of the Item

Once the item has been selected, attention has to be given to its usage. Although it is now generally accepted that lexis and grammar are a whole which Halliday qualifies as a lexico-grammar, choosing the item and learning how it should be inserted into the text has purely formal aspects which it is the dictionary's task to teach. These formal aspects of *encoding* refer mainly to **syntax** and **collocates**, and this information can be transmitted in an implicit or an explicit way depending on the case and on the personal preferences of the learner.

The *explicit* way to transmit information on *syntax* is to list with every word the 'formula' of its use, which are the constraints it is subjected to. The *explicit* way to

transmit *collocational information* is by means of a list of all the lexical items that combine with the item in the target language.

The *implicit* way of transmitting this information is by means of *examples*. Different kinds of examples were described above. They have the ability to combine different kinds of information. I will discuss the advantages of a few kinds of examples for the purpose of *encoding*.

5.2.2.2.1. Explicit Information

Syntax

Syntax, for the purpose of this thesis, is a set of rules which governs the use of a lexical item within a sentence. These rules have a varying degree of intertwining with the meaning of the item and I am aware of the work of Sinclair and Francis on this issue (Francis, 1996). In most cases they are likely to appear to a learner as arbitrary and simply typical of a particular language, whereas in other cases these rules may be arbitrary in themselves, but intimately linked with the meaning. I suggest that the way in which the dictionary should handle each of these cases will depend on the category these items belong to. I will give a few examples of this.

In Japanese, *because* followed by a verb is translated by the formula [*nazenara* ... <verb in the dictionary form> *kara desu*]. This is a pattern which undoubtedly has meaning in itself within the context of the Japanese language, but for the learner it is a purely technical, 'meaningless' question. It should therefore be treated accordingly by the dictionary, giving the 'formula' explicitly, and a number of examples which illustrate this formula. Another example is the use of the French *subjonctif* after *désirer*

and other verbs expressing *will*. Although there is a remote connotation of 'unreality' involved in the use of the *subjonctif*, this is of little concern to *encoding* dictionary users and their queries should not be hampered by this sort of information.

In other cases, however, syntax is intimately intertwined with the meaning of the word. In English, *indulge* is used in the patterns *verb + in + noun; verb + noun*; and *pronominally*. In each of these constructions, *indulge* means something different and the equivalent of each of these patterns is a different word, for instance in French: *céder à; donner libre cours à; gâter; se laisser tenter; boire de l'alcool; se livrer à; se faire plaisir; s'offrir; se faire une gâterie en faisant*... .

In the first case, in which the syntax is arbitrary, a dictionary should give *explicit* information, examples still being important, but only secondarily. One can be given a number of examples of the Japanese *because*-construction without grasping that the verb has actually to be in the *dictionary form* (and not in the *-masu, -te, -nai,* or *-ta* forms). This is, however, an essential piece of information.

Syntax information is typical of production dictionaries and can be found in two forms: general information in a separate part of the dictionary, or specific information concerning specific items under the appropriate entries. I am very much in favour of considering a dictionary a word-based tool, particularising all the information so as to fit specific lexical items, i.e., concentrating everything in the entries. Explicit syntax indications in a dictionary are controversial because of their necessarily abstract and compact character. Abbreviations are partly to blame for some of the dictionaries' unpopularity. The reason for this is, as always, lack of space. This is likely to change in

electronic dictionaries although the ones published to date still follow the tradition. Furthermore, dictionary users feel instinctively put off by grammar because it demands a fair amount of mental effort and one cannot always be sure it is the best way of coping with the problems a specific item poses. When the syntax indications are, moreover, not followed by an example illustrating them, some will not consider it worth the effort.

Syntax indications should therefore not be used indiscriminately but only in those cases in which they really help and are vital to show how a word must be used: in those cases in which the syntax of a word has no bearing on its meaning (*because* in Japanese). In the opposite case *(indulge),* implicit syntax information in the form of examples will often be more efficient than a grammatical formula.

Collocations

Collocation is a neglected problem. Although the advent of the corpora era should have made it easier to deal with this issue, there does not seem to be much awareness of its importance. Advanced learners are supposed to have a vast passive knowledge. This is why simple solutions as E. H. Mikhail's *Dictionary of Appropriate Adjectives*, which lists 4000 nouns with the adjectives that can be used with them, is highly effective. This explicit way of giving information on collocates can be even more effective than examples. Collocates are, by definition, the words which occur most often with the target word and it is therefore likely that if advanced learners want to use a word, they will recognise the word that goes with it and applies to the context at hand. Examples could additionally help to dispel any doubts.

5.2.2.2.2. Implicit Information: Examples

Examples are a key feature for *encoding* learners. They can make the meaning of a word more precise, show in which contexts it is used, what the collocates are and what its syntax is. Since I dedicated a chapter (Ch.4) to the current situation of examples and established a few criteria, I will now concentrate on how different kinds of examples can assist specifically *encoding*. As I said before, a learner needs an enormous number of examples to be able to use a word correctly. Even the *Longman Language Activator,* which claims to concentrate on *encoding*, does not give more than three examples per word, and this is often not enough. Surely, lack of space and not of awareness is, in paper dictionaries, the main reason for this.

For the purpose of *encoding* I distinguish three kinds of examples: made-up, authentic, and controlled. The distinction between made-up and authentic is clear. What I call 'controlled' examples are those examples which can be retrieved from publications which target an audience of not only native speakers, like Encyclopedias and other reference works, mainly in English. I will deal with this in more detail later in this chapter.

Made-up examples

Made-up examples are the most common type of examples and until recently they were the only ones available. The first examples of the *Oxford Learner's Dictionary* were supposed to be a confirmation of the definition and corroborate it. They were, in this sense, *decoding* aids. At the same time they also taught something about the syntax of the word. However, as Silvia Bernardini mentioned to me in a personal communication,

made-up examples presume the lexicographer knows what the problem of the learners is and what has to be taught to them. This may be true for the most frequent senses of the most frequent words, but maybe not for the rest of the lexis.

House and *home* are frequent words which are both translated by *casa* in Portuguese. It is unlikely that these words will be looked up more by advanced learners than by beginners. These are the examples for *house* in the *Oxford Wordpower*, aimed at intermediate students.

1 a building that is made for one family to live in
Is yours a four-bedroomed or a three-bedroomed house?
Note: Look at **bungalow**, **cottage** and **flat**. Your **home** is the place where you live, even if it is not a house:
Let's go home to my flat.
Your home is also the place where you feel that you belong. A house is just a building:
We've only just moved into our new house and it doesn't feel like home yet.
You can **build**, **do up**, **redecorate** or **extend** a house. You may **rent** a house from somebody or **let** it out to somebody else. If you want to **move house** you go to an **estate agent**.
2 [*usually singular*] all the people who live in one house
Don't shout. You'll wake the whole house up.

Figure 31 Examples for house *in the Oxford Wordpower.*

The first example seems rather inadequate. The word *bedroomed* is more difficult than *house* and the structure *Is yours a...?* must sound a bit strange even to intermediate students. The example *Let's go home to my flat* is better because it illustrates the helpful comment on the distinction *house-home*. It is presumably not a very natural example. However, it fulfils the function that *paradigmata* would fulfil in traditional Latin grammar teaching. If learners learn by heart *Let's go home to my flat*, it is unlikely they will say *Let's go house to my flat*. The examples for *home* counterbalance this with the example *That old house would make an ideal family home.* This is not a very natural sentence either but it meets the specific needs of the words *home* and *house*, allowing

learners to make grammatically correct if not very natural sentences. The example *We've only just moved into our new house and it doesn't feel like home yet* reinforces this teaching point which was clearly the intention of the lexicographers.

Particularly in the case of very frequent words like *house*, authentic examples do not fulfil any immediately evident purpose, at least not if the dictionary is aimed at beginning students. COBUILD2 includes the following ones for the first sense of *house*:

She has moved to a small house and is living off her meagre savings. ...her parents' house in Warwickshire.

Figure 32 COBUILD2 examples for the first sense of house.

In the first example *meagre* is, in COBUILD's own classification, a one diamond frequency word and *live off* should not be much more frequent. The second example is too short to be helpful. No reference is made to *home* and the examples do not make clear the difference between *house* and *home*.

On the other hand, when made-up examples are used for less frequent words, the superiority of authentic examples shows clearly. *Impetus* is a very low frequency word. The entry in the *Oxford Wordpower* is:

impetus noun [uncountable] [singular] something that encourages something else to happen *I need fresh impetus to start working on this essay again.* (sic)

Figure 33 Impetus *in the* Oxford Wordpower.

The example for this entry has probably been made up. Any foreigner looking up *impetus* for *encoding* reasons supposedly knows what the word means, and at the most wants to confirm this. The need for information of this kind of learner concerns syntax and collocates (*gain, sustain, add, strong, fresh,* etc.) and an example only illustrating

what the definition says is of not much use.

Authentic examples

Since I dedicated a lot of attention to the problem of authentic examples in chapter 4, I will be brief here. Authentic examples seem to me to be extremely useful for the analysis of language particularly when the intention is to teach it to learners. For the purpose of *encoding*, authentic examples fulfil their aim in the long run, provided there are enough of them. It does not seem to me to be the most economical way of proceeding, but there is no doubt a corpus solves any *encoding* problem.

Another kind of authentic example is in my experience of great utility. Always in search of examples, I tried the CD-ROM packages which include all kinds of reference works and are becoming increasingly popular. I will discuss this kind of 'controlled example' in the next section.

Controlled examples

Controlled examples are a kind of example which are not being used by any dictionary and which learners themselves can glean from a number of Encyclopedias on CD-ROM. Although these were not conceived for this purpose, they are a fast way of retrieving a great number of examples. They have the advantage of being authentic and at the same time to give a broader context for the word at hand. Moreover, the language used by this kind of Encyclopedia can be considered mid-range both from the point of view of the register as from the point of view of the difficulty of the language. The following are only five of the 140 examples which the *Grolier Encyclopedia* came up with for a rather infrequent word like *impetus*:

1. The introduction of powered flight by means of the dirigible, a cylindrical balloon driven by propellers, and more important, by the first primitive airplanes gave *impetus* to the development of military air forces. (Air Force)

2. The California GOLD RUSH of the 1850s provided the *impetus* for the initial wave of immigrants from China. (Asian Americans)

3. The development of digital computers, which can monitor external conditions and make appropriate adjustments to a system, added further *impetus* to the applications of automation. (Automation.)

4. In the wake of the abortive coup d'etat attempted by hard-line political elements in the USSR in mid-August 1991 the three Baltic States received new *impetus* in their struggle for independence. (Baltic States)

5. Although Bacon was not a great scientist, he gave *impetus* to the development of modern inductive science. (Francis Bacon)

Figure 34 Five examples for impetus *in the* Grolier Encyclopedia.

These five randomly chosen examples all obey a regular, uncomplicated syntax, use an advanced type of vocabulary and provide learners with a number of collocates *(give, provide, add further, receive new)*. If a learner has any doubt about the meaning, or the range of applicability of the item, s/he can read as much of the context as required (and consult the built-in dictionary).

By contrast, we can compare these examples to the two examples COBUILD2 gives for *impetus*. They are of course authentic as well, but of a totally different character, as I hope to show below.

This decision will give renewed *impetus* to the economic regeneration of East London...

She was restless and needed a new *impetus* for her talent.

Figure 35 Examples for impetus *in COBUILD2.*

Apart from the fact that two examples may be too few to start using any new word with sufficient confidence, the examples themselves are learner-unfriendly because there is no way a learner can gauge with any certainty the type of publication –or speech– where these examples come from. The first one reads like a newspaper article,

whereas the second one sounds like fiction. With so little information, learners will have difficulty finding out whether *impetus* is a suitable word for their particular context.

In spite of this and because *impetus* happens to belong to a rather formal register, there is not much chance a learner would make a register error taking the COBUILD examples as a model. As for *collocates*, however, it is not clear what are the suitable collocates for *impetus*. Since there is no repetition, there is no way in which learners can know what the collocates are. From these examples they can only deduce that *give, renewed, need* and *new* can safely be used with *impetus*, not whether they are particularly frequent. After looking at a few examples from the *Grolier Encyclopedia*, on the other hand, it is immediately clear what the collocates are. Moreover, the uncontroversial style in which this kind of Encyclopedias are written gives learners a trustworthy model. It is unlikely that learners would want to use a word in a very original and experimental way right from the beginning. What they need are examples showing how the word is used, in a comprehensible style. This is not the case with corpus examples, at least in the way they are presently made available. The following are the first five examples from a total of 14 retrieved from the corpus of the COBUILD Dictionary on CD-ROM:

1. This weekend is our reunion when we meet again to reflect upon the tour with the *impetus* of our slides and photographs to invoke memory. (ephemera)

2. The committee agreed. Whispers that the committee had been wanting to name Arnie for years, waiting impatiently for Jack to reach the same conclusion. Whatever the *impetus*, the good deed was done. (magazine)

3. But you have to dive at a spectacular angle to gain sufficient *impetus* for a loop or roll, the control forces becomes (sic) high, and the speed drops off quickly and the nose pointed upwards. (magazine)

4. 5 per cent and that provided the *impetus* for an upward surge in trading at the St Leger Sales after a series of disasters had set the market back a full decade. (magazine)

5. There were growing fears that the rift would sour the annual Group of Seven summit and prevent the leaders of the West from providing fresh *impetus* for the long-running Gatt round. (newspaper)

Figure 36 Five examples for impetus *from the COBUILD CD-ROM.*

In the first example, the use of *we* indicates a rather informal setting and the words within the nine-word span *(of our slides)* are clearly not regular collocates. Examples 2 and 3 present a peculiar syntax which might even be considered wrong in the mouth of a foreigner. Example 4 gives too little information on the context, making it hard for the learner to understand what is meant, and example 5 includes two proper names, presupposes a lot of general knowledge, and the metaphors used by the author make the sentence simply hard to understand. I do not think this kind of example is particularly useful.

In the case of the examples retrieved from the COBUILD CD, learners have the advantage that the sources of the examples have been mentioned, which provides them with some supplementary information. However, apart from the sometimes difficult vocabulary and the problem of grasping what the text is about through only one sentence, what stands out from these examples is the fact that there is an author behind them, somebody expressing his or her personality. Still, what has to be welcomed in a magazine, newspaper or any other book, is supposed to be avoided in a reference work,

which is why dictionary definitions all seem to be written by one and the same person. One of the ways in which a personality expresses itself is by infringing the rules, by dashing if only slightly the expectations of the reader with the use of an unusual piece of syntax, a metaphor, or an unexpected collocate. The result is a number of sentences which are not always very clear. This does not mean that no collocates can be deduced from these COBUILD examples. It does mean that the learner will not feel confident to deduce them from the examples.

The *Grolier* examples, on the other hand, might not be immediately clear, but learners can read as much of the context as they want, or they can solve their problem by jumping to a next example since there are so many of them. An Encyclopedia might not reflect the whole of the language in all its intricacies, but since it was not written to illustrate the use of particular words, there is no danger that these examples were made up only to suit the lexicographers' argument.

Conclusion

Encoding is for advanced learners as difficult a problem as for beginning learners. Beginning learners have difficulty producing easy sentences, advanced learners have difficulties producing difficult sentences, and the tools for both could be improved.

As regards advanced learners, their look-up problem can be subdivided into two stages: the *choice* of the item, and its *usage*. These two steps are more clearly distinguishable in the case of advanced learners because in many cases they will be almost certain as to which item they want to use, and only be unclear as to its usage.

As for the *choice* of the item, learners may want to solve their lexicographic problem in two different ways: via the *source* language or via the *target* language. If they start from the language they know best, they will use a bilingual dictionary and will need some help so as to make a choice between the different alternatives. *Labels* are best suited to disambiguating content words. *Synonyms* are able to distinguish quickly the different translations for a word if they are clearly distinct entities. *Examples* are less appropriate for distinguishing various translation options for the kind of word which advanced learners need. Indeed, infrequent words depend less on the co-text to be disambiguated. Finally, *definitions* are useful to distinguish *synonyms (modernise/upgrade)*, but not when different translation options are the result of the fact that one word may indicate two completely different things (*lead* (metal) and *lead* (guide).

Not infrequently, *advanced learners*, used to expressing themselves and thinking in the target language, will have a fairly good idea of the lexical item they need. A dictionary should allow them to find an item starting from a few elements that are part of its definition. It should also give the possibility to vary the register of a given word, change expressions from congruent to metaphorical, and facilitate the use of fixed expressions.

Turning passive into active knowledge is one of the main challenges of an *encoding* dictionary. In practice, this means to provide the learner with sufficient information on the usage of an item. This information can be given in two basic forms: explicit and implicit. Explicit information on usage consists of grammatical information in its crudest form: rules. This is an appropriate method in cases in which the

grammatical structure of the item has no influence on its meaning. If not, an example is a more appropriate procedure. As for the issue of *collocations*, since *advanced* learners have an extended *passive* knowledge, it will be easy for them to recognise the collocate they are looking for and lists of them will be highly effective.

Another way of learning about the usage of a word is by means of examples. They have my personal preference. I distinguish three types: made-up, authentic and 'controlled'. Each of these examples have their own uses. For advanced learners, made-up examples are not very adequate since they mostly teach one specific point, or only illustrate the definition. Authentic examples have the advantage of combining several kinds of information but are sometimes confusing. 'Controlled' examples, as I call them, can be found in Encyclopedias and the like. They have the advantage of being authentic, guarantee the register, bring with them a whole context, and are written in a clear and uncontroversial way.

CHAPTER 6 CONCLUSION

In this thesis I have addressed a number of issues which I think are essential to improve dictionaries for foreign language learners.

I suggested that needs should drive the conception of a tool such as a dictionary, but when I situated foreign language dictionaries in the broader field of historical lexicography, I suggested that the relationship between needs and the tools supposed to meet them was no longer as clear as when dictionaries were first invented. Bilingual dictionaries still fulfil an obvious need, but the writings of Baudrillard (1968, 1970, 1972) convinced me of the fact that not everything has necessarily a usage value, and that every object also functions in a symbolical circuit in which the token value predominates. Over the centuries, monolingual native speaker's dictionaries were greatly influenced by ideological considerations and this has made their relationship with their audience an unclear one.

As a result, learner's dictionaries, which have come to assist non-native speakers in their learning of a foreign language and were derived from native speaker's dictionaries, have inherited some of their defects. They are still adaptations of the original idea, instead of tools directly modelled on the analysis of learner's needs. At the same time, the needs of present-day foreign language learners have changed. This change consists basically in a switch of interest from merely understanding texts written in a foreign language, to expressing oneself in that language.

An analysis of the learner's needs is an essential ingredient of any project aiming at improving the efficiency of foreign language learning tools. It was to have a better

idea of what these needs are and the methods used to discover them that I resorted to the specialised literature on the subject. After a while I realised that part of this literature dealt with problems which modern technology had almost made obsolete. Another part of it dealt with what I considered to be the right questions, but research on the matter did not seem to come to any convincing conclusions.

This was, in my opinion, largely due to the application of inadequate research methods. It also meant that there was no consensus on what exactly had to be discovered. Reading Kuhn (1970), I had the idea that, in the field of theoretical lexicography, a clear awareness of the foundations of the discipline was missing, and that the techniques used to acquire new knowledge were epistemologically flawed. It is not enough to study 'habits', if you want to discover 'needs'. Popper (1994, 1995) convinced me of the fact that methods borrowed from the exact sciences, and which are at present the norm, were inappropriate.

Indeed, the two methods used by researchers to discover learners' needs in terms of dictionaries were *questionnaires* and *tests*. Since the results to which these techniques led were not entirely satisfactory and sometimes did not agree with my own experience, I concluded the matter should be tackled differently. Agreeing with Hillary Nesi (1996), I thought a qualitative approach was a better way of dealing with the problem. I decided therefore to privilege introspection and close-reading.

In my research this expressed itself in two ways. Where I could I would ask learners for their experience with dictionaries. This was not always a rewarding exercise, since a number of learners would answer the way they thought they were

expected to answer either by me, or by society. This reflected secondarily the *questionnaire* situation and confirmed to me that this methodology was indeed not always suitable. It was only by 'cross-examining' my subjects, asking questions about what they effectively did, that I had a few interesting answers. More results, however, I obtained from a second method, by examining dictionaries myself, trying to solve specific problems. It is in these moments of 'réflexion jointe à l'usage' that I had, I hope, a few 'idées nettes...'.

In one particular case, I set out to analyse qualitatively a feature which I deem to be of the utmost importance in foreign language lexicography: *examples*. Without examples it is impossible for a learner to start using a new word. However, no agreement exists as to the best kind of example and over the last few years there has been some controversy on the issue. The main cause of the dispute was the publication of the first COBUILD dictionary in 1987. Because of the controversial character of the issue, I thought it was an adequate topic to demonstrate how qualitative research could be carried out.

In this research, I investigated the nature of examples and what characteristics can be expected from them. I also took a critical look at the treatment given to examples in traditional and less traditional foreign language lexicography, bilingual and learner's. In both cases, I had to conclude that example policies still suffer from a few shortcomings, the main reason being that no clear distinction is made between *encoding* and *decoding*. In the absence of any clear idea of what *examples* are supposed to assist with, their choice remains in some cases rather casual and learners miss out on one of the main forms of help a dictionary can give them. In terms of collocates and syntax, examples do

rather poorly in all of the dictionaries I investigated. These typical requirements of *encoding* learners are still not taken into due consideration and suffer from a seemingly haphazard treatment.

Finally, I had a closer look at *authentic examples* in connection with the COBUILD experience. My conclusions were that authentic examples were certainly useful and indeed a revolution in the world of foreign language lexicography, but that their use should not be elevated into a dogma. Depending on the case, made-up examples may be more suited.

In the last part of this thesis, I give an idea of what I would consider to be the ideal foreign language dictionary. An essential characteristic of this new dictionary is the distinction between *encoding* and *decoding* which I think is a basic requirement. Starting from this distinction, I discuss several aspects of both processes. The case of *decoding* is easier because the context in which a hard word is found always helps comprehension. Even so I discuss a few of the problems that can arise in the course of a *decoding* process. *Polysemy* is one of the most important ones and I hope to have made a few helpful suggestions.

As for *encoding*, this is clearly a more difficult undertaking. The number of expedients invented over the centuries to assist learners with *encoding* is high and generations of lexicographers have dedicated themselves to perfecting them. All this material should be re-used although applied in a way more adapted to present-day needs. *Labels, synonyms, examples* and *definitions,* all convey a particular kind of information but, depending on the item, one is more useful than the other.

Furthermore, it is necessary to make a distinction between the *encoding* problems *beginners* have, and those of *advanced* learners. Beginners will have a great number of grammatical problems which have to be attended to and will often not know what exactly to look up. Whereas it is the task of language classes to teach macro-grammatical concepts, it is the task of the dictionary to show the grammatical constraints of specific words. Function words are in this aspect the most complex items and examples are the most suitable way to teach them.

Encoding advanced learners are constantly in search of the right collocate, and are often in the dark as to the appropriateness of the item they want to use. From my own experience and from discussions with other people who need to express themselves in a foreign language at a proficient level, I draw the conclusion that in the field of reference works much still has to be done.

This work has to be based on an analysis of the advanced learners' needs. In the section on advanced *encoding* I have analysed a few of these needs and I have suggested ways of meeting them. *Encoding* learners have at their disposal several different ways of obtaining the item they need. A ideal reference tool could integrate various approaches and show flexibility. This flexibility will be greatly enhanced in dictionaries destined to be used on computers.

With this thesis I had the intention to make a contribution to the improvement of dictionaries which I hope may have some practical consequence. In order to have this kind of outcome, I am aware that much additional research will have to be carried out and this will justify my academic existence for a few more years to come. More research

is needed as to the frequency of an item and the kind of example best suited for it. This has some bearing on the relation frequency/polysemy which I am eager to start investigating. The study of Japanese has convinced me that a number of problems go unnoticed if we are familiar with only one type of language and I want to pursue this study further. Many of the insights I had on lexicography were due to this experience of *dépaysement* which had and will have consequences for the way in which I see dictionaries and how I would like them to be. Finally, I would like to do some more research on 'controlled examples'. It is still not clear to me what makes them so different from plain authentic examples.

Some of this research, I hope, will be of profit for the *Dicionário de Uso Português-Espanhol,* in progress. As in everything, a combination of theory and practice is the best way to further both.

BIBLIOGRAPHY

1. Dictionaries

1.1. Bilingual Dictionaries

1. *Basic Japanese-English Dictionary.* 1994. The Japan Foundation, Oxford University Press, Oxford.
2. *Cassell Colloquial Handbooks Spanish.* 1992. Cassell, London.
3. *Collins* COBUILD *Students' Dictionary. Bridge Bilingual English-Portuguese.* 1995. HarperCollins, London.
4. *Collins Gads Danish Dictionary.* 1994. HarperCollins, London.
5. *Collins German Dictionary.* 1991. HarperCollins/Pons, London.
6. *Collins Norstedts Swedish Dictionary.* 1995. HarperCollins/Norstedt, London.
7. *Collins Robert French Dictionary.* 1994. HarperCollins/Robert, London.
8. *Collins Russian Dictionary.* 1995. HarperCollins, London.
9. *Collins Sansoni Italian Dictionary.* 1995. HarperCollins/Sansoni, London.
10. *Collins Shubun English-Japanese Dictionary.* 1993. HarperCollins/Shubun, London.
11. *Collins Spanish Dictionary.* 1994. HarperCollins/Grijalbo, London.
12. De Garcia H. 1958 (1947). *Dicionário Português-Espanhol*, Globo, Porto Alegre.
13. *Diccionario de bolsillo portugués-español.* 1978. Sopena, Barcelona.
14. *Dicionário de Ouro Espanhol-Português.* (no date), Ediouro, Rio de Janeiro.
15. *Dicionário Escolar Francês-Português-Francês.* 1980. Fename, Rio de Janeiro.
16. *Dicionário Inglês-Português-Inglês.* 1982. Record, Rio de Janeiro.
17. *Dicionário Japonês-Português.* 1990. Rideel, São Paulo.
18. *Dictionary of English Synonyms and Antonyms.* 1992. Penguin, Harmondsworth.
19. *Genius English-Japanese Dictionary.* 1988. Taishukan, Tokyo.
20. *Il Nuovo Dizionario Hazon-Garzanti Inglese-Italiano-Inglese.* 1991 (first edition; 1990). Hazon-Garzanti, Milano.,.
21. *Larousse Français-Espagnol-Français.* 1987. Larousse, Paris.
22. *Longman Dictionary of Contemporary English.* 1995. Longman, Harlow.
23. *Martin's Concise Japanese Dictionary.* 1994. Charles Tuttle Company, Rutland, Tokyo.
24. Morsbach H. and Kurebayashi K. 1994. *Japanese Phrase Book and Dictionary*, HarperCollins, London.
25. *New Proceed English-Japanese Dictionary.* 1994. Benesse Corporation.

26.*Nouveau dictionnaire français-japonais.* 1921. Hakusuisha, Tokyo.

27.*Oxford-Duden, Oxford-Hachette, and Oxford Spanish Dictionary on CD-ROM.* 1996. Oxford University Press, Oxford.

28.*Pequeno dicionário de provérbios Alemão/Francês/Inglês/Português.* 1983. Moraes, Lisboa.

29.*Pocket English Thesaurus.* 1985. Penguin, Harmondsworth.

30.*Pocket Greek Dictionary.* 1995. HarperCollins, London.

31.*Portuguese Dictionary.* 1991. HarperCollins, London.

32.*Portuguese-English Dictionary.* 1963. Record, Rio de Janeiro.

33.*Prisma Handwoordenboek Duits Nederlands.* 1971. Prisma, Utrecht, Antwerpen.

34.*Proceed Japanese-English Dictionary.* 1988. Fukutake, Tokyo.

35.Santos Saraiva, (no date) *Diccionário Latino-Portuguêz.* Garnier, Rio de Janeiro.

36.*Slovene Phrase Finder.* 1996. HarperCollins, London.

37.*Taschenwörterbuch der Spanischen und Deutschen Sprachen.* 1927. Langenscheidt, Berlin.

38.Tochtrop L. 1984. *Dicionário Alemão-Português.* Globo, Rio de Janeiro.

39.*Van Dale Nederlands Engels Woordenboek.* 1991. Van Dale, Utrecht.

40.Van Dam C. F. A. 1968. *Spaans Handwoordenboek.* Van Goor, Den Haag/Brussel.

41.*Word Selector Inglés-Español.* 1995. Cambridge University Press, Cambridge.

1.2. Learner's Dictionaries

1. Benson, M. & al. 1986. *The BBI Combinatory dictionary of English. A guide to word combinations*, John Benjamins Publishing Company, Amsterdam/Philadelphia.
2. *Cambridge International Dictionary of English.* 1995. Cambridge University Press, Cambridge.
3. COBUILD *Dictionary of Idioms* 1995. HarperCollins, London.
4. *Collins* COBUILD *English Language Dictionary.* First edition. 1987. HarperCollins, London.
5. *Collins* COBUILD *English Language Dictionary.* Second edition. 1995. HarperCollins, London.
6. *Collins English Dictionary and Thesaurus.* 1992. Electronic version 1.5, HarperCollins, London.
7. Hornby A. S. *Advanced Learner's Dictionary of Current English.* 1972 (ed. 1963). Oxford University Press, Oxford.
8. Konder R. 1982 (ed. 1993). *Longman English Dictionary for Portuguese Speakers.* Ao Livro Técnico, Rio de Janeiro.
9. *Longman Dictionary of Contemporary English.* 1995. Longman, Harlow.

10.*Longman Language Activator*. 1993. Longman, Harlow.

11.*Macmillan Learner's Dictionary of American English*. 1994. Macmillan, Lincolnwood.

12.*Oxford Advanced Learner's Dictionary of Current English*. 1989. Oxford University Press, Oxford.

13.*Oxford Advanced Learner's Dictionary of Current English*. Oxford University Press, Oxford.

14.*Oxford Advanced Learner's Encyclopedic Dictionary*. 1993 (ed. 1989). Oxford University Press, Oxford.

15.*Oxford Wordpower*. 1993. Oxford University Press, Oxford.

1.3. Monolingual Native Speakers Dictionaries

1. Brouwers L. 1988. *Het juiste woord*. Standaard, Antwerpen.
2. Casares J. 1979. *Diccionario ideológico de la lengua española*. Ed. Gustavo Gili, Barcelona.
3. *Cassell's Modern Guide to Synonyms and Related Words*. 1984. Book Club Associates, London.
4. *Concise Oxford Dictionary of Current English*. 1996. Oxford University Press, Oxford.
5. *Concise Oxford Thesaurus*. 1995. Oxford University Press, Oxford.
6. Corominas J. 1961 (ed. 1983). *Breve Diccionario Etimológico de la Lengua Castellana*, Gredos, Madrid.
7. Corripio F. 1979. *Gran diccionario de sinónimos*, Bruguera, Barcelona.
8. Covarrubias S. (ed. 1989). *Tesoro de la Lengua Castellana o Española*, Ed. Alta Fulla, Barcelona.
9. *Der Sprach-Brockhaus*. 1965. Brockhaus, Wiesbaden.
10. *Diccionario temático de la lengua española*. 1975. Biblograf, Barcelona.
11. *Dictionary of English Idioms*. 1986. Penguin, Harmondsworth.
12. *Dizionario Garzanti de la Lingua Italiana*. 1982 (ed. 1963). Garzanti, Milano.
13. Dos Santos Azevedo F. 1983. *Dicionário analógico da língua portuguesa*. Thesaurus Ed. e Coordenada Ed., Brasília.
14. *Fowler's English Usage*. 1990. Oxford University Press, Oxford.
15. *Franck's Etymologisch woordenboek der Nederlandse taal*. 1971. Martinus Nijhoff, 's Gravenhage.
16. *Longman Dictionary of the English Language*. 1992. Longman, Harlow.
17. Meldau R. 1972. *Schulsynonymik der Deutschen Sprache*. Julius Groos, Heidelberg.
18. Mikhail E. H. 1994. *Cassell Dictionary of Appropriate Adjectives*. Cassell, London.
19. Moliner M. 1975. *Diccionario de uso del español*. Gredos, Madrid.

20.Morínigo M. 1993. *Diccionario del español de América.* Anaya/Muchnik, Madrid.

21.*Nuevo Diccionario Lunfardo.* 1994. Corregidor, Buenos Aires.

22.*Pequeño diccionario Kapelusz de la lengua española.* 1980. Kapelusz, Buenos Aires.

23.*Pequeño Larousse en color.* 1976. Larousse, París.

24.*Petit Robert 2.* 1977. Société du nouveau Littré, Paris.

25.*Petit Robert.* 1973. Société du nouveau Littré, Paris.

26.*Roget's Thesaurus of English words and phrases.* 1852 (ed. 1987). Penguin, Harmondsworth.

27.Stoett F. A. 1901 (1943). *Nederlandsche Spreekwoorden, spreekwijzen, uitdrukkingen en gezegden.* Thieme, Zutphen.

28.*The compact Oxford English Dictionary.* 1994 (2nd ed.). Oxford University Press, Oxford.

29.*Van Dale Groot Woordenboek der Nederlandse Taal.* 1976. Martinus Nijhoff, 's Gravenhage.

30.*Wahrig Deutsches Wörterbuch.* 1987 (ed. 1966). Mosaik Verlag, München.

31.*Wahrig Wörterbuch der Deutschen Sprache.* 1981. Deutscher Taschenbuch, Berlin.

1.4. Electronic Reference Works

1. *Bibliorom Larousse.* 1996. Microsoft/Larousse.
2. *Bookshelf '94.* 1994 Microsoft Corporation, Redmond.
3. Buarque de Holanda A. *Aurélio eletrônico* V. 1.3. Nova Fronteira, Rio de Janeiro.
4. *Chambers Dictionary on CD-ROM. V. 2.1.* 1993.
5. *Collins* COBUILD *English Collocations CD-ROM.* 1995. HarperCollins, London.
6. *CyberDico. Le dictionnaire multimédia de la langue française.* 1996. Goto Informatique, Tourcoing.
7. *Encarta '95.* 1995. Microsoft Corporation, Redmond.
8. *Euroglot Professional V.1.* 1994. Linguistic Systems, Nijmegen.
9. *Larousse Référence Électronique V1.11.* 1992-93. Larousse, Paris.
10. Moliner M. 1996. *Diccionario de uso del español.* Edición en CD-ROM. Gredos, Madrid.
11. *Oxford Wordpower on Diskette.* 1993. Oxford University Press, Oxford.
12. *Petit Robert sur CD-ROM.* 1996. Société du nouveau Littré, Paris.
13. *Random House Webster's Electronic Dictionary and Thesaurus*, V.1. 1992. Webster's, New York.
14. *The Software Toolworks Multimedia Encyclopedia.* 1992. version 1.5, Grolier.
15. Zingarelli N. 1998. *Lo Zingarelli 1998 in CD-ROM. Vocabolario della Lingua Italiana*, Zanichelli, Bologna .

2. Secondary Bibliography

1. Aarts J. and Meijs W. (Eds.) 1983. *Corpus Linguistics: recent developments in the use of computer corpora in English language research.* Rodopi, Amsterdam.
2. Ahmad K. et al. 1992. The elaboration of special language terms: the role of contextual examples, representative samples and normative requirements. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere. Finland.* Part I, Studia Translatologica Ser. A Vol. 2, Tampere. 139-149.
3. Al B. 1990. Transfert automatique de données lexicales par le biais d'un dictionnaire bilingue. *Cahiers de Lexicologie*, 56-57: 123-130.
4. Al B. et al., 1990. Thème I. Questions et remarques. Communication de H. Kromann. *Cahiers de Lexicologie*, 56-57: 63-71.
5. Al-Kasimi A. M. 1983. The interlingual/translation dictionary. Hartmann R. R. K. 1983. *Lexicography: Principles and Practice. Academic Press, London,* 153-162.
6. Antoine F. 1994. Un nouveau bilingue anglais-français (BIS). Le Dictionnaire Hachette-Oxford. *Cahiers de Lexicologie*, 1994/2 183-196.
7. Apresyan Y. D. Mel'cuk I. A. and Zolkovsky A. K. 1969. Semantics and lexicography: Towards a new type of unilingual dictionary. Kiefer, F. (Ed.) 1969. *Studies in syntax and semantics*, Reidel, Dordrecht-Holland, 1-33.
8. Atkins B. T. and Knowles, F. E. 1990. Interim report on the *Euralex*/AILA research project into dictionary use. Magey, I. and Zigany, J. *Budalex 88 Proceedings* ed. By Y. Magay and J. Zigány. Budapest: Akadémiai Kiadó 381-392.
9. Atkins B. T. et al. 1987. A research project into the use of learner's dictionaries. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 29-43.
10. Atkins B. T. S. 1996. Bilingual dictionaries: past, present and future. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg, Sweden*, Göteborg University, Department of Swedish, 515-546.
11. Atkins B. T. S. and Varantola, K. 1997. Monitoring dictionary use. *International Journal of Lexicography*, Vol. 10, 1:1-45.
12. Atkins B. T. S. and Zampolli, A. 1994. *Computational Approaches to the Lexicon*, Oxford University Press, Oxford.
13. Atkins B. T. S. Levin, B. and Song G. 1996. Making sense of corpus data: a case study. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 345-354.
14. B. A. A. L. 1979. *Seminar on Lexicography, Dictionaries and their users.* University of Exeter.
15. Bacon F. 1952 (1620). *Advancement of learning. Novum Organum, New Atlantis*, William Benton, Encyclopedia Britannica, The University of Chicago, Chicago.

16. Baganz L. 1987. Problems of a Hindi-German Dictionary. Ilson, R. (Ed.) *A Spectrum of Lexicography. Papers from AILA BRUSSELS 1984.* John Benjamins, Amsterdam/Philadelphia 101-114.

17. Bailey, R. (ed.) 1987. *Dictionaries of English: Prospects for the Record of our Language.* Ann Arbor: University of Michigan Press.

18. Barthes R. 1957. *Mythologies.* Seuil, Paris.

19. Basnett-McGuire S. 1991. *Translation Studies.* Methuen, London.

20. Battenburg J. D. 1991. *English monolingual learner's dictionaries: a user oriented study*, Lexicographica, Series Maior 39, Niemeyer, Tübingen.

21. Baudrillard J. 1968. *Le système des objets.* Gallimard, Paris.

22. Baudrillard J. 1970. *La société de consommation.* Gallimard, Paris.

23. Baudrillard J. 1972. *Pour une critique de l 'économie politique du signe.* Gallimard, Paris.

24. Béjoint H. 1981. The foreign student's use of monolingual English dictionaries: a study of language needs and reference skills. *Applied Linguistics,* Vol. II Number 3, Autumn 1981, 207-222.

25. Béjoint H. 1983. On fieldwork in lexicography. Hartmann R. R. K. 1983. *Lexicography: Principles and Practice. Academic Press, London,* 67-76.

26. Béjoint H. 1990. Review of: Harrap's Pocket French-English Dictionary. *International Journal of Lexicography*, Vol. 3, 1.

27. Béjoint H. 1993. *Tradition and innovation in Modern English Dictionaries*, Oxford University Press, Oxford.

28. Béjoint H. and Moulin, A. 1987. The place of the dictionary in an EFL programme. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 97-114.

29. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot, Louvain-la-Neuve,.

30. Béjoint M. 1979. The use of informants in dictionary making. Hartmann, R. R. K. (Ed.) *Dictionaries and their users.* Exeter Linguistic Studies 1 (4) : 25-30. University of Exeter.

31. Benson M. 1989. The structure of the collocational dictionary. *International journal of lexicography*, Vol. 2. 1, Oxford.

32. Bensoussan M., Sim D. and Weiss R. 1984. The effect of dictionary usage on EFL test performance compared to the student and teacher attitudes and expectations. *Reading in a Foreign Language* 2 (1984) 15-32.

33. Berkov V. 1996. Passive vs. active dictionary. A revision. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 547-550.

34. Binon J. 1991. Un dictionnaire d'apprentissage du français des affaires. Projet et propositions de réalisation. S.Verlinde (Ed.) *Proceedings of the Symposium on*

differentiation in LSP, Learning and Teaching. Leuven, Instituut voor Levende Talen. 143-188.

35. Binon J. 1995. Lexicographie pédagogique et apprentissage/enseignement du français des affaires. Les dictionnaires d'apprentissage comme générateurs d'exercices. Thessalonique, septembre 1995.
36. Binon J. and Verlinde S 1994. The *Dictionnaire contextuel du français économique*. A production oriented dictionary of business French. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex International Congress on Lexicography in Amsterdam. The Netherlands,* 523-529.
37. Binon J. and Verlinde S. 1992. Le dictionnaire d 'apprentissage du français des affaires. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Finland. Part I, Studia Translatologica ser. A Vol. 2, Tampere.43-50.
38. Binon J., Dancette J., Verlinde S. 1997. Comment optimaliser le traitement des (para)synonymes dans un dictionnaire d'une langue de spécialité? (forthcoming).
39. Blanco X. 1996. Élaboration et réutilisation des exemples dans la lexicographie bilingue. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot, Louvain-la-Neuve, 103-110.
40. Bloom H. 1994. *The western canon*, Harcourt Brace & Company, New York.
41. Bogaards P 1995. Dictionnaires et compréhension écrite. *Cahiers de Lexicologie*, 67(1995), 37-53.
42. Bogaards P. 1988. A propos de l 'usage du dictionnaire de langue étrangère, *Cahiers de Lexicologie*, 52:131-152.
43. Bogaards P. 1990. Deux langues, quatre dictionnaires. Dolezal F. M. et al. (Eds.), Lexicografica 6/1990. Niemeyer, Tübingen.
44. Bogaards P. 1992. 'French dictionary users and word frequency. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Finland. Part I, Studia Translatologica ser. A Vol. 2, 51-62.
45. Bogaards P. 1992. French dictionary users and word frequency. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Finland. Part I, Studia Translatologica ser. A Vol. 2, Tampere.51-62.
46. Bogaards P. 1994. Synonymy and bilingual lexicography. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex International Congress on Lexicography in Amsterdam. The Netherlands,* 612-618.
47. Bogaards P. 1996. Dictionaries for learners of English. *International Journal of Lexicography* Vol. 9, 4:277-320.
48. Boisson C. 1996. L'antiquité et la variété des dictionnaires bilingues. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot,

Louvain-la-Neuve, 17-30.

49. Boswell J. 1791 (edition 1897). *The life of Samuel Johnson, LL. D.* Edited by Percy Fitzgerald, Bliss Sands, London.

50. Botha W. 1992. The lemmatisation of expressions in descriptive dictionaries. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Part 2, 465-472.

51. Boysen G. 1990. Information syntaxique dans les dictionnaires bilingues. *Cahiers de Lexicologie*, 56-57: 43-49.

52. Bray L. 1988. La lexicographie bilingue Italien-Allemand, Allemand-Italien du dix-septième siècle. *International Journal of Lexicography*, Vol. 1, 4: 313-342.

53. Broeders T. 1987. The treatment of phonological idioms. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 246-256.

54. Broughton G. 1978. Native speaker insight. English Language Teaching Journal 32 (4) : 253-6.

55. Brown S. 1994. Hard words for the ladies; the first English dictionaries and the question of readership. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex international Congress on Lexicography in Amsterdam. The Netherlands.* 355-364.

56. Burchfield R. 1987. *Studies in Lexicography*, Oxford, Clarendon Press.

57. Callebaut B. De l'interférence au transfert, ou 'peut-on enseigner le SAE? Ilson, R. (Ed.) A Spectrum of Lexicography. *Papers from AILA BRUSSELS 1984*, John Benjamins, Amsterdam/Philadelphia 115-127.

58. Calzolari N. et al. 1987. The use of computers in lexicography and lexicology. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 55-77.

59. Carter R. 1987. *Vocabulary. Applied Linguistic Perspectives*, Allen & Unwin, London.

60. Carter R. 1988. Review of LDOCE and COBUILD. *International Journal of Lexicography*, Vol. 2, 1:30-43.

61. Catford J. C. 1965. *A Linguistic Theory of Translation.* Oxford University Press, Oxford.

62. Chi A. 1995. A critical user-oriented review of the *Collins* COBUILD *English Language Dictionary*: definitions and examples. Wong Y. Kitty and Green F. Christopher 1995. *Thinking language. Issues in the study of language and language curriculum renewal*, Language Centre, The Hong Kong University of Science and Technology, 45-56.

63. Chi A. and Yeung S. 1996. A case study of the use of dictionaries to assist the teaching and learning of English in Hong Kong. (forthcoming).

64. Choul J. C. 1987. Contrôle de l 'équivalence dans les dictionnaires bilingues. Ilson,

R. (Ed.) *A Spectrum of Lexicography. Papers from AILA BRUSSELS 1984*, John Benjamins, Amsterdam/Philadelphia 75-90.

65. Collison R. L. 1982. *A History of Foreign-Language Dictionaries*. Deutsch, London.
66. Coulthard M. (Ed.) 1994. *Advances in written text analysis*, Routledge, London.
67. Cowie A. (Ed.) 1987. The dictionary and the language learner. *Papers from the Euralex Seminar at the University of Leeds*, Tübingen, Niemeyer 11-28.
68. Cowie A. 1981. The treatment of collocations and idioms in learner's dictionaries. *Applied Linguistics,* Vol. II Number 3, Autumn 1981, 222-235.
69. Cowie A. 1983. The Pedagogical/Learner's dictionary. Hartmann R. R. K. 1983. *Lexicography: Principles and Practice. Academic Press, London,* 135-143.
70. Cowie A. 1983b. English dictionaries for the foreign learner. Hartmann, R. R. K. (Ed.) *Lexicography: principles and practice*: 153-43. Academic Press, London,.
71. Cowie A. 1984. EFL Dictionaries: past Achievements and present Needs. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 155-164.
72. Cowie A. 1987. Syntax, the dictionary and the learner's communicative needs. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 183-192.
73. Cowie A. 1989. Pedagogical descriptions of language: Lexis. *Annual Review Of Applied Linguistics*, Cambridge University Press, Cambridge.
74. Cowie A. 1989. The language of examples in learner's dictionaries. James G. (ed) *Lexicographers and their Works.* Exeter Linguistic Studies volume 14. University of Exeter 55-65.
75. Creamer T. 1984. Recent and forthcoming Chinese-English and English-Chinese Dictionaries from the People's Republic of China. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 176-182.
76. Creamer T. 1987. Beyond the definition: Some problems with examples in recent Chinese-English and English-Chinese bilingual dictionaries. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Niemeyer, Tübingen 238-245.
77. Cruse, D. A. 1986. *Lexical Semantics*. Cambridge University Press.
78. Crystal D. 1995 *The Cambridge Encyclopedia of the English Language*. Cambridge University Press.
79. Dalgish G. M. 1995. Learner's dictionaries: Keeping the learner in mind. Kachru B. and Kahane H. (Eds.) *Cultures, ideologies and the Dictionary, Studies in honour of Ladislav Zgusta*, Lexicographica Series Maior 64, Niemeyer, Tübingen 329-340.
80. Darbelnet J. 1970. Dictionnaires bilingues et Lexicologie différentielle. *Langages*, 19:92-102.
81. de Stadler L. 1992. Syntagmatic lexical relations: a lexicographical perspective. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international*

Congress on Lexicography in Tampere. Part 2, 411-418.

82. Dendien J. 1990. Eléments de réflexion pour l 'informatisation d 'un dictionnaire. *Cahiers de Lexicologie*, 56-57: 191-200.
83. Descartes R. 1951 (1637). *Discours de la méthode*. Éditions Montaigne, Paris.
84. Diab T. 1990. *Pedagogical Lexicography. A case study of Arab nurses as dictionary users*, Lexicographica Series Maior, Niemeyer, Tübingen.
85. Dickens A. and Salkie R. 1996. Comparing bilingual dictionaries with a parallel corpus. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 551-560.
86. Dobrovol'skij D. 1994. Idioms in a semantic network: towards a new dictionary-type. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex international Congress on Lexicography in Amsterdam. The Netherlands.*263-270.
87. Dodd S. 1989. Lexicomputing and the dictionary of the future. James, G. (Ed.), *Lexicographers and their Works.* 83-93 Exeter: University of Exeter Press.
88. Drysdale D. 1987. The role of examples in a learner's dictionary. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 213-223.
89. Dubois J. 1981. Evolution in dictionary design. *Applied Linguistics,* Vol. II Number 3, Autumn 1981, 236-249.
90. Dubois J. et Cl. 1971. *Introduction à la lexicographie*. Larousse, Paris.
91. Dubois-Charlier, F. 1988. Review of *Collins-Robert* French-English dictionary. *International Journal of Lexicography*, Vol. 1, 1:63.
92. Duval A. 1990. Nature et valeur de la traduction dans les dictionnaires bilingues. *Cahiers de Lexicologie*, 56-57: 27-33.
93. Ellis R. 1994. *The study of second language acquisition*, Oxford University Press, Oxford.
94. Euralex Colloquium, 1989. *Translation and Lexicography*, Benjamin, Amsterdam.
95. Fillmore Ch. J. 1989. Two Dictionaries. *International Journal of Lexicography.* Vol. 2, 1:57-83.
96. Foucault M. 1966. *Les mots et les choses*. Gallimard, Paris.
97. Fox G. 1987. The case for examples. Sinclair J. (Ed.) *Looking Up*, London and Glasgow: Collins ELT.
98. Francis G. 1993. A corpus-driven approach to Grammar. Principles, methods and examples. Baker, M. et al. (Eds.) 1993. *Text and technology: in honour of John Sinclair*, John Benjamins, Amsterdam/Philadelphia.
99. Francis G. et al. 1996. *Collins* COBUILD *Grammar Patterns 1: Verbs*, HarperCollins, Glasgow.
100. Gak V. G. 1992. De la langue-système à la réalisation discursive dans un

dictionnaire bilingue 'de type actif. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Part 2, 329-335.

101.Gimson A. C. 1981. Pronunciation in EFL dictionaries. *Applied Linguistics,* 2/3: 250-62.

102.Gleason H. A. Jr. 1967. The relation of lexicon and grammar. Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

103.Gold D. L. 1982. Review of the *University of Chicago Spanish Dictionary.* Reviewed Edition (1972). *Dictionaries,* 1982, 4:238-249.

104.Gold D. L. Review of *The Bantam New College Spanish and English Dictionary* (1986). *Dictionaries*, 1985, 8:293-316.

105.Green J. 1996. *Chasing the sun*, Henry Holt and Company, New York.

106.Grundy V. 1996. L 'utilisation d 'un corpus dans la fabrication des dictionnaires bilingues. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot, Louvain-la-Neuve,.

107.Haas M. R. 1967 What belongs in the bilingual dictionary? Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

108.Haiman J. 1980. Dictionaries and encyclopedias. *Lingua* 50 1980. 329-357.

109.Hale J. 1994. *The civilization of Europe in the Renaissance*, Atheneum, New York.

110.Halliday M. A. K. 1989. *Spoken And Written Language*, Geelong, Vi.: Deakin University Press; Oxford University Press, Oxford.

111.Halliday M. A. K. 1994. (second edition) An introduction to functional grammar, Edward Arnold, London.

112.Halliday M. A. K. and Hassan R. 1976. *Cohesion in English.* Longman, London.

113.Hanks P. 1979. To what extent does a dictionary definition define? Hartmann, R. R. K. (Ed.) *Dictionaries and their users.* Exeter Linguistic Studies 1 (4) : 32-8. University of Exeter.

114.Haraldson H. 1996. Problems of dictionary grammar. The Zaliznyak solution: a boon or a burden? *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 561-572.

115.Harrell R. S. 1967. Some notes on bilingual lexicography. Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

116.Hartmann R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings.* Niemeyer, Tübingen.

117.Hartmann R. R. K. 1980. *Contrastive textology*, Heidelberg, Groos.

118.Hartmann R. R. K. 1981a. Dictionaries, learners, users: some issues in lexicography: review article. *Applied Linguistics,* 2 (3) : 297-303.

119.Hartmann R. R. K. 1981b. Style values: Linguistic approaches and lexicographical practice. *Applied Linguistics,* Vol. II Number 3, Autumn 1981, 261-273.

120.Hartmann R. R. K. 1983. *Lexicography: Principles and Practice.* Academic Press,

London.

121.Hartmann R. R. K. 1987. Four perspectives on dictionary use: a critical review of research methods. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 11-28.

122.Hartmann R. R. K. 1987. Dictionaries of English: The user's perspective. Bailey, R. (ed.) 1987. *Dictionaries of English: Prospects for the Record of our Language.* Ann Arbor: University of Michigan Press, 121-135.

123.Hartmann R. R. K. 1988. Review of *Collins concise German-English Dictionary. International Journal of Lexicography*, Vol. 1, 1:67-69.

124.Hartmann R. R. K. 1992. Learner's references: from the monolingual to the bilingual dictionary. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere. Finland.* Part I, Studia Translatologica ser. A Vol. 2, 63-70.

125.Hartmann R. R. K. 1994. The use of parallel text corpora in the generation of translational equivalents for bilingual lexicography. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex International Congress on Lexicography in Amsterdam. The Netherlands,* 291-297.

126.Hartmann R. R. K. 1996. *Contrastive textology and corpus linguistics: on the value of parallel texts. Language Sciences*, Vol. 18, 3-4:947-957.

127.Harvey K. and Yuill D. 1992. The COBUILD testing initiative: the introspective, *encoding* component. Unpublished COBUILD paper.

128.Hatherall G. 1984. Studying Dictionary Use: Findings and Proposals. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen 183-198.

129.Hausmann F. J. and Gorbahn, A. 1989. COBUILD and LDOCE II, A Comparative Review. *International Journal of Lexicography*, Vol. 2, 1:44-56.

130.Hausmann F. J. et al. (Eds.) 1989. *International Encyclopedia of Lexicography.* W. de Gruyter, Berlin, New York.

131.Hayward T. and Moulin A. 1984. False friends invigorated. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 190-198.

132.Heid U. 1996. Creating a multilingual data collection for bilingual lexicography from parallel monolingual lexicons. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 573-590.

133.Herbst T. 1996. On the way to the perfect learner's dictionary: a first comparison of OALD5, LDOCE3, COBUILD2 and CIDE. *International Journal of Lexicography* Vol. 9, 4:321-357.

134.Herbst T. and Stein G. 1987. Dictionary-using skills: a plea for a new orientation in language teaching. Cowie, A. (Ed.) 1987 *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 115-127.

135.Heyn M. 1992. *Zur Wiederverwendung maschinenlesbarer Wörterbücher*, M. Niemeyer, Tübingen.

136.Hoenigswald H. M. 1967. Lexicography and grammar. Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

137.Householder F. W. and Saporta S. (Eds.) 1967. *Problems in lexicography. Indiana University*, Bloomington.

138.Hüllen W. and Haas R. 1992. Adrianus Junius on the Order of his NOMENCLATOR. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Part 2, 581-588.

139.Humblé Ph. 1996. Cobuild Student's Dictionary Bridge Bilingual Portuguese: a new bilingual dictionary concept. *Cahiers de Lexicologie*, 1996-1, 193-202.

140.Iannucci J. E. 1967. Meaning discrimination in bilingual dictionaries. Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

141.Iannucci J. E. 1985. Sense discriminations and translation complements in bilingual dictionaries. *Dictionaries*, 1985, 7:57-65.

142.Ilson R. (Ed.) 1987. A Spectrum of Lexicography. *Papers from AILA BRUSSELS 1984*, John Benjamins, Amsterdam/Philadelphia.

143.Ilson R. 1985. *Dictionaries, Lexicography and Language Learning,* Oxford University Press, Oxford.

144.Ilson R. 1987. Illustrations in dictionaries. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 193-212.

145.International Conference on Lexicography, 1983. Lexeter '83, Proceedings, Max Niemeyer, Tübingen.

146.Jain M. 1981. On meaning in the foreign language learner's dictionary. *Applied Linguistics*, Vol. II Number 3, Autumn 1981, 274-286.

147.Jakobson R 1959. On linguistic aspects of translation. *Selected Writings*, II. Mouton, The Hague.

148.James G. 1989. *Lexicographers and their Works.* Dictionary Research Centre, University of Exeter.

149.Jansen, J. et al. 1987. Controlling LDOCE's controlled vocabulary. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 78-94.

150.Kasimi A. M. 1977. *Linguistics and bilingual dictionaries.* Brill, Leiden.

151.Keisuke N. 1988. English-Japanese learner's dictionaries. *International Journal of Lexicography*, Vol. 1, 1:295-314.

152.Kernerman L. 1994. A 3-part, fully trilingual tri-directional dictionary. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex international Congress on*

*Lexicography in Amsterdam. The Netherlands.*472-478.

153.Kernerman L. 1996. English learner's dictionaries: how much do we know about their use? *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 405-414.

154.Kharma N. N. 1984. Contextualization and the bilingual dictionary. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen 199-206.

155.Kibbee D. A. 1985. Progress in bilingual lexicography during the Renaissance. *Dictionaries*, 7:21-31.

156.Kilgariff A. 1997. Putting frequencies in the dictionary. *International Journal of Lexicography*, Vol. 10, 2:135-155.

157.Kipfer B. A. 1984. *Workbook on Lexicography*, Dictionary Research Centre, University of Exeter.

158.Kipfer B. A. 1987. Dictionaries and the intermediate student: communicative needs and the development of user reference skills, Cowie, A. (Ed.) 1987 *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 44-54.

159.Kirchmeier-Andersen S., Sandford Pedersen B. Schøsler L. 1994. Combining semantics and syntax in monolingual dictionaries. Attacking the enemy from both flanks. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex international Congress on Lexicography in Amsterdam. The Netherlands.*137-146.

160.Knowles F. 1996. L'informatisation de la fabrication des dictionnaires bilingues. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot, Louvain-la-Neuve,.

161.Krashen S. 1996. *Notes on a polyglot: Kato Lomb. System*, Vol. 24, 2:207-210.

162.Krishnamurthy R 1987. The process of compilation. Sinclair J (Ed.) *Looking Up*, London and Glasgow: Collins ELT.

163.Krishnamurthy R. 1996. The data is the dictionary: Corpus at the cutting edge of Lexicography. *Papers in computational lexicography. Complex '96*, Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, 117-144.

164.Kromann H. et al. 1984. 'Active' and 'passive' bilingual dictionaries: the Scerba concept reconsidered. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen 207-215.

165.Kromann H-P 1990. Selection and presentation of translational equivalents in monofunctional and bifunctional dictionaries, *Cahiers de Lexicologie*, 56-57: 17-26.

166.Kuhn S. T. 1970. *The structure of scientific revolutions.* University of Chicago Press, Chicago.

167.Landau S. I. 1984. *Dictionaries. The art and craft of lexicography.* Scribner Press, New York.

168.Larsen-Freeman D. and Long M. H. 1991. *An introduction to Second Language*

Acquisition Research. Longman, London.

169.Laufer B. 1992. Corpus-based versus lexicographer examples in comprehension and production of new words. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Finland. Part I, Studia Translatologica ser. A Vol. 2, 71-76.

170.Laufer B. 1993. The effect of dictionary definitions and examples on the use and comprehension of new L2 words. *Cahiers de Lexicologie*, 63:131-142.

171.Laufer B. and Melamed L. 1994. Monolingual, bilingual and 'bilingualised' dictionaries: which are more effective, for what and for whom? Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex International Congress on Lexicography in Amsterdam. The Netherlands,* 565-576.

172.Laufer B. and Nation 1995. Vocabulary size and use: Lexical richness in L2 production. *Applied Linguistics*, Vol. 16, 3:307-322.

173.Laufer B. Appropriation du vocabulaire: les mots faciles, les mots difficiles et les mots impossibles. Acquisition et interaction en Langue Etrangère, 3(1993-94), 97-114.

174.Lepschy G. 1984. Léxico. Romano R. (Ed.) 1984. *Enciclopédia Einaudi* Vol. 2 'Linguagem-Enunciação' Imprensa Nacional Casa da Moeda.

175.Lorentzen H. 1996. Lemmatisation of multi-word lexical units: in which entry? *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 415-422.

176.Lovatt E. A. 1984. Illustrative examples in a bilingual colloquial dictionary. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen 216-220.

177.Macklovitch E. 1996. Les dictionnaires bilingues en-ligne et le poste de travail du traducteur. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot, Louvain-la-Neuve, 169-180.

178.Magay T. 1984. Technical or general: problems of vocabulary selection in a medium size bilingual dictionary. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen 221-225.

179.Maingay S. and Rundell, M. Anticipating learner's errors — Implications for dictionary writers, in: Cowie, A. (Ed.) 1987 *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 128-135.

180.Malkiel Y. 1967. A typological classification of dictionaries on the basis of distinctive features. Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

181.Malone K. 1967. Structural linguistics and bilingual dictionaries. Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

182.Marello C. 1987. Examples in contemporary Italian bilingual dictionaries. Cowie, A. 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Max Niemeyer, Tübingen, 224-237.

183.Marello C. 1996. Les différents types de dictionnaires bilingues. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot, Louvain-la-Neuve,, 31-52.

184.Martín L. 1984. Lexical fields and stepwise lexical decomposition in a contrastive English-Spanish verb valency dictionary. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 226-236.

185.Martin S. E. 1967. Selection and presentation of ready equivalents in a translation dictionary. Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

186.Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex 1994, Proceedings, Papers Submitted to the 6th Euralex international Congress on Lexicography in Amsterdam, The Netherlands.*

187.Mdee J. S. 1984. Constructing an entry in a learner's dictionary of standard Swahili. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 237-241.

188.Meara P. and Nesi H. 1994. Patterns of misinterpretations in the productive use of EFL dictionary definitions. *System*, Vol. 22, 1:1-15.

189.Mel'cuk I. A. and Wanner L. 1994. Towards an efficient representation of restricted lexical co-occurrence. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex international Congress on Lexicography in Amsterdam. The Netherlands.*325-338.

190.Mel'cuk I., Clas A. and Polguère A. 1995. *Introduction à la Lexicologie explicative et combinatoire*. Éditions Duculot, Louvain-la-Neuve,.

191.Mel'cuk I. et al. 1984, 1988. *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexicosémantiques I et II.* Montréal, Les Presses de l'Université de Montréal.

192.Minaeva L. 1992. Dictionary examples: friends or foes? *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Finland. Part I, Studia Translatologica ser. A Vol. 2, 77-80.

193.Montemagni S. Federici Stefano and Pirelli V. 1996. Example-based word sense disambiguation: a paradigm driven approach. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 151-160.

194.Moon R. (forthcoming) Vocabulary connections: multi-word items in English. McCarthy M. J. & Schmitt N. (Eds.) *Second Language Vocabulary,* Cambridge University Press, Cambridge.

195.Moon R. 1987. Monosemous words and the dictionary. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.*

Niemeyer, Tübingen, 173-182.

196.Moon R. 1989. Objective or objectionable: Ideological aspects of dictionaries. *English Language Research* 3, Birmingham. 59-94.

197.Moon R. 1992. There is reason in the roasting of eggs: a consideration of fixed expressions in native-speaker dictionaries. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Finland. Part I, Studia Translatologica ser. A , 493-502.

198.Moon R. 1996. Data, descriptions and idioms in corpus lexicography. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 245-256.

199.Moulin A. 1983. The LSP learner's lexicographical needs. Hartmann R. R. K. 1983. *Lexicography: Principles and Practice. Academic Press, London,*, 145-152.

200.Mounin, G 1963 *Les problèmes théoriques de la traduction.* Gallimard, Paris.

201.Nesi H. 1991. English Dictionaries in the Classroom: a critical survey. *CLIE Working Papers* Number 12 ISSN 0964-8275.

202.Nesi H. 1994. The effect of language background and culture on productive dictionary use. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex International Congress on Lexicography in Amsterdam. The Netherlands,* 578-585.

203.Nesi H. 1996. The role of illustrative examples in productive dictionary use. *Dictionaries* , 17:198-206.

204.Nesi H. 1996a. Defining a shoehorn: the success of learner's dictionary entries for concrete nouns. Paper submitted to *Lexicographica Special Edition* (papers from AILA Symposium on Dictionary Use, 1996) B. Atkins (ed).

205.Nesi H. 1997 For future reference: A review of current electronic learner's dictionaries.-to appear. *System.*

206.Nguyen D. H. 1980-81. Teaching culture through bilingual dictionaries. *Dictionaries*, 2-3:57-68.

207.Niedzielski H. 1984. Semantic considerations of 'set' and some of its French equivalents. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 242-252.

208.Nistrup Madsen B. 1990. Computerised registration of the structure of a printed bilingual dictionary and the establishment of a database. *Cahiers de Lexicologie*, 56-57: 211-225.

209.Noël B. Devos F. and Defrancq B. 1996. Towards a more grammatical bilingual dictionary. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg, Sweden*, Göteborg University, Department of Swedish, 597-608.

210.Nuccorini S. 1992. Monitoring Dictionary Use. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in*

Tampere. Finland. Part I, Studia Translatologica ser. A Vol. 2, 89-102.

211.Nuccorini S. 1994. On dictionary misuse. Martin W. Meys W. Moerland M. Plas ten E. Sterkenburg van P. Vossen P. (eds.) 1994. *Euralex 1994, Proceedings, Papers Submitted to the 6 th Euralex International Congress on Lexicography in Amsterdam, The Netherlands* 586-597.

212.O'Connor D. 1990. *A History of Italian and English bilingual Dictionaries*, Leo S, Olschki, Firenze.

213.Ogasawara L. 1984. Problems in improving Japanese-English learner's dictionaries. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 253-257.

214.Osselton N. E. 1983. On the history of dictionaries. The history of English-Language dictionaries. Hartmann R. R. K. 1983. *Lexicography: Principles and Practice. Academic Press, London,*, 13-21.

215.Osselton N. E. 1996. The dictionary label 'Figurative ': Modern praxis and the origin of a tradition. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 239-250.

216.Otani Y. 1996. Cross language equivalence: between lexical and translation equivalents in the case of English-Japanese dictionaries. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 609-614.

217.Pedersen V. H. 1984. Reflections on the treatment of prepositions in bilingual dictionaries, and suggestions for a statistical approach. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 258-267.

218.Peters C. Federici S. Montemagni S. Calzolari N. 1994. From machine readable dictionaries to lexicons for NLP: the COBUILD dictionaries-a different approach. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex international Congress on Lexicography in Amsterdam. The Netherlands.*147-157.

219.Popper K. 1994. *In search of a better world. Lectures and essays from thirty years.* Translated by Laura J. Bennett, Routledge, London and New York.

220.Popper K. 1994b. *Knowledge and the body-mind problem. In defence of interaction.* Ed. by M. A. Notturno, Routledge, London and New York.

221.Popper K. 1995. (ed. 1959) *The logic of scientific discovery.* The University of Chicago Press, Chicago.

222.Quemada B. 1990. Les données lexicographiques et l'ordinateur' *Cahiers de Lexicologie*, 56-57: 171-189.

223.Rasmussen J. 1990. Sélection des entrées pour un grand dictionnaire bilingue danois-français. *Cahiers de Lexicologie*, 56-57: 161-170.

224.Renouf A. J. 1983. Corpus development at Birmingham University. Aarts J. and Meijs W. (Eds.) *Corpus Linguistics: recent developments in the use of computer corpora in English language research.* Rodopi, Amsterdam.

225.Rey 1982. A. Encyclopédies et dictionnaires, Presses Universitaires de France, Paris.

226.Rivers W. 1975. *A metodologia do ensino de línguas estrangeiras*. Trad. Marchi H. S. Livraria Pioneira Editora, São Paulo.

227.Riverside Conference on Lexicography 1969, 1972. *New Aspects of Lexicography*, Carbondale, Southern Illinois University Press.

228.Roberts R. 1996. Le traitement des collocations et des expressions idiomatiques dans les dictionnaires bilingues. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot, Louvain-la-Neuve, 181-198.

229.Roberts R. and Montgomery C. 1996. The use of corpora in bilingual lexicography. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 457-464.

230.Rosenstein R. 1985. Jean Nicot's Thrésor and Renaissance multilingual lexicography. *Dictionaries,* 7:33-56.

231.Rousseau J. J. 1973. *Confessions*. Gallimard, Folio, Paris.

232.Rousseau J. J. 1966. *Émile*, GF Flammarion, Paris.

233.Särkkä H. 1984. Improving the usability of the Finnish comprehension dictionary. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 268-273.

234.Schorr G. 1987. Deux types de dictionnaires bilingues de poche. Ilson, R. (Ed.) *A Spectrum of Lexicography. Papers from AILA BRUSSELS 1984*, John Benjamins, Amsterdam/Philadelphia 91-99.

235.Sciarone A. G. and Ahmadi M. R. 1996. Towards a universal dictionary. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 465-472.

236.Scott D. S. 1987. What should we demand of electronic dictionaries? *The uses of large text databases. Proceedings of the conference. Third Annual Conference of the University of Waterloo Centre for the New Oxford English Dictionary.* November 9-10, 1987, Waterloo, Canada, 91-112.

237.Sharpe A. 1988. Pragmatic considerations for an English-Japanese dictionary. *International Journal of Lexicography*, Vol. 2, 4:315-323.

238.Sinclair J. M. 1984. Lexicography as an Academic Subject. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings*, Niemeyer, Tübingen 3-12.

239.Sinclair J. M. (Ed.) 1987. *Looking Up*, Collins, London and Glasgow.

240.Sinclair J. M. 1988. Naturalness in Language. McCarthy, M. (Ed.) 1988. *Naturalness in Language*. English Language Research Journal Vol. 2. University of Birmingham.

241.Sinclair J. M. 1991. *Corpus, Concordance, Collocation*, Oxford University Press, Oxford.

242.Sinclair J. M. 1993. Written Discourse Structure. Sinclair, J. M. 1993 *Papers in Written Discourse Analysis*, London: Routledge.

243.Sinclair J. M. 1996. Multilingual databases. An international project in multilingual lexicography. *International Journal of Lexicography*, Vol. 9, 3:179-196

244.Singh R. A. 1982. *An introduction to lexicography*, Mysore, Central Institute of Indian Languages.

245.Snell-Hornby M. 1984. The bilingual dictionary — help or hindrance? Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen 274-281.

246.Snell-Hornby M. 1987. Towards a learner's bilingual dictionary. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 159-170.

247.Soelberg N. 1990. Les informations du dictionnaire bilingue. Equivalents ou champs sémantiques? *Cahiers de Lexicologie*, 56-57: 51-61.

248.Stein G. 1985. *The English Dictionary before Cawdrey*, Niemeyer, Tübingen.

249.Steiner G. 1971. *After Babel: Aspects of Language and Translation.* Oxford University Press, Oxford.

250.Steiner R. J. 1984. Guidelines for reviewers of bilingual dictionaries. *Dictionaries*, 5:166-181.

251.Steiner R. J. 1989. The absence of text: the bilingual dictionary as an index. *International Journal of Lexicography*, Vol. 2, 3:249-257.

252.Steiner R. J. 1990. Review of: Colin Smith et al. *Collins Spanish-English / English-Spanish dictionary.* Second edition, London, Collins. *International Journal of Lexicography*, Vol. 3, 1:65-71.

253.Stock 1984. Polysemy. Hartmann, R. R. K. (Ed.) *LEXeter '83 Proceedings*: 131-40. Niemeyer, Tübingen.

254.Summers D. 1988. The role of dictionaries in language learning. Carter R. and McCarthy M. 1988. *Vocabulary and Language Teaching*. Longman, London and New York.

255.Svensén B. 1983. A computerised concordance based on a bilingual dictionary: a case study. Hartmann R. R. K. 1983. Lexicography: Principles and Practice. Academic Press, London, 296-208.

256.Svensén B. *Practical Lexicography-Principles and Methods of Dictionary Making*, Oxford University Press, Oxford.

257.Swanson D. C. 1967. Recommendations on the selection of entries for a bilingual dictionary. Householder F. W. and Saporta S. (Eds.), *Problems in lexicography. Indiana University*, Bloomington.

258.Szende T. 1996. Problèmes d'équivalence dans les dictionnaires bilingues. Béjoint H. and Thoiron Ph. (Eds.) 1996. *Les dictionnaires bilingues*, Aupelf-Uref-Éditions Duculot, Louvain-la-Neuve,.

259.Taylor A. and Chan A. 1994. Pocket electronic dictionaries and their use. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994.

Euralex 1994, Proceedings, Papers Submitted to the 6 th Euralex international Congress on Lexicography in Amsterdam, The Netherlands 598-605.

260.Thompson S. A. 1984. Bilingual dictionaries and the notion of 'lexical categories' in Chinese. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen 282-288.

261.Tomaszczyk J. 1981. Bilingual pedagogical lexicography. *Applied Linguistics* Vol. II Number 3, Autumn 1981, 287-296.

262.Tomaszczyk J. 1983. *On bilingual dictionaries.* Hartmann R. R. K. 1983. *Lexicography: Principles and Practice.* Academic Press, London.

263.Tomaszczyk J. 1984. The culture-bound element in bilingual dictionaries. Hartmann, R. R. K. (ed. 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen 289-297.

264.Tomaszczyk J. 1987. FL 'Learner's communication failure: implications for pedagogical lexicography. Cowie, A. (Ed.) 1987. *The dictionary and the language learner. Papers from the Euralex Seminar at the University of Leeds.* Tübingen, Niemeyer, 136-145.

265.Tommola H. et al. (Eds.) 1992. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Finland. Part I, Studia Translatologica ser. A, Vol. 2, Tampere.

266.Traupman J. C. 1980-81. Review of Collins German-English dictionary (1980). *Dictionaries,* 2-3:162-166.

267.van der Meer G. 1996. The treatment of figurative meanings in the English learner's dictionaries. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 423-430.

268.Varantola K. 1994. The dictionary user as decision maker. Martin W., Meys W., Moerland M., Plas ten E., Sterkenburg van, Vossen (Eds.) 1994. *Euralex '94 Proceedings. Papers Submitted to the 6th Euralex International Congress on Lexicography in Amsterdam. The Netherlands,* 606-611.

269.Veisbergs A. 1996. False friends dictionaries: a tool for translators or learners or both. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 627-634.

270.Walter E. 1996. Parallel development of monolingual and bilingual dictionaries for learners of English. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 635-644.

271.Wexler J. 1987. Affix-entries in bilingual dictionaries. *Cahiers de Lexicologie*, 50:237-243.

272.Wierzbicka A. 1992-93. What are the uses of theoretical lexicography? *Dictionaries,* 14:44-78.

273.Wittgenstein L. 1963. *Tractatus logico-philosophicus.* Suhrkamp, Frankfurt-am-

Main.

274. Wittgenstein L. 1981. *Philosophical investigations*. translated by G. E. M. Anscombe, Basil Blackwell.

275. Wittgenstein L. 1984. *Über Gewißheit. Zettel. Vermischte Bemerkungen.* Suhrkamp, Frankfurt am Main.

276. Wood, E. R. 1979. Review of Ali M. Kasimi *Linguistics and bilingual dictionaries. Dictionaries* 1979, n°1, p.163-164

277. Wouden van der T. 1992. Prolegomena to a multilingual description of collocations. *Euralex '92 Proceedings I-II. Papers submitted to the 5th Euralex international Congress on Lexicography in Tampere.* Part 2, 449-456.

278. Zgusta L. 1971 *Manual of lexicography*. Mouton, The Hague.

279. Zgusta L. 1980. Some remarks on the context of lexicography. Zgusta L. 1980. (Ed.) *Theory and method in lexicography*, Hornbeam Press, Columbia, 3-29.

280. Zgusta L. 1984. Translational Equivalence in the Bilingual Dictionary. Hartmann, R. R. K. (Ed.) 1984. *LEXeter '83 Proceedings,* Niemeyer, Tübingen.

281. Zgusta L. 1987. Equivalence in a bilingual dictionary: Bahukosyam. *Dictionaries,* 9:1-47.

282. Zgusta L. 1988. *Lexicography Today*, Series Maior Lexicographica, Niemeyer, Tübingen.

283. Zgusta L. 1996. The lexicographer's creativity. *Euralex '96 Proceedings. Papers submitted to the seventh Euralex international congress on lexicography in Göteborg*, Sweden, Göteborg University, Department of Swedish, 323-336

284. Zöfgen E. 1994. *Lernerwörterbücher in Theorie und Praxis*, Max Niemeyer, Tübingen.

APPENDIXES FOR 'EXAMPLES IN LEARNER'S DICTIONARIES.

Appendix 1 'Road' [25]

COBUILD I

1. Cross the main road, then go on down the lane to the village.
They took the road that led up the hill.
.the road from Belfast to Londonderry.
She was studying a road map when I got back into the car.
By road Luxembourg is about 225 miles from the ferry.
The ancient ruins were accessible by road.
2. There is an antique shop at the top of my road.
.the quiet Edgbaston road where he had lived for some thirty years.
The museum was in a side street leading off from a road of shops.
3. The hotel was just a little farther along the road.
There are shops just down the road.
.her cousins from across the road.
I was talking to Mr Marks from along the road.
.Janet from down your road.
4.159, London Road.
This is the Oxford road, and the Watford road turns off to the right.
They crossed the Yugoslav border on the Budapest-Zagreb road.
5. The number of road accidents was greatly reduced.
6. I was again on the road, again at the wheel of the old blue sedan.
I was stiff after seven hours on the road.
7. It could well be the only one of its type still on the road.
8. We'll take the play for a few weeks on the road before it opens in London.
9. She was well on the road to recovery.
Surely you're on the road to recognition, even if it's only as head of department.
10. New information is probably the surest road to new ideas.
We have science and technology to help us along the road to peace and plenty.
This is the first step on the road to victory.
11. Let's have one for the road before we go.

COBUILD II

1. There was very little traffic on the roads...
We just go straight up the Bristol Road...
He was coming down the road the same time as the girl was turning into the lane...
Buses carry 30 per cent of those travelling by road...
You mustn't lay all the blame for road accidents on young people.
2. We are bound to see some ups and downs along the road to recovery.
3. I was relieved to get back in the car and hit the road again.
4. He still hoped someday to get a new truck and go back on the road.
5. The government took another step on the road to political reform.
...the stunning fashion pictures which launched unknown teenager Jane March on the road to stardom.

[25] Bolds and italics are mine.

Longman

1. a busy road
at the end of the road
We live just down the road.
It takes three hours by road.
Take the main road out of town and turn left at the first light.
He was killed in a road accident.
Kids of that age have no road sense.
A road safety campaign.
2. Maple Road.
3. I've been on the road since 5:00 a.m. this morning.
It costs a lot of money to keep these old cars on the road.
4. It was this deal that set him on the road to his first million.
5. You could move your pension to a private scheme, but I wouldn't advise going down that road.

Oxford Advanced Learner's Dictionary

1. The road to Bristol / the Bristol road
main/major/minor roads
a quiet suburban road
2. a road-map of Scotland
be considerate to other road-users.
3. 35 York Rd, London SW16.
4. the Southampton Roads.
5. all roads lead to Rome.
6. It's a long way by road – the train is more direct
It's cheaper to ship goods by road than by rail.
7. get the show on the road.
8. hit the road.
9. one for the road.
10. The band has been on the road for almost a month.
11. the road to success/ruin.
12. The road to hell is paved with good intentions.
13. take to the road.

Oxford Wordpower

1. Is this the right road to Beckley?
Take the London road and turn right at the first roundabout.
Turn left off the main (= big, important) road.
major/minor roads.
If you get onto the ring road you'll avoid the town centre.
road signs.
a road junction.
a road-map of England.
Bayswater Road, London.
60 Marylebone Road, London.
It's going to be a terrible journey by road – let's take the train.
We were on the road for 14 hours.

Appendix 2 'Collocates in Cambridge and COBUILD II'

face

Cambridge

They face/are faced with financial penalties.
We'll have to face her with this new information and see what she has to say.
I can't face climbing those stairs again.
He'll have to face the music when his parents find out he's been missing school.
Their houses face each other across the street.
We'll have to face the facts and start cutting costs.
He thinks he would lose face if he admitted the mistake.
She tried to save face by inventing a story about being overseas at the time.
They gave him the title of company president as a face-saving gesture, although he no longer had any power.
They agreed that there should be no attempts at face-saving.
The hospital charges £4000 for a full facelift.
The bank is planning to give its 1930s building a complete facelift.

collocate	occ.
down	0
now	0
let	0
value	0
look	0
smile	0
changes	0
man	0
see	0
problems	0
turned	0
eyes	0
across	0
back	0
off	0
hands	0
put	0
still	0
made	0
looked	0
total	0

collocates/example 0/12=0

stopword	occ.
the	2
to	4
of	0
a	3
and	1
in	0
his	0
's	0
her	1
on	0
total	11

stopwords/example 11/12=0.91

face

COBUILD

He rolled down his window and stuck his face out.
A strong wind was blowing right in my face.
He was going red in the face and breathing with difficulty.
She had a beautiful face.
He was walking around with a sad face.
The priest frowned in the light, his face puzzled.
Harrer was one of the first to climb the north face of the Eiger.
He scrambled 200 feet up the cliff face.
The changing face of the British country side.
This would change the face of Malaysian politics.
Brothels, she insists, are the acceptable face of prostitution.
With the collapse of communism, the ugly face of capitalism to some extent is appearing again.
England doesn't want a war but it doesn't want to lose face.
To cancel the airport would be a loss of face for the present governor.
Children have an almost obsessive need to save face in front of their peers.
Can't you see this could blow up in your face?
You can criticise him until you're blue in the face, but you'll never change his personality.
All the time Stephen was lying face down and unconscious in the bath tub.
Charles laid down his cards face up.
No human being on the face of the earth could do anything worse than what he did.
If a nuclear war breaks out, every living thing will be wiped off the face of the Earth.
We were strolling into the town when we came face to face with Jacques Dubois.
It was the first face to face meeting between the two men.
Eventually, he came face to face with discrimination again.
I was gradually being brought face to face with the fact that I had very little success.
Scientific principles that seem to fly in the face of common sense.
He said that the decision flew in the face of natural justice.
The Prime Minister has called for national unity in the face of the violent anti-government protests.
Roosevelt was defiant in the face of the bad news.
With juveniles under eighteen, there's little we can do. We can't keep them in custody. They just laugh in your face.
He came to me with a very long face.
Opening the door, she made a face at the musty smell.
Katrhyn pulled a face at Thomasina behind his back.
On the face of it that seems to make sense. But the figures don't add up.
It is, on the face of it, difficult to see how the West could radically change its position.
Friends will see you are putting on a brave face and might assume you've got over your grief.
Scientists are putting a good face on the troubles.
This Government has set its face against putting up income tax.
If she shows her face again back in Massachusetts she'll find a warrant for her arrest waiting.
I felt I ought to show my face at her father's funeral.
What went through Tom's mind I can't imagine, but he did manage to keep a straight face.
You have to wonder how anyone could say that seriously and with a straight face.
Her opponent called her a liar to her face.
Relief and gratitude were written all over his face.
I could just see the pain written across her face.
Stand up. Face the wall.
He was hauled in to face the judge.
Although your heart is breaking, you must face the truth that a relationship has ended.
He accused the Government of refusing to face facts about the economy.
I have grown up now and have to face up to my responsibilities.
They were having to face up to the fact that they had lost everything.

I couldn't face the prospect of spending a Saturday night there, so I decided to press on.
My children want me with them for Christmas Day, but I can't face it.
I couldn't face seeing anyone.
She was always attracted to younger men. But, let's face it, who is not?
Nothing gives a room a faster facelift than a coat of paint.
The decision seems to be a face-saving compromise which will allow the government to remain in office.

Collocate	occ.
down	1
now	0
let	0
value	0
look	0
smile	0
changes	0
man	0
see	(1)
problems	0
turned	0
eyes	0
across	1
back	1
off	1
hands	0
put	0
still	0
made	1
looked	0
total	4

collocates/example 4/66=0,06

stopword	occ.
the	31
to	14
of	14
a	11
and	6
in	10
his	4
's	0
her	3
on	5
total	98

stopwords/example 98/66=1.48

fact

Cambridge

I'm not interested in hopes and plans, I just want you to tell me the plain/bare facts.
Can I regard what you have just told me as fact?
The play was closely based on fact.
The fact is that they are the stronger team and are sure to win.
No, I don't work. In fact, I've never had a job.
Have you always lived here? As a matter of fact I've only lived here for the last three years.
Going bald is a fact of life.
We are getting some facts and figures together and we will then have a full board meeting, and hopefully make a decision.

Collocates	occ.
despite	0
matter	1
no	0
very	0
only	1
some	1
many	0
people	0
most	0
think	0
life	1
even	0
much	0
(erm)	0
never	1
quite	0
both	0
finding	0
remains	0
women	0
total	5

collocates/example 5/8=0.62

stopword	occ.
the	
in	1
that	1
is	2
it	0
was	0
I	6
he	0
they	1
are	3
total	14

stopwords/example 14/8=1.75.

fact

COBUILD

His chances do not seem good in view of the fact that the Chief Prosecutor has already voiced his public disapproval.
Despite the fact that the disease is so prevalent, treatment is still far from satisfactory.
No amount of encouragement can hide the fact that talking about very personal issues with a stranger is intimidating.
In Rome, meeting him every morning, he soon became aware of the fact that Erter was ill.
My family now accepts the fact that I don't eat sugar or bread.
The fact that he had left her of his own accord proved to me that everything he'd said was true.
We've had a pretty bad time while you were away. In fact, we very nearly split up this time.
He apologised as soon as he realised what he had done. In actual fact, he wrote a nice little note to me.
Mr Major didn't go to university. In fact he left school at 16.
That sounds rather simple, but in fact it's very difficult.
They complained that they had been trapped inside the police station, but in fact most were seen escaping over the adjacent roofs to safety in nearby buildings.
Why had she ever trusted her? In point of fact she never had, she reminded herself.
A statement of verifiable historical fact.
How much was fact and how much fancy no one knew.
There is so much information that you can almost effortlessly find the facts for yourself.
His opponent swamped him with facts and figures.
The lorries always left for China in the dead of night when there were few witnesses around to record the fact.
The local people saw all the sufferings to which these deportees were subjected. And as a matter of fact, the local people helped the victims of these deportations.
'I guess you haven't eaten yet'. 'As a matter of fact, I have', said Hunter.
I know for a fact that baby corn is very expensive in Europe.
I know for a fact that Graham has kept in close touch with Graham.
The fact is blindness hadn't stopped the children doing many of the things that sighted children enjoy.
I found that election rallies were being very poorly attended. But the fact of the matter is that they're not terribly interested in this election.
The fact remains, however you measure it, is unacceptably high.
His admirers claim that he came to power perfectly legally, but the fact remains that he did so by exploiting an illegal situation.
We aren't playing well as a team, and that's a fact.
He's a dull writer and that's a fact.
'I'm still staff colonel'. — 'Is that a fact?'.
A UN fact-finding mission is on its way to the region.
Stress is a fact of life from time to time for all of us.
There comes a time when children need to know more than the basic facts of life.

collocates	occ.
despite	1
matter	3
no	0
very	3
only	0
some	0
many	0
people	0
most	1
think	0
life	2
even	0
much	2
(erm)	0
never	1
quite	0
both	0
finding	1
remains	2
women	0
total	16

collocates/example 16/31=0.51

stopword	occ.
the	
in	6
that	11
is	4
it	0
was	2
I	4
he	3
they	0
are	0
total	30

stopwords/example 30/31=0.96

failed

Cambridge

I tried to persuade him to come, but I failed.
She moved to London in the hope of finding work as a model, but failed.
He failed dismally/miserably in his attempt to break the record.
The two sides in the negotiation have failed to come to an agreement.
She failed to reach the Wimbledon Final this year.
He's a failed writer.
After two failed marriages, he is planning to marry for a third time.
She has been given the task of sorting out the government's failed taxation policy.
He promised to help, but he failed to arrive on time.
Her parents failed to understand that there was a problem.
The club had been promised a grant from the council, but the money failed to materialize.
He never failed to take a disapproving stand at people who got divorced.
He failed her when she most needed him.
When I looked down and saw how far I had to jump, my courage failed me.
'Did you pass?' 'No, I failed'.
I passed in history but failed in chemistry.
The examiners failed him because he hadn't answered enough questions.
The brakes failed and the car crashed into a tree.
After talking non-stop for two hours, her voice failed.
The wheat failed last year because of the lack of rain.

