Dialectologia 1 (2008), 121-135.

MAPPING MORPHOLOGICAL AND PHONETIC FEATURES OF CATALAN: A GENERAL TEMPLATE FOR CONTEMPORARY ATLASES AND CORPUS[1]

Maria-Pilar Perea
Universitat de Barcelona
mpilar.perea@ub.edu

Abstract

In Catalonia, from a general point of view and concerning Geolinguistics, three assessments can be done: *a*) no new initiatives for creating a general linguistic atlas are expected; on the contrary, the tendency would be to create regional or local atlases or, disregarding cartography, to develop of monographs concerning several linguistic aspects of a certain dialectal area; *b*) there is no perceived need for an electronic publication of the atlas or the release of an internet version (the general format used is paper); and *c*) there is a possibility of computerising the data contained in old atlases. The main aim of this paper is to describe the processes of systematisation and mapping of dialectal data based on "La flexió verbal en els dialectes catalans". The paper is structured in five parts: *a*) The corpus of morphological and phonetic data; *b*) Mapping the data; *c*) Using the program; *d*) Sound maps; *e*) Conclusions.

Key words

Dialectology, geolinguistics, morphology, automatic mapping.

1. Introduction

Traditionally, the main aim of geolinguistics was to study dialectal variation from a geographical point of view, using as its main tool the linguistic atlas, understood as a set of maps that show the spatial distribution of the phonetic, morphological, and particularly the lexical data obtained from questionnaires.

The largest and most recent Catalan atlas, currently in publication, is the *Atlas Linguistic del Domini Català*, by Joan Veny and Lídia Pons. Two volumes (of a total of ten) have so far been published. It is a traditional atlas, essentially of lexical typology,

[1] A first version of this paper was presented at the 119th meeting of the Variation Theory Forum of Japan on June 30 of 2006, held at Seisen University of Tokyo. I am very grateful to Asahi Yoshiyuki and Danny Long for their invitation.

but it also covers some morphological and syntactic aspects. Surveys for the preparation of this atlas began during the 1960s. 190 towns were visited, a single informant was consulted, and most of the data were recorded on tape while one of the interviewers noted the answers on a hard copy of the questionnaire.

If we consider that, during the 20th century, only three atlases have been published on Catalan dialects – the first, the *Atlas Lingüístic de Catalunya* (ALC), by Antoni Griera, of doubtful quality; the second, *the Atlas Lingüístico de la Peninsula Ibérica* (ALPI), by Tomás Navarro Tomás, in which the Catalan language was partially covered and will subsequently be more comprehensively treated thanks to the efforts of Dr. David Heap; and the third, the abovementioned *Atlas Lingüístic del Domini* Català (ALDC), a valuable tool but one that reflects data gathered more than thirty years ago – it is plausible to begin a project to develop a new atlas that compiles dialectal information on contemporary Catalan. I may be proved wrong, