

Tradition and Innovation in Catalan Lexicography*

Joaquim Rafel i Fontanals

Universitat de Barcelona and Institut d'Estudis Catalans

The Organizing Committee of the XIII International Congress of the European Association for Lexicography has given me the honor of delivering the inaugural address. The organizers suggested to me that this occasion was an excellent opportunity to provide congress participants and readers of EURALEX volumes from around the world with a view of Catalan lexicography. The result is the following paper, which is divided into two parts. In the first part I provide a brief history of our lexicographic tradition; in the second, I discuss Catalan lexicography in the 20th century and concentrate on recent activities in the field of lexicography undertaken by the Institut d'Estudis Catalans. Our lexicographic tradition aims both to satisfy the specific needs of a society that uses Catalan as its language of communication and to adopt rigorous methodological criteria grounded in linguistic analysis. As I hope to show, our dictionaries are on a par with current lexicographic practise and incorporate several innovative characteristics.

From the beginnings to the 19th century

In the early stages of the language, interest in lexis is no different in Catalan than it is in other European languages, especially Western European languages. Alongside medieval glosses in Latin, which were included to improve the comprehension of texts, glosses in Catalan appeared throughout the XII, XIII and XIV centuries. These are the first rudimentary manifestations of what would later become glossaries and bilingual dictionaries. These glosses, together with lexical equivalents in Catalan that appear in some grammars written to facilitate the study of Latin, are the first testimonies we have of interest in Catalan lexis, although these works are not truly lexicographical in nature (Colon & Soberanas 1985[1991]: 11-38).

Lexicographic works begin to appear in the 15th century in the form of bilingual dictionaries concentrating on Catalan and Latin. These dictionaries were designed for decoding or studying Latin. A significant exception worthy of note is the thematic Catalan-German/German-Catalan vocabulary published in 1502. This anonymous work is an adaptation of the Italian-German vocabulary compiled by Adamo di Roudila, which had been published some years earlier, in 1477 (Colon & Soberanas 1985 [1991]: 56-59). Besides this isolated case, the publication of bilingual dictionaries, first for Catalan and Latin and then for Catalan and Spanish,

* I am grateful to the congress organisers for their kind invitation. The original paper was written and presented orally at the congress in Catalan, and I would like to take this opportunity to thank Janet DeCesaris for her English translation.

would continue through the mid 19th century. During this long period Catalan dictionaries are conceived of merely as collections of words with equivalents in other languages (in most cases, Latin and Spanish). Among the most noteworthy books were the *Liber elegantiarum* by Joan Esteve, published in 1489, which is the first collection of phrases in Catalan with the corresponding phrases in Latin, and the Catalan adaptation by Gabriel Busa of Antonio Nebrija's Latin-Spanish/Spanish-Latin dictionary. Busa's adaptation was first published in 1507 and would be reedited several times; it was to be a strong influence on other dictionaries. We should also mention the *Tesaurus Puerilis* by Onofre Pou, published in 1575; the *Fons verborum et phrasium*, by Antoni Font, published in 1637; the *Thesaurus verborum ac phrasium* by Pere Torra, published in 1640; and the *Gazophilacium catalano-latinum*, by Joan Lacavalleria, published in 1696. The situation of Catalan lexicography in the 18th century is, according to Colon & Soberanas, disheartening, to say the least; there is the *Diccionario valenciano-castellano* by Carles Ros, which appeared in 1764, as well as a few other minor works. In all of these dictionaries, however, Catalan is treated as a function of the other languages, without the autonomy that would make it the object of study itself with lexicographic criteria, i.e., from the standpoint of the meaning of lexical units and use in Catalan.

This was the situation of Catalan lexicography when European languages in more important cultures were already the object of autonomous, independent lexicographic works—monolingual dictionaries. It is useful to recall here the most representative works produced in Europe in the 16th, 17th, and 18th centuries. First, there was the *Vocabolario degli academici della Crusca*, written between 1591 and 1612, which was to become the model and stimulus for several dictionaries of reference in other languages; in French, there appear three dictionaries at the end of the 17th century that are testimonies to three different attitudes towards language and social position: the *Dictionnaire français* by Pierre Richelet (1680), the *Dictionnaire Universel* by Antoine Furetière (1690) and the *Dictionnaire de l'Académie Française* (1694) (Collinot-Mazière 1997: 25). In the case of Spanish, there was the *Tesoro de la Lengua Castellana o Española* by Sebastián de Covarrubias (1611) and later the first dictionary of the Spanish Academy (*Diccionario de la lengua castellana, en que se explica el verdadero sentido de las voces, su naturaleza y calidad, con las frases o modos de hablar, los proverbios o refranes y otras cosas convenientes al uso de la lengua*), commonly known as the *Diccionario de Autoridades*, which was published in six volumes between 1726 and 1739. Samuel Johnson published his *Dictionary of the English Language* a few years later, in 1755, and Johan Cristoph Adelung published his *Grammatisch-kritisches Wörterbuch der Hochdeutschen Mundarten* between 1774 and 1786.

We find nothing similar for Catalan during this period. A series of mainly socio-political factors was responsible for this situation. I shall not go into detail here, but suffice it to say that during this time (16th-18th centuries), the use of Catalan in writing and in the public arena declined in favour of Spanish. This process, by which Spanish supplanted Catalan in many contexts, reached its high point at the beginning of the 18th century with the *Nova Planta* ('New Design') decrees (1707

in Valencia, 1716 in Catalonia and 1717 in Majorca and Eivissa). In addition, as the Bourbon dynasty came to power in Spain, more laws were put into place to prevent Catalan from having any role in education or in public administration.

At the beginning of the 19th century, after a series of unsuccessful initiatives (especially, the dictionary projects of the Acadèmia de Bones Lletres de Barcelona ('The Academy of Letters of Barcelona'), the 2-volume *Diccionario catalan-castellano-latino* by Joaquim Esteve, Josep Bellvitges and Antoni Juglà is appeared (in 1803 and 1805). This dictionary, however, is not a monolingual dictionary along the lines of the European dictionaries mentioned above; rather, it is a trilingual dictionary with brief definitions and notes in Catalan for only a part of the entries. The fact that the title of the dictionary is in Spanish gives us an indication of the social status of Catalan at that time. Nevertheless, this dictionary represents the first and only successful effort in Catalan 17th century lexicography, despite the fact the dictionary was printed at the beginning of the 18th century (Colon & Soberanas 1985 [1991]: 135). This dictionary, even though it was not written by a language academy, was written by people who were members of the Acadèmia de Bones Lletres and who had belonged to the academic commissions that had (unsuccessfully) tried to produce a dictionary for the Acadèmia. The *Diccionario catalan-castellano-latino* has been described as the starting point for modern Catalan lexicography (Colon & Soberanas 1985 [1991]: 144) and we can consider it an indirect testimony to what might have been the first institutional or academic dictionary for the language.

We must look to the mid-19th century to find the first dictionary of Catalan with all the words defined in the same language as the headwords. The work by Pere Labèrnia, the *Diccionari de la llengua catalana, ab la correspondència castellana y llatina*, published in 1839-1840, is truly conceived of as a monolingual dictionary, although—as the title indicates—it also contains translation equivalents in Spanish and Latin. Pere Labèrnia was also a member of the Acadèmia de Bones Lletres and his dictionary has a certain institutional flavour to it; he dedicates the dictionary to the Acadèmia and asks the Academy to endorse it (see Rafel 2004: 108 for further discussion). The front page of the dictionary explicitly states that the dictionary is published under the auspices of the Acadèmia de Bones Lletres. It is important to stress that Labèrnia's work was the reference dictionary for Catalan for many years. Several subsequent editions appeared, and this dictionary was the basis for the encyclopaedic dictionary published by the Salvat publishing firm at the beginning of the 20th century (Colon & Soberanas 1985 [1991]: 155-156).

The first decades of the 20th century

The beginning of the 20th century was witness to two events that were to have a significant impact on Catalan lexicography.

The first important event was an initiative by Antoni Maria Alcover, published in 1901 in the document *Lletra de convit* ('Invited letter'), to undertake a large dictionary project. This dictionary project, as described by Alcover, would include Old Catalan, modern Catalan (with forms from the oral and written language),

literary variants and all the dialectal forms found within the language's geographic area (Alcover 1911: 370).

The second event was the creation of the Institut d'Estudis Catalans, the Catalan Academy of Sciences, in 1907. The Institut is an academic, scientific and cultural institution for scientific research on all aspects of Catalan culture. The goals of the Institut, according to the by-laws, are as follows: a) to study the Catalan language, establish its standard grammar and vocabulary and ensure that the standardisation of the language is coherent in all geographic areas in which the language is spoken; b) to contribute to the planning, coordination, and dissemination of research in science and technology; and c) to support the progress and development of society with its activities and to provide advisory studies for public officials and other institutions, as may be necessary. The Institut is currently structured into five thematic sections: History-Archaeology, Biological Sciences, Science and Technology, Philology, and Philosophy and Social Sciences. Although the establishment of the standard language is the responsibility of the Institut as a whole, this specific task—in addition to other language-related work—is of special concern to the Philological Section, which was created in 1911.

What we know about the early days of the Philological Section's activity is that it assumed Alcover's 1901 project as its own. It also approved the preparation of a *Diccionari català provisional* ('Provisional Catalan Dictionary') based on the written language with little attention given to dialectal varieties; this dictionary was to include the most common words of the literary language. Seen in retrospect, we can understand these decisions as a sort of compromise, or as a way to satisfy two quite different positions on the language and the actions that were considered most necessary at that point in time. The two opposing positions were represented by Alcover (whose dictionary would include all variants of the language, including geographic and historical forms) and by Pompeu Fabra (whose dictionary would unify and, to some extent, purify the language) (Rafel 1996: 219-220 and ensuing discussion).

It would take us too far afield to discuss the various vicissitudes these two projects experienced over the years. I shall only mention here that Alcover left the Institut in 1918 because of disagreements with other members and took with him the materials gathered up to that time. Francesc de Borja Moll continued this project after Alcover's death in 1932, and the result is the *Diccionari català-valencià-balear*, which was published in 10 volumes between 1926 and 1962. This important work is considered to be one of the most complete collections of lexical items and expressions in the Romance languages, and is particularly rich in data from Old Catalan and Catalan dialects.

The second project, led by Pompeu Fabra, became part of the Institut's task of establishing the standard language. The first important tangible result was the *Diccionari ortogràfic* ('Spelling Dictionary'), which appeared in 1917 under the direction of Fabra. The publication of this work was extremely important: on the one hand, it represented the consolidation of the orthographic rules that had been

approved in 1913; on the other, it meant that the form of the words considered most necessary for communication had been agreed upon and as such responded to the Institut's mission of providing a standard for the language. The Philological Section continued working on the language's dictionary of reference, which was known as the *Diccionari de la llengua literària* ('Dictionary of the Literary Language'). Work on this dictionary, however, did not progress as quickly as expected, mostly because of the inherently slow nature of academic work. The dictionary had been begun (up to the letter C) and the first fascicles were about to be published when the dictatorship of Primo de Rivera (1923-1930) rose to power in Spain. The dictatorship was opposed to Catalan language and culture, and consequently work on the dictionary ceased. Given the need for a standard reference dictionary, Pompeu Fabra took on the project himself, with financial backing from private sources. He used and completed the materials that had been collected for the *Diccionari de la llengua literària*, and the result was the publication of the *Diccionari general de la llengua catalana* in 1932. The dictionary appeared at the beginning of Spain's republican period, at a time when society was restoring Catalan institutions and the language was beginning to return to its proper social status. In this context, Fabra's dictionary provided an essential service to Catalan society. The Philological Section corrected and added new entries to the dictionary for the 1954 edition, and several later editions of the dictionary were published, up through the 32nd edition (which appeared in 1994 (Mir & Solà: 11). Fabra's dictionary was the reference dictionary for the standard language until the Institut d'Estudis Catalans published an updated dictionary in 1995. Dictionaries that Fabra consulted and took as models include the *Diccionario de la lengua española* of the Spanish Royal Academy (1925), *Webster's New International Dictionary of the English Language* (1909), and especially the *Dictionnaire general de la langue française* by A. Hatzfeld, A. Darmesteter and A. Thomas (no date, published in fascicles between 1890 and 1900) (Colon 2007).

The second half of the 20th century

It is well known that the results of the Spanish Civil War were extremely prejudicial for Catalan culture in general and for the Catalan language particularly. The Institut d'Estudis Catalans was officially disbanded, although some clandestine activity was carried out and even tolerated at certain points. This illegal status of the Institut began at the end of the Spanish Civil War in 1939 and continued to the beginning of the transition to democracy in 1976, when a royal decree returned the Institut to its original status.

During this time, as I mentioned above, the Philological Section revised and published the second edition of Fabra's dictionary, which would be reprinted several times after 1962. In the following years, the Philological Section limited itself to publishing short lists of words that were circulated informally, and these entries were incorporated in the dictionary as of the 4th edition (1966) and in an appendix to it as of the 5th edition (1968), which commemorated the centennial of Fabra's birth.

At the same time private publishing firms begin to publish bilingual dictionaries and occasionally a monolingual dictionary, which was always much indebted to Fabra's work. Of special note here is the publication in 1982 of the *Diccionari de la llengua catalana* by the firm Enciclopèdia Catalana. This dictionary includes all the entries in Fabra's dictionary, and adds a substantial number of new words that the company was able to obtain as it worked on the publication of the *Gran enciclopèdia catalana*, a 15-volume work published between 1962 and 1980. For several years, this initiative was particularly important in Catalonia because it assumed a leading role in academic circles, because the Institut had not yet returned to its normal research activities. In the 1980s, by which time Catalan had regained a public presence in the media, in education, in the public administration, etc., several dictionaries of varying types and length are published. Special mention should be given to the *Gran diccionari de la llengua catalana*, published in 1988 by Enciclopèdia Catalana, because it includes many more entries than the earlier dictionary. We must also note the many dictionaries published by the firm Edicions 62 in the year 2000 (*Gran diccionari 62 de la llengua catalana*, *Diccionari 62 manual de la llengua catalana*, *Diccionari 62 essencial de la llengua catalana* i *Petit diccionari 62 de la llengua catalana*). More details about these and other dictionaries of Catalan may be found in Haensch (1990) and, for those dictionaries published between 1940 and 1988, in Cabré & Lorente (1991).

None of these projects represented progress in methodological terms. In many ways, lexicography in Catalan was at a standstill. The dictionaries published by commercial companies during this period did not show much interest in incorporating methodological innovations. In fact, they did quite the opposite: they tended to adopt a conservative attitude, maintaining traditions acquired over the years. I can only agree with Colon & Soberanas when they somewhat discouragingly state that Catalan lexicography did not respond to the current demands of modern society.

This situation was all the more disheartening given that the 20th century was witness to many significant advances in linguistics, almost none of which had an impact on our lexicographic practise. Moreover, in the second half of the 20th century many researchers discussed lexicographic theory and methods that could be used in dictionary preparation, which resulted in a significant bibliography; practically none of these works had any influence on the dictionaries named above. The technological revolution, with its many advances in computerisation, permits a type of language processing that is directly applicable to lexicography, but in our context these advances had only a minor effect. Colon, as far back as 1977, noted the need for modernisation of Catalan lexicography: "el català no hauria de trobar-se absent de cap de les metodologies modernes" ('Catalan should not be without any modern technology') (Colon 1977: 16).

Recent dictionary projects undertaken by the Institut d'Estudis Catalans

As the Institut d'Estudis Catalans slowly moved towards regaining its place as a research institution, in 1983 the Philological Section of the Institut d'Estudis Catalans debated the best way to organise its activities in the field of lexicography. It decided to undertake a major dictionary project employing modern lexicographic techniques and principles. This project would be called the *Diccionari del català contemporani*, and aimed both to incorporate the main advances in linguistic analysis and to apply the possibilities of modern technology. At the same time, while this project was in progress, the Philological Section would also produce a new edition of Pompeu Fabra's dictionary, which was still the reference for the standard language even though it was clearly out-of-date.

At this point, then, the Institut d'Estudis Catalans becomes very active in lexicography. The by-laws of the Institut state that its missions include (1) setting the standard language and (2) describing actual use of the language. In terms of the lexicon, the Institut addressed these missions by undertaking two dictionaries: the *Diccionari de la llengua catalana* ('Dictionary of the Catalan Language') as the reference for the standard language and the *Diccionari del català contemporani* ('Contemporary Catalan Dictionary') project as the description of the language's vocabulary based on actual usage.

The Diccionari de la llengua catalana ('Dictionary of the Catalan Language')

Once institutional activity in Catalan was reinstated, the society needed an up-to-date dictionary for the standard language. The solution adopted by the Institut was to write a dictionary based on Pompeu Fabra's *Diccionari general de la llengua catalana* that would not change the basic criteria that Fabra had used; rather, the new dictionary would include new vocabulary and update the dictionary. Communicative needs had changed dramatically since 1932, and as a result a large number of new words and senses had come into the language. Many of these words and senses had been included in commercial dictionaries but were not present in the official dictionary of the standard language. This updated version of Fabra's work was published in 1995.

Following the publication of the *Diccionari de la llengua catalana*, the Philological Section decided that it needed to publish a shorter dictionary concentrating on commonly used vocabulary, a dictionary without so many technical words that could be used at schools and by people who did not need such a large work. It thus published the *Diccionari manual de la llengua catalana* ('Manual Dictionary of the Catalan Language') in 2001. This dictionary is not simply an abridged edition of the earlier dictionary, but rather includes some changes designed to make it easier to consult for the general public.

As soon as the first edition of the *Diccionari de la llengua catalana* appeared in 1995, work began on preparing the second edition, which was published in March, 2007. At the same time, a digital version of the dictionary was made available free

to the public on Internet. The digital edition allows many complex searches that are particularly useful for language professionals (teachers, language consultants, editors, etc.).

The Diccionari del català contemporani ('Contemporary Catalan Dictionary')

The *Diccionari del català contemporani* is the name of the ongoing project at the Institut d'Estudis Catalans. The Philological Section of the Institut realised that it needed to start a new period in its history, and that its work needed to benefit from the scientific, methodological, and technological advances that had occurred during the long period of forced inactivity. The project is structured into two stages: the first stage concentrated on developing language resources, while the second concentrates on producing a descriptive dictionary, the *Diccionari descriptiu de la llengua catalana*, based on those resources.

The First Stage: Creating language resources for Catalan

The first stage of the project, which is now completed, included the development of the Computerised Text Corpus of Catalan (in Catalan, the *Corpus Textual Informatitzat de la Llengua Catalana* or CTILC) and the Lexicographic Database (*Base de Dades Lexicogràfica* or BDLex), which contains the text of the most important dictionaries published in the 19th and 20th centuries.

a) Main features of the CTILC

From a chronological point of view, the CTILC includes texts spanning the period 1832-1988, in other words, more than 150 years in the history of Catalan. The date 1832 marks the beginning of the cultural movement to recover the use of the literary language, and 1988 was the date of the most recent texts included in the corpus. In terms of text typology, the corpus includes both literary and non-literary texts. These two groups were then divided into several subgroups, which allowed us to make a balanced choice of texts for each group (literary and non-literary) and for each subgroup (for literary texts, the four traditional genres: narrative, poetry, theatre, and essay; for non-literary texts, ten thematic domains: philosophy; religion and theology; social sciences; the press; pure and natural sciences; applied sciences; art, leisure and sports; language and literature; history and geography; and, correspondence). The corpus contains 52 million words, 44% of which correspond to literary text and 56% to non-literary text. The 3,299 texts in the corpus are of many different lengths. One of our main concerns was to ensure that the corpus would be as representative as possible, in other words, we wanted the set of texts included to reflect the written language as best as possible. Our objective was to strike a balance among the various text types that had been produced between 1832 and 1988.

A significant feature of the CTILC is that the entire corpus is lemmatised, which means that we have analysed each token and classified it grammatically so that it could be linked to a base reference form. This allowed us to achieve two important goals: on the one hand, graphic forms corresponding to different grammatical forms (such as verb forms belonging to a single conjugation or inflected forms

belonging to different lemmas) were disambiguated; on the other, forms belonging to an inflectional series were all linked to the lemma in question. The CTILC can be used in many different ways precisely because the entire corpus is lemmatised: for example, by looking up one lemma we obtain information on all the inflected forms, including the entire range of spelling variants that appeared in texts from different periods; we can also obtain all the derived forms containing evaluative or intensifying affixes which, according to morphological criteria in Catalan (as in other Romance languages) do not produce new lemmata. The number of lemmata in the corpus is 149,185, corresponding to 678,386 grammatical forms and 51,253,680 tokens.

The CTILC was developed during the period 1985-1997, which makes it a pioneering project in southern Europe. Once it was completed, a tool specially designed to consult it to write a dictionary was developed, although obviously this tool can also be used for other types of studies based on empirical data. The CTILC is currently available for consultation at no cost via Internet on the website of the Institut d'Estudis Catalans, and the consultation procedure allows you to obtain a series of contexts for each lemma looked up. As part of the corpus project, between 1996 and 1998 the Institut published a frequency dictionary of Catalan (Rafel 1996-1998). This frequency dictionary includes all the lexical data and corpus statistics from the project, and was made available both in a 3-volume printed edition and on 2 CD-ROMs.

b) Main features of the BDLex

The Lexicographic Database (BDLex) contains the thirteen most important Catalan dictionaries from the 19th and 20th centuries. This database was created to facilitate quick, systematic access to the information contained in these dictionaries, and was conceived as a complementary resource within the overall project. The various elements that constitute the structure of the dictionaries included in the BDLex have been identified and labelled systematically, so that they can be consulted individually or collectively.

The Second Stage: Writing a descriptive dictionary

The goal of the second stage of the ongoing DDC project is to write a descriptive dictionary of Catalan based primarily on the analysis of the data in the CTILC. We understand a 'descriptive dictionary' to be a dictionary that defines the lexical units in the language based on actual usage, regardless of prescriptive criteria.

One of the reasons a language academy decided to undertake a descriptive dictionary of this nature is the belief that prescriptive norms will be better justified if the language as used is objectively studied; in the case of the Institut d'Estudis Catalans, the justification is twofold, in that the Institut's by-laws specify that the Institut is not only responsible for establishing what constitutes standard usage but also for furthering the scientific study of the language (to quote the by-laws, "ocupar-se de l'estudi de la llengua"). With a descriptive dictionary of these characteristics,

the Institut d'Estudis Catalans has initiated a project guided by principles that are generally accepted in modern lexicography, and aims to create a valuable tool to comply with its mission as the academic institution charged with updating standard Catalan.

General characteristics of the DDLC

In an overview such as this, we are not able to detail all the structural features of the *Diccionari descriptiu de la llengua catalana* (DDLC); here discussion will be limited to the most general characteristics (see Rafel 2007 for more details).

The DDLC does not aim at being a theoretical dictionary, although it shares some of the properties associated with theoretical dictionaries. The DDLC does not attempt to serve a particular pragmatic purpose nor does it have a specific pedagogical goal. Rather, the information in the dictionary is the result of a specific research programme throughout which we have attempted to maintain the highest scientific standards. We specifically want to ensure that the information presented in the dictionary is expressed as precisely as possible, which implies a certain degree of formalism in the language used and in the layout of the dictionary. In this sense, the DDLC comes close to the concept of a 'purely language dictionary' or a 'linguistic dictionary' as proposed by Dirk Geeraerts (1985), as opposed what is generally called a 'language dictionary'. As such, the ideal users of the dictionary are language professionals. It is important to note, however, that this work is not only aimed at experts: rather, the dictionary strives to be useful for experts yet accessible to any educated person interested in language issues. In other words, the dictionary aims to combine the highest standards in explaining the information presented in the work and at the same time aims to present it in a clear, easily understood fashion.

As mentioned above, the DDLC strives to incorporate the methodological and technological advances of modern lexicography. Because it is a multifunctional database, it has the characteristics of an electronic dictionary that is not only useful for research but also as the basis for several types of works addressed to the public. The database structure allows us to make the dictionary available over the Internet, in addition to the traditional book format.

One of the most important characteristics of the DDLC is that it is based entirely on the CTILC, that is, it is a strictly corpus-based dictionary. This is neither the time nor place to argue for the advantages of a corpus-based dictionary; I should note, however, that writing a corpus-based dictionary implies that the corpus not only determines which headwords are included, but also the meanings of the lexical units as actually used, as opposed to using a priori methods (based on consulting other dictionaries) or lexicographers' intuition. Using a corpus allows us to use statistical data to set the headword list and other aspects of the dictionary. In addition, it allows us to include illustrative examples taken from citations with specific references, so that we can locate a particular sense or meaning in time. Since a corpus was used as our reference, the DDLC not only includes frequency data but also complete and systematic information on the syntactic structures used with the words included in the dictionary.

Given the type of dictionary we are working on, the definitions in the DDLC do not include encyclopaedic information (based on the description of reality) and concentrate instead on linguistic information (description of the meaning, the lexical restrictions, and the syntactic properties associated with each word). The defining text differentiates between the elements strictly reflecting the conceptual meaning (elements that are intrinsic to the definition) and those that refer to conditions or selective restrictions (extrinsic elements) that derive from the argument structure of the lexical unit defined (Rafel 2006). These extrinsic elements are related to the notion of *entourage* as described by Rey-Debove (1971) or that of *contorno* developed by Seco (1979).

Each entry in the DDLC contains quantitative information on use as shown by the corpus. This information is shown in a simplified fashion, and is represented graphically (and not numerically); each entry belongs to one of the five groups we established. For the first three frequency groups, the DDLC also contains percentages of the morphological categories in the corpus.

Another characteristic that is specific to this work is the writing process. The DDLC is not being produced in strict alphabetical order. Starting from the basic alphabetical order, we write—together with the entry for the word corresponding to alphabetical order—the entries for other lexical units that are related in form or in meaning, regardless of where those entries will be in the dictionary. This means that each lexicographer is responsible for writing a series of headwords that are related to one another, and the lexicographers work in parallel. This procedure attempts to avoid to the extent possible the lack of structural and descriptive coherence that characterises most existing dictionaries.

Current status of the DDLC

At the time of this writing, we have reached the midway point in the dictionary and have completed some 88,000 entries. As a result of the special procedure used to write the dictionary, approximately a third of the finished entries correspond to the first letters of the alphabet (A and B), and the remaining two-thirds belong to the other letters.

The DDLC was made available on-line in January, 2005. The digital edition of the dictionary can be consulted at no charge on the website of the Institut d'Estudis Catalans, and is continually being updated. The number of entries in the on-line edition increases as each group of entries is completed. Furthermore, any changes made to previously completed sections are incorporated into the dictionary automatically.

With this digital edition, the public can consult the DDLC using a dynamic tool that relates elements both within and across entries. The electronic edition also includes several hyperlinks to elements that are external to the dictionary.

* * *

As I come to the conclusion of this paper, it becomes obvious that space limitations prevent me from discussing all aspects of the development of Catalan lexicography. Nevertheless, I hope to have shown that our lexicographic tradition is characterized by a strong spirit to overcome adversity. Despite the significant obstacles in its path, Catalan lexicography has made significant accomplishments and today is well on its way to incorporating the conceptual and methodological advances required by modern lexicography.

References

- Alcover, A. M. (1911). "Crònica de la Secció Filològica de l'Institut d'Estudis Catalans". *Butlletí del Diccionari de la Llengua Catalana* VI (nº 20, August-September 1911). 368-372.
- Cabré, M. T.; Lorente, Mercè (1991). *Els diccionaris catalans de 1940 a 1988*. Barcelona: Publicacions de la Universitat de Barcelona.
- Collinot, A.; Mazière, F. (1997). *Un prêt à parler: le dictionnaire*. Paris: Presses Universitaires de France.
- Colon, G. (1977). "La lexicografia catalana: realitzacions i esperances". In Colon, G. (ed.). *Actes del quart Col·loqui Internacional de Llengua i Literatura Catalanes, Basilea, 22-27 de març de 1976*. Barcelona: Publicacions de l'Abadia de Montserrat. 11-35.
- Colon, G. (2007). "Introducció al Diccionari general de la llengua catalana". In Mir, J.; Solà, J. (dirs.) (2007). *Pompeu Fabra. Obres completes. Vol. 5, Diccionari general de la llengua catalana*. Barcelona: Institut d'Estudis Catalans. 15-40.
- Colon, G.; Soberanas, A.-J. (1985). *Panorama de la lexicografia catalana. De les glosses medievals a Pompeu Fabra*. Barcelona: Enciclopèdia Catalana. [2nd ed., 1991]
- Geeraerts, D. (1985). "Les données stéréotypiques, prototypiques et encyclopédiques dans le dictionnaire". *Cahiers de Lexicologie* 46 (1). 27-43.
- Haensch, G. (1990). "Katalanische Lexicographie". In Hausmann, F. et al. (eds.). *Wörterbücher. Ein internationales Handbuch zur Lexicographie. Dictionaries. An International Encyclopedia of Lexicography. Dictionnaires. Encyclopédie internationale de lexicographie*. Berlin - New York: Walter de Gruyter. II, 1770-1788.
- Mir, J.; Solà, J. (dirs.) (2007). *Pompeu Fabra. Obres completes. Vol. 5, Diccionari general de la llengua catalana*. Barcelona: Institut d'Estudis Catalans.
- Rafel, J. (1996). "El diccionari de l'Institut i el Diccionari Fabra". In *Estudis de Lingüística i Filologia oferts a Antoni M. Badia i Margarit*. Barcelona: Publicacions de l'Abadia de Montserrat i Departament de Filologia Catalana. 3, 217-269.
- Rafel, J. (dir.) (1996-1998). *Diccionari de freqüències* [3 volumes and 2 CD-ROMs]. Barcelona: Institut d'Estudis Catalans.
- Rafel, J. (2004). "La lexicografia institucional: el cas del català". In Battaner, P.; DeCesaris, J. (eds.). *De Lexicografia. Actes del I Symposium Internacional de Lexicografia (Barcelona, 16-18 de maig de 2002)*. Barcelona: Institut Universitari de Lingüística Aplicada - Universitat Pompeu Fabra. 103-122.

- Rafel, J. (2006). “Els elements extrínsecs en les definicions lexicogràfiques: teoria i aplicació”. In Bernal, E.; DeCesaris, J. (eds.). *Palabra por palabra. Estudios ofrecidos a Paz Battaner*. Barcelona: Institut Universitari de Lingüística Aplicada - Universitat Pompeu Fabra. 201-218.
- Rafel, J. (2007). “Prescripción y descripción en la actividad académica: el *Diccionari descriptiu de la llengua catalana*”. In Campos, M. et al. (eds.). *Reflexiones sobre el diccionario [Actas del I Congreso Internacional de Lexicografía Hispánica, A Coruña, setembre de 2004]*. A Coruña: Universidade da Coruña. 9-33.
- Rey-Debove, J. (1971). *Étude linguistique et sémiotique des dictionnaires français contemporains*. Paris-The Hague: Mouton.
- Seco, M. (1979). “El contorno en la definición lexicográfica”. In *Homenaje a Samuel Gili Gaya. In memoriam*. Barcelona: Biblograf. 183-191. Also in Seco, M. (1987). *Estudios de lexicografía española*. Madrid: Paraninfo. 35-45; 2nd ed. Madrid: Gredos. 47-58.