

Some Lexicographic Concepts Stemming from French Training in Lexicology (Part One)

Jean Pruvost



Jean Pruvost is the Vice-President of Université de Cergy-Pontoise, where he teaches linguistics, lexicology and lexicography, directs the laboratory of the CNRS [National Centre for Scientific Research] (Metadif, UMR 8127) that is devoted to dictionaries and their history, and organizes the annual international dictionary conference, *La Journée des dictionnaires*, a meeting point for lexicologists, lexicographers and lexicographers. He is the author of *Dictionnaires et nouvelles technologies* [*Dictionaries and New Technologies*], for which he received the international linguistics prize *Logos* in 2000, and of two titles in the *Que sais-je?* [*What Do I Know?*] series, one on French language dictionaries, the other on neologisms (with Jean-François Sablayrolles), as well as of numerous articles. Professor Pruvost co-directs, along with Bernard Quemada, two collections at Éditions Honoré Champion (*Études de lexicologie, lexicographie et dictionnaire*; *Lexica*), serves on several editorial boards (*Cahiers de lexicologie*; *Études de linguistique appliquée*; *International Journal of Lexicography*), and presides the Association des Sciences du Langage.
pruvost.jean@wanadoo.fr

Introduction

On the occasion of the Euralex Congress held in Copenhagen in August 2002, as a French lexicologist and lexicographer in contact with my colleagues who express themselves in English, three strong ideas crossed my mind:

- first, there was obviously a great deal for a French specialist to learn from my non-French colleagues in their specific approach to our discipline;
- then, from the outset it seemed to me that so far there have not been many exchanges between the French attitude and English and American ones;
- finally, I realized that if I have learnt a lot from my colleagues and friends, I may have also certain points of view and methods that are peculiar to my training, which could effectively take part in our collaborative study.

Also, encouraged by Ilan Kernerman for whom this conference was particularly stimulating, as well as by Tony Cowie whose benevolent dynamism I admire, I agreed to try to present some of these ideas which form a part of my creed and my training within the framework of this newsletter. Three perspectives seemed to me particularly interesting to develop.

The first corresponds to the distinction to be set between “lexicography” and “dictionary”. The recent notion of “dictionary” has actually been introduced by Bernard Quemada, director of the *Trésor de la langue française* [*Treasure of the French Language*] (16 volumes: 1971-1994), and it has been adopted fruitfully by numerous French lexicologists. This seems fundamental to me.

The second perspective is the one developed by Robert Galisson with regard to “lexiculture”. Galisson is one of our most original and efficient lexicologists of French as a foreign language. Actually, lexiculture is probably one of the most neglected components in the editing of entries in French or English dictionaries, sometimes even completely forgotten.

The third perspective is what I call the “triple dictionary investigation”. Some lectures I gave on the subject have convinced me that this particular approach may very likely have its virtues for the improvement of our dictionaries.

1. The useful distinction between lexicography and dictionary

In order to understand the difference between lexicography and dictionary and to perceive their essential complementarity, it is necessary to situate it in the recent history of French dictionaries which, more or less, is not very far from the history of lexicography in other western countries. One can actually distinguish four successive moments during the second half of the 20th century.

1.1 Lexicology disassociated from lexicography, in the traditional sense of the term

From 1950 to 1965, a first period distances itself on the whole wherein a distinction is made between “lexicology”, the scientific study of words, and “lexicography”, in the traditional sense of the term, that is the actual developing of dictionaries. We do know that lexicology as a study of words has not really attained the range of a scientific discipline until the second half of the 20th century. In France, a certain date is symbolic in this regard, it is actually in 1959 that the first issue of the *Cahiers de lexicologie* [*Journal of Lexicology*] appeared, and this scientific journal that was established and run by Bernard Quemada would cross the mark of the 21st century, with no less than 78 issues and an undeniable scientific success.

During this first period, lexicology and lexicography in the classic meaning of the word have each redefined itself and the one in relation to the other, lexicology becoming simply a scientific discipline, and lexicography clearly assimilating simultaneously to both a know-how and a science.