Collocates	occ.
because	2
coup	0
make	0
last	1
get	0
any	0
government	0
win	0
reach	1
attempt	1
take	1
find	0
far	0
also	0
having	0
tried	1
test	0
yesterday	0
again	0
even	0
total	7

collocates/example 7/28=0.25.

stopwords	occ.
to	6
the	8
have	1
has	0
that	0
had	0
he	6
but	5
they	0
his	0
total	26

stopwords/example 26/20=1.3.

failed

COBUILD

The Worker's Party failed to win a single governorship.
He failed in his attempts to take control of the company.
Many of us have tried to lose weight and failed miserably.
The truth is, I'm a failed comedy writer really.
We tried to develop plans for them to get along, which all failed miserably.
After a failed military offensive, all government troops and police were withdrawn from the island.
He failed to file tax returns for 1982.
The bomb failed to explode.
The lights mysteriously failed, and we stumbled around in complete darkness.
In fact many food crops failed because of the drought.
So far this year, 104 banks have failed.
a failed hotel business.
Here in the hills, the light failed more quickly.
communities who feel that the political system has failed them.
For once, the artist's fertile imagination failed him.
Their courage failed a few steps short and they came running back.

Collocates	occ.
because	1
coup	0
make	0
last	0
get	0
any	0
government	1
win	1
reach	0
attempt	1
take	1
find	0
far	1
also	0
having	0
tried	2
test	0
yesterday	0
again	0
even	0
total	8

collocates/example 8/16=0.5.

stopwords	occ.
to	4
the	7
have	2
has	1
that	1
had	1
he	2
but	0
they	0
his	1
total	19

stopwords/example 19/16=1.18.

fail

Cambridge

The reluctance of either side to compromise means that the talks are doomed to fail.
You couldn't fail to be saddened by the distressing reports on the famine victims.
I fail to see what you're getting at.
I fail to see what this has to do with the argument.
Be there by ten o'clock without fail.
Every morning, without fail, she used to sit in the park and read her newspaper.
A lot of people fail their driving test the first time.
They had a track record of success and they never imagined the business could fail.
a fail-safe device/mechanism/system

Collocate	occ.
because	1
see	2
many	0
people	1
without	2
make	0
test	1
often	0
get	1
even	0
talks	1
meet	0
take	0
understand	0
might	0
too	0
safe	1
likely	0
never	1
succeed	0
total	11

collocates/example 11/9=1.2.

stopword	occ.
to	4
if	0
they	1
that	0
will	0
you	1
we	0
but	0
would	0
not	0
total	6

stopwords/example 6/9=0.66

fail

COBUILD

He was afraid the revolution they had started would fail.
Some schools fail to set any homework.
We waited twenty-one years, don't fail us now.
It's the difference between a pass and a fail.
That's how it was in my day and I fail to see why it should be different now.
He attended every meeting without fail.
On the 30th you must without fail hand in some money for Alex.
Tomorrow without fail he would be at the old riverside warehouse.

Collocate	occ.
because	0
see	1
many	0
people	0
without	3
make	0
test	1
often	0
get	0
even	0
talks	0
meet	0
take	0
understand	0
might	0
too	0
safe	0
likely	0
never	0
succeed	0
total	5

collocates/example 5/8=0.62.

stopword	occ.
to	2
if	0
they	1
that	0
will	0
you	1
we	1
but	0
would	2
not	0
total	

stopwords/example 7/8=0.87.

fade

Cambridge

You'll fade that tablecloth if you wash it in hot water.
If you hang your clothes out in the bright sun, they will fade.
He had a lovely suntan when he got back from his holiday, but it soon faded.
They arrived home just as the light was fading.
After his girlfriend left him, Bill faded from the picture/scene.
The horse riders gradually faded from view/sight.
Day slowly faded into night.
The voice on the radio faded out.
The children's memories of their fathers slowly faded away.
Hopes of saving the trapped miners are fading away fast.
He was wearing a pair of faded jeans and an old T-shirt.
A faded beauty is a woman who was beautiful in the past.

Collocate	occ.
away	2
under	0
began	0
just	1
beginning	0
before	0
never	0
die	0
hopes	1
quickly	0
soon	1
begun	0
flowers	0
memories	1
begins	0
total	6

collocates/example 6/12=0.5.

stopword	occ.
to	0
and	1
will	1
as	1
out	1
they	1
into	1
or	0
then	0
total	6

stopwords/example 6/12=0.5.

fade

COBUILD

All colour fades – especially under the impact of direct sunlight.
No matter how soft the light is, it still plays havoc, fading carpets and curtains in every room.
fading portraits of the Queen and Prince Philip.
a girl in a faded dress.
faded painted signs on the sides do some of the buildings.
Seaton lay on his bed and gazes at the ceiling as the light faded.
The sound of the last bomber's engines faded into the distance.
They observed the comet for 70 days before it faded from sight.
They watched the familiar mountains fade into the darkness.
We watched the harbour and then the coastline fade away into the morning mist.
She had a way of fading into the background when things got rough.
The most prominent poets of the Victorian period had all but faded from the scene.
Margaret Thatcher will not fade away into quiet retirement.
Sympathy for the rebels, the government claims, is beginning to fade.
Prospects for peace had already started to fade.
fading memories of better days.
Jay nodded, his smile fading.
He thought her campaign would probably fade out soon in any case.
You'll need to be able to project two images onto the screen as the new one fades in and the old image fades out.

collocate	occ.
away	2
under	1
began	0
just	0
beginning	1
before	1
never	0
die	0
hopes	1
quickly	0
soon	1
begun	0
flowers	0
memories	1
begins	0
total	8

collocates/examples 8/19=0.42.

stopword	occ.
to	2
and	2
will	1
as	2
out	2
they	0
into	5
or	0
then	1
total	15

stopwords/examples 15/19=0.78

Appendix 3 'Proposal'

Monolingual learner's dictionaries

Oxford Advanced Learner's Dictionary (Second Edition)

a proposal for peace
proposals for increasing trade between two countries
a girl who had five proposals in one week

Oxford Advanced Encyclopedic Learner's Dictionary (Updated Fourth Edition)

the proposal of new terms for a peace treaty
a proposal for uniting the two companies
Various proposals were put forward for increasing salaries
a proposal to offer a discount to regular customers
She had had many proposals (of marriage) but preferred to remain single

COBUILD I

There is controversy about a proposal to build a new nuclear power station
The two governments discussed a proposal for ending hostilities
I heard about some proposals for cheaper flights to the United States

COBUILD II

The president is to put forward new proposals for resolving the country's constitutional crisis
...the government's proposals to abolish free health care...
The Security Council has rejected the latest peace proposal
After a three-weekend courtship, Pamela accepted Randolph's proposal of marriage

Cambridge

Congress has rejected the latest economic proposal put forward by the president
There has been an angry reaction to the government's proposal to reduce unemployment benefit
Have you read Steve's proposals for the new project?
There was anger at the proposal that a UN peace-keeping force should be sent to the area
She refused his marriage proposal/proposal of marriage

Bilingual dictionaries

Hazon-Garzanti

English-Italian

the proposal was never carried out
to make a proposal
she had a proposal

Italian-English

proposal for the selection of an arbitrator

Collins-Robert

French-English (proposition)

propositions de paix/peace proposals
à la proposition de/on the proposal of

English-French

proposals for the amendment of this treaty

Collins German

German-English

to make somebody a proposal
his proposal of this plan surprised his colleagues
his proposal of John as chairman was expected

English-German (Vorschlag)

nothing

Collins Spanish

English-Spanish

proposal of marriage
to make a proposal
to make the proposal that

Spanish-English

nothing

Oxford-Hachette

French- English (proposition)

faire des propositions concrètes: to make concrete proposals
proposition technique/commerciale: technical/business proposal

English-French

to make/put forward a proposal faire
a proposal for changes
a proposal for doing *ou* to do
the proposal that everybody should get a pay rise
to receive a proposal

Oxford-Duden

English-German

make proposals for
make a proposal for doing sth or to
his proposal for improving the
draw up proposals/a proposal
proposal [of marriage]
he was interrupted in the middle of his proposal to her/the committee

German-English

a conciliatory proposal

Oxford Spanish

English-Spanish

to put or make a proposal to sb.

Spanish-English

proposal of marriage

New Proceed

make a proposal for peace
His proposal to put off the meeting was rejected
Reluctantly she accepted their proposal that she should be operated on
He made a proposal of marriage to her

Taishukan's Genius

make [offer] proposals for peace
We accepted a proposal to repair [for repairing, that we (should) repair] a road
receive a proposal (of marriage) from him
he made a proposal to her

Appendix 4 *Herald Tribune* article

Smart Arms In Gulf War Are Found Overrated
Pentagon's Reliance On High-Tech War Questioned in Review

By Tim Weiner New York Times Service

WASHINGTON - During and after the Gulf War, the Pentagon dramatically oversold the effectiveness of its most expensive high-tech aircraft and missiles, the most thorough independent study to date has found.

The Pentagon and its principal military contractors made claims for the precision of their most impressive new weapons - the Stealth fighter jet, the Tomahawk land-attack missile and laser-guided "smart bombs" - that "were overstated, misleading, inconsistent with the best available data or unverifiable," the study by the nonpartisan General Accounting Office found.

The accounting office concluded that new, costly "smart" weapons systems did not necessarily perform better in the Gulf War than old-fashioned, cheaper "dumb" ones. It called into question the wisdom of the military's plans to depend increasingly on weapons that extend the state of the art of war at a cost of tens of billions of dollars.

The accounting office analyzes government programs for Congress. Its secret four-year study of the air war conducted during Operation Desert Storm is the most detailed analysis of its kind to be made public.

It used more than one million pieces of information: Defense Department databases compiled for commanders, intelligence reports, after-action analyses and reports from military contractors. The accounting office also interviewed more than 100 Desert Storm pilots, war planners and battlefield commanders.

An unclassified summary of the 250-page secret report is scheduled to be published this week. The report was commissioned in 1992 by Senator David Pryor, Democrat of Arkansas, and Representative John D. Dingell, Democrat of Michigan, to help Congress decide what future weapons to buy.

The secret report contains facts and figures to buttress the 13-page unclassified summary, which was made available to The New York Times by a government official familiar with the underlying report.

During the war, Pentagon briefers treated the public to videotapes showing a smart bomb diving down the air shaft of a Baghdad building and told anecdotes about the extraordinary accuracy of Tomahawk missiles launched from afar. The study concluded that while some of those stories were true, they were not the whole truth.

The Pentagon did not dispute the new report's main conclusions. In an April 28 letter to the accounting office, the Defense Department said it "acknowledges the shortcomings" of its precision-guided munitions, the aircraft that carry them, the Tomahawk missiles and the department's ability to assess the effectiveness of its bombing campaign in the Gulf War.

It said it would deal with those shortcomings by building improved smart weapons, studying whether it has the right mix of weaponry in its arsenal and proposing new ways to locate and destroy targets.

American air power overwhelmed the Iraqi military during the 1991 Gulf War.

The United States deployed nearly 1,000 combat aircraft and unleashed nearly as many tons of bombs each day as were dropped on Germany and Japan daily during World War II.

But for all their superior technology, pilots often could not tell whether a presumed target was a tank or a truck or whether it already had been destroyed, the report said. Their sensors - laser, electro-optical and infrared systems - could not see clearly through clouds, rain, fog, smoke or high humidity, the report said.

The sleek black F-117 Stealth fighter jet, despite its high cost and its highly touted ability to get close to a target while evading detection, did not necessarily outperform older, cheaper aircraft.

The U.S. Air Force claimed an 80 percent success rate on bombing runs by the Stealth fighter, but the reality was closer to 40 percent, the report found.

"It is inappropriate, given aircraft use, performance and effectiveness demonstrated in Desert Storm, to characterize higher-cost aircraft as generally more capable than lower-cost aircraft," the summary said.

Nor did smart bombs necessarily deliver bang for buck, the summary said.

Only 8 percent of the bomb tonnage dropped on Iraq were smart bombs, or guided munitions. But they accounted for 84 percent of the cost of munitions in the war, the summary said.

Despite their cost, "the air campaign data did not validate the purported efficiency or effectiveness of guided munitions, without qualification," the summary said. "'One-target, one-bomb' efficiency was not achieved."

The cost of smart bombs being built by the Pentagon and planned for the future is now estimated to be more than $58 billion, more than triple what the government will spend this year on the FBI, the war on drugs, immigration control, customs, federal courts and prison construction.

"The cost of guided munitions," the summary concluded, "and the limitations on their effectiveness demonstrated in Desert Storm need to be addressed by the Department of Defense."

This is not the first time praise for Pentagon weaponry in the flush of victory in the Gulf has been questioned. In 1991, President George Bush said the Patriot missile system had been nearly perfect, shooting down 41 out of 42 Iraqi missiles aimed at Israel and Saudi Arabia.

Defense officials later said that the Patriot was far from perfect, knocking out perhaps 40 percent of the Scuds aimed at Israel and 70 percent of those aimed at Saudi Arabia.

Index