The lexicologists, while fully adhering to the continuity from philology, then assess the new-born structuralism and those present-day technologies, technologies offered at that period by punch card machines. To study the lexis and the vocabularies of big corpora with the aid of punch cards, such is the pioneer issue of this period. It is notably at Besançon, in the laboratories equipped with punch card machines, where the lexicologists from all over Europe were trained. Thus, in June 1961, a symposium, which today bears symbolic value in my eyes, is organized by Quemada at the University of Besançon, about the mechanization of lexicologic

research, a symposium altogether representative of the new state of mind which sets in. What does actually declare one of the participants, the Reverend father Busa, director of “Centro per l’automazione dell’analisi letteraria” from Gallarate, near Torino? “One is aware that all of us who take part in this conference, are pioneers of the automation of lexical analysis. We illustrate a necessary role in the evolution, which is the process of the book [...]. Today, alongside the printed journals and books, finds a place for itself the *electronic book*”. Such a declaration, in 1961, deserves to be qualified as visionary! It shows in any case that lexicology is assuming a new dimension derived from the new technologies that are being born.

As for the lexicography of this period, it is distinguished in France by the awareness that dictionaries should rely on a greater analysis of the criteria which define them. The dictionaries escape the isolated role of tools in order to become the object of a new reflection.

There is interest notably in the idea of *Dictionnaire du français fondamental* [*Dictionary of Basic French*] (1958), in the lign of “basic English”: this dictionary of basic French derives from a scientific experience with didactic perspective, being based on an analysis of vocabulary frequency. While awaiting a new great dictionary of the French language, a big dictionary that is symbolic of the 19th century is republished, the *Dictionnaire de la langue française* [*Dictionary of the French Language*] by Littré (first edition, 1873; reprinted in 1956). At the same time, in preparation, through the *Grand Larousse encyclopédique* [*Great Larousse Encyclopedic*] in ten volumes (1960-1964), the first defining steps are established for using new technologies, those of that time, namely 400,000 punch cards formed upstream of this paper dictionary. An *encyclopedic dictionary* but functioning also in the domains of language and technicalities, this *Grand Larousse encyclopédique* deserves to be remembered as one of the seminal dictionaries of this period. We do not tackle yet the computer era in the precise sense of this term, but the very rigorous methods, based on the algorithmic analysis, are already at work.

However, for the moment, lexicography may still keep its traditional meaning: it is comparable actually to the compilation of dictionaries, making use according to the circumstance of the best adapted technologies, and based on teams that are increasingly professional.

1.2 The birth of *metalexigraphy* and the new distinction *lexicography*/*dictionary*

The second period runs in general from 1965 to 1980, marking a moment when the dictionary benefits from a new status, being largely recognized as an object of scientific research. A French thesis entitled *Les dictionnaires du français moderne (1539-1863)* [The dictionaries of modern French (1539-1863)] (Didier, 1968), a dissertation made thanks to Quemada, sets itself as a parting point for numerous studies that will flourish concerning this or that past dictionary. A new discipline is thus given birth: *metalexigraphy*. *Lexicography*, until now, mainly tied to a daily need, observed above all as a tool, is henceforth part of corpora that is studied in order to better understand the history of the genre and the functioning of the language. In so doing, the dictionaries begin to be not only the creation of philologists and excellent craftsmen, they become a matter for linguists as well.

This second period coincides with a moment of intense commercialization of the dictionaries towards the general public, and a real revolution in information technologies, elevating the dictionary, in the classification of data and their interpretation. Actually, the research domain of the lexicon assumes a new scope, just as it becomes easier to produce dictionaries based on different computerized databases, adapting them for different sorts of public. Bernard Quemada introduces then a new dichotomy, between “*lexicography*”, to which he gives a new meaning in relation to its traditional sense, and “*dictionary*”, both concepts forming a useful dichotomy while being complementary.

Within the new contrast instituted between *lexicography*, in its new definition, and *dictionary*, *lexicography* exceeds well beyond the action of editing a dictionary to be compared to a real scientific research, driven by the words and their inventory, with all the defining works that correspond to it.

Dictionary represents in contrast all that is related to the concrete aspects of the production, of the presentation, for a given public, with all of the commercial imperatives that are imposed in order to please the public.

With *lexicography*, one is really placed in the domain of the research, without being preoccupied about according value for a non-initiated public, without worrying about adapting the content for readers who buy a product. One is somehow well above the dictionary that is put into shape in

Parts Two and Three of this contribution will be published in issue Number 12, July 2004.

All general references to lexicographers and users in this article, using *he/him/his*, apply to both male and female persons.

This article has been translated from the French language. The original French manuscript is available online: <http://kictionaries.com/newsletter/kdn11-03fr.html>

I am most obliged to Anthony Cowie for his dedicated and competent assistance with the English language version. Inefficiencies that might concern any aspect of the translation are my own. IJK

order to be sold, set in pure research. There can even be *lexicography* that, as opposed to the common definition of lexicography comparable to the compilation of a dictionary, does not necessarily lead to a sold dictionary. Corresponding to such or such research on the groups of words, on their definition, it may very well not leave the laboratory and correspond, for example, to the computerized bases designed uniquely for researchers. There isn't here the concern for grading the information for a seductive product in a purchasable form.

Dictionaric – a word that Charles Nodier has already used in the 19th century, but which has fallen into oblivion until Quemada resurrected it – defines for its part the act of developing a dictionary as a product, offered for sale, with all the constraints and problems related to each production, as an instrument for consultation, cultural media conceived with intention for a determined public of potential buyers. Thus, one must never forget that the dictionary represents a technic-commercial product whose content is defined in function of the means that are granted to it for a defined clientele, in the framework of a study of a specific market.

Thus, let us take for example two dictionaries that are very widespread in the French-speaking world and which are considered as having high quality, in this case the *Petit Robert* [*Little Robert*] or the *Petit Larousse* [*Little Larousse*] (be reminded that on average 200,000 *Petit Robert* are sold each year, and 800,000 *Petit Larousse* annually, thus over one million copies in 2001). When a new edition is offered (every year since they are bearing a date) and it is necessary to add a new word in a certain page, it is out of the question to recompose at the beginning of each school year the entire printed dictionary, so this or that example is simply removed from another entry on the same page, or such or such meaning, to gain the several lines which will allow to insert this new word without touching the beginning of the page and its end, and therefore not having to modify the preceding pages and those that follow. Here one is plainly in *dictionaric*: these are the practical restrictions that are imported to the defining quality and precision.

It is possible also, to better illustrate the difference that exists between *lexicography* and *dictionaric*, to assert that one may be an excellent lexicographer, that is carry out effective researches on the groups of words, on their definitions, and still turn out to be a lousy dictionarist, that does not succeed to respect the production timetables and the

inevitable material restrictions imposed. You therefore see great dictionaries that in their first editions have the advantage of enormous entries, making them almost illegible, then as one goes along, because the space will be lacking and it has already been necessary to increase the number of volumes that was initially expected, the entries get thinner, and you can even find yourself at the end of the alphabet with poor entries.

The publisher does not get confused with being a researcher, they must necessarily sell the product according to a selected size for a public to be seduced in a given price, during a certain period: the dictionaric is their first priority. The rule has no mystery: if the product is inappropriate, excessive, inconsistent in the density of information provided, the dictionary as a product will have no success, it will not sell, and the publishing house will be in danger.

Whatever the case is, lexicography and dictionaric are complementary: there is actually no interesting dictionaric if it is not based on solid lexicography, and lexicography is sometimes more efficient if it knew how to account for the dictionaric's constraints of time and of space which, in a certain way, maybe frame it and enhance it to more homogeneity in the description of a large group of words.

1.3 A revealing distinction of basic principles

It is possible to retain several lessons from this necessary distinction between the lexicography and the dictionaric.

First of all, it is important to separate the two perspectives, lexicographical and dictionaric. A dictionary, therefore a product, in which the two perspectives will be confused risks being very disappointing with respect to the user's expectations. The user wishes in general to have precise information, yet not stifling. If he buys for instance a thousand-page book, he will rather have useful and clear information than information that tends to be exhaustive and that transforms each entry to a reduction exercise, into a dreadful digest. To want always to give the maximum information in the minimal space, is to condemn the reader to reading with a magnifying glass, to an intellectualized reading of "researcher". Has one reflected, for example, that the dictionary represents a genre in which the editor refuses in principle any stylistic verbosity, that being considered out of place? This is the chase of the superfluity, the object is scientific and in this capacity, it should, one assumes, be austere.

Note, however, that the first monolingual French dictionaries, those of the 17th and

18th centuries from which they exude great charm, do not seem at this point restricted by a scientific rule, of nearly monastic nature, governing the entire work. Always to save space, by condensing as much as possible, in order to add new information, is not a good habit. Outside the “dictionary” genre, in the works with didactic nature, the verbosity is indeed very present, not to say essential, for the explications. It allows to space out the information, to make it accessible, digestible, it offers also the possibility to propose diverse approaches. Too much dense information definitely contradicts the efficient information, while harming the pleasure of the consultation. In the absence of a dictionary that enables straightaway this flexibility of editing, in the absence of a dictionary that knows how to limit the lexicography to a certain degree in order to add all the dictionaric that is convenient, that which will make the reading of the dictionary pleasant, you no doubt lose one of the primary functions of the dictionary: to make the information clear but also pleasant, legible. It is easy to add the dense information in the name of the lexicography, it is difficult to limit yourself and to choose in good dictionaric the most suitable text.

Then, to consider the result of the research as the editing of an article that should account for it in a hundred percent, is to confuse the stages. There is a time to conduct the research, to thus do the lexicography, with in the end an entry intended for the sole researcher; there is another time to adapt the results for the user, to thus install in dictionaric, not wanting necessarily to regive everything that has been found in lexicography. The entry compiled then is aimed at a reader who is not a linguist, nor willing to reread and reread definitions that are too dense. The information given to the reader should not be confused with the plain and simple recapture of the scientific and austere speech that is expected by the linguist. Thus, the absolute meticulousness and the care for exhaustivity which reign in the research are no longer necessarily the primary criteria: there is a need to adapt in order to explain better. The lexicographer-researcher can write for his peers when he is in the domain of the research, but when he becomes dictionarist, he doesn't write any more at all for his peers, he writes for all the readers and especially those who are not linguists. The dictionary has a didactic vocation as a tool for everybody.

Finally, have we reflected sufficiently on the fact that if it is good that the researcher knows everything possible about the functioning of the word in the language,

he should, when becoming dictionarist, not necessarily summarize it in as little space as possible, but on the contrary refer as much as possible to the particular questions that the reader asks concerning that word? Yet, the often systematic treatment of the information, in the way we do in linguistics, does not always respond to the majority of specific questions that the dictionary users ask themselves on such and such word.

There are as a matter of fact various words in the word, the “word of the language”, the “word of the speech”, the “literary word”, the “reference word”, etc. And yet, let us not forget, the word registered in the dictionary that often corresponds to a more or less successful synthesis of all these “words” hidden in a sole word, is not the word itself. The word described in the entry is not only that which has been analyzed between language and speech, it can first be perceived as the “consulted” word. And, in this capacity, the consulted word has in part its proper difficulties which often escape the homogenous rules of description, conceived for the entire group of the words in the dictionary.

A summary list was drawn up as example of these proper difficulties for certain French words, and it was noted that such a word was almost always consulted in the dictionary to raise the same problem. It is curious that there have not been organized groups of non-linguist observers, of dictionary users, noting systematically the questions that they pose to the dictionary. A big investigation of this type should be quite revealing. There is such an idea for online electronic dictionaries, when organizing an automatic observation of the questions posed, but the studies on the needs that were shown is lacking.

Thus, how do you write the French verb *rejeter* [*reject*] in the future tense: *rejetera*, *rejettera*? Nearly no dictionary considers introducing this in an example, although 80% of look-ups of the entry are for this question. For the commonly-used abbreviation, *pro* (a professional), the plural is never found, and you hesitate, can you write “pros”. Here too, consulting this word responds on the major part to this question of orthography. Indeed, for the linguist, the problem does not exist. He has in fact thought in terms of rules on the scale of the entire work, and he considers that, if he does not provide this or that information, it's because in his eyes it goes without saying. If he does not mention any particular note, it's because the general rule is followed. Obsessed as we are as lexicographers to economize space, any

saving of space is good to take, the general rule serves to gain on typographic spaces: too bad for the reader who is not aware of our obsession and who looks up a word just because he is not aware of the general regulation. This is treating with disdain the anxious look up of the user. The truth is that statistically the user never reads the introductions and he wishes for a direct reply to his queries. Thus, for the particular and complicated rules(4) of the pronominals in French which cause that “ils se sont développés” [“they have developed themselves”] does not take an ‘s’, here too, it is quite rare that the dictionaries offer the illustration hoped for in the examples. The same is true semantically where one does not expect necessarily and systematically the exhaustive description of all the semantic components of the word, but sometimes the examples clarify the referent, yesterday and today. Hardly any French person looks up the entry “chaise” [chair] for its orthography or its use in the language: here, it’s the referent that conveys it. It is therefore the referent that should be further developed in good dictionaric, yet a French language dictionary in principle does not offer an illustration. It is therefore necessary already to know what is a “chaise haute” [highchair], a “chaise longue” [deckchair], a “chaise à porteur” [sedan chair], a “chaise percée” [commode], a “chaise roulante” [wheelchair], the “chaise d’une meule” [?] in order to take advantage of these words listed in the language dictionaries, matching (not always) the definition of the more summaries.

This apparently essential distinction between the notion of *lexicography-research* and *dictionaric*, concerned with providing a pleasant and efficient tool for the user and not for the researcher, has not disappeared during the two subsequent periods. If during the second period mentioned were born, in fact, great dictionaries such as the *Dictionnaire du français contemporain* [Dictionary of Contemporary French] (1966), of a distributionalist nature, the *Trésor de la langue française* [Treasure of the French Language] (1971-1994), of a philologic nature, based on unequaled computerized textual documentation of the French language, the *Grand Larousse de la langue française* [Great Larousse of the French Language] (1971-1978), of a likewise distributionalist approach, the *Dictionnaire alphabétique et analogique de la langue française* [Alphabetic and Analogic Dictionary of the French Language] (1964; *Supplément* in 1970) by Paul Robert, in the renewed continuity

of Littré, actually, the third period that emanates from 1980 till about 1995, does further reinforce the useful distinction to be set between the lexicography and the dictionaric.

One can witness then a connection between, on the one part, the domains appropriate to the dictionaries designed for the human consultation, and, on the other part, the lexicomatic [= computational lexicography], a discipline reserved until then to computer scientists, this last discipline associating all that constitutes the base of lexical knowledges and all that refers to the machine-dictionaries for the computerized treatment of the languages and the language industries. The research assumes its full flight, the computerized means allow the works of great extent, the lexicography in its Quémadien meaning of the term is in full swing.

On the other hand, the information technology, even before the birth of the first CD-ROMs and Internet, enabled the gearing-down of the dictionaries designed for the public departing from well-nourished databases. Many small dictionaries thus appear, diversified according to the ages, the “dictionaric” may henceforth even exceed the data offered by the research, to sometimes have its autonomy, outside of the lexicography. It’s no longer the research, it’s the adaptation of the data, with so many “mixtures” with the data as its potential publics. Diversify to better sell. And it is by sometimes adapting with talent, efficiency, that the publics are acquired. This is the case of the *Dictionnaire historique de la langue française* [Historical Dictionary of the French Language] (Le Robert, 1992), for example, which presents in a pleasant way etymological information offered by the researchers, of CNRS notably. Then by contrast, providing nothing more than a fairly dull re-intermingling of information, selected, targeted at a perfectly calculated commercial profit, in the manner of a well packaged product.

To make sure that the *lexicography* does not close on itself, that the *dictionaric* does not auto-reproduce itself, this is then the course that should not be lost. Both perspectives, *lexicographical* and *dictionaric*, should remain united and complementary. Without research, there is actually no interesting future for the dictionaries. And without good dictionaric, the lexicography might stiffen and be of benefit to very few, without really attracting new competences.

As for the last period, of the very end of the 20th century and the beginning of the 21st century, marked by the

development of the Internet, it distinguishes itself primarily by the revival of editorial strategies, extended and adapted for the new virtual spaces, infinite spaces of information accessible in real time. It defines itself also by a profound metamorphosis of the look-up habits.

A problem remains: for the time being, it is mostly electronic adaptations of products, offered not long ago on paper, which are in the process of development or being offered on the market. This is the “redictionarization”, moving from paper on to computer, adding to it all the proper tags for the richest and most reciprocal look-up possible, and matching to it internet links. It nevertheless remains to invent the dictionaries conceived from the outset for the computerized medium, with no doubt real opening-ups for the hypertextual means between the encyclopedia and the language, between the synchrony and the diachrony, between the general vocabulary and the specialized vocabulary, between the textual examples and the visual, sound example, all in all synesthetic. Adding to it the lexiculture which we will expose later.

A new lexicography and a new dictionary are to be developed: the field of activity is immense. Many begin to prefer the muddled search on the Internet, certainly rich but unpredictable, on the consultation of a real “dictionary” based on this opening up, starting by not confusing *lexicography* with the *dictionary*. Already, some works take up the challenge, especially on the side of learner dictionaries. All together, on the global scale, we will not be cautious.

References

Books and articles

Galisson R. et J. Pruvost (dr.), 1999. *Vocabulaire et dictionnaires en français langue maternelle et en français langue étrangère*. Étude de linguistique appliquée, n° 116 (octobre-décembre). Paris: Didier Érudition.

Pruvost J., 2002. ‘À la recherche de la norme: sa représentation lexicographique et dictionnaire chez Larousse et Robert et la triple investigation’. In *La représentation de la norme dans les pratiques terminologiques et lexicographiques*, Langues et sociétés n° 39. Montréal: Office de la langue française, Québec, pp. 139-170.

Pruvost J., 2002. *Les dictionnaires de langue française*. Collection ‘Que sais-je?’ n° 3622. Paris: Presses Universitaires de France.

Quemada B., 1961. ‘Actes du Colloque sur la mécanisation des recherches lexicologiques’. In *Cahiers de lexicologie*, n°3. Paris: Didier Érudition.

Quemada B., 1987. ‘Notes sur lexicographie et dictionnaire’. In *Cahiers de lexicologie*, n° 51. Paris: Didier Érudition, pp. 235-245.

Dictionaries

Dictionnaire alphabétique et analogique de la langue française, 5 vol. and Supplement, 1964-1970. Paul Robert. Paris: Société du Nouveau Littré.

Dictionnaire de la langue française, 4 vol., 1863-1873. Littré É. Paris : Librairie Hachette.

Dictionnaire du français contemporain, 1966. Dubois J. (dir.). Paris: Larousse.

Dictionnaire du français fondamental, 1958. Gougenheim G. Paris: Didier.

Dictionnaire historique de la langue française, 2 vol., 1992. Rey A. (dir.). Paris: Le Robert.

Grand Larousse encyclopédique, 10 vol., 1960-1964. Dubois Cl. (dir.). Paris: Larousse.

Le Petit Robert, Dictionnaire alphabétique et analogique de la langue française, since 1967, revised every year. Paris: Société du Nouveau Littré, Société Le Robert.

Le Petit Larousse illustré, since 1905, revised every year. Paris: Larousse.

Trésor de la langue française, Dictionnaire de la langue du XIXe et du XXe siècle (1789-1960), 16 vol., 1971-1994. Paris: Klincksieck Vol. 1 à 10. Paris: Gallimard, vol. 11 à 16. Directeurs: Imbs P. (vol. 1-7), Quemada B. (vol. 8-16). Online: <http://www.inalf.fr/tlfi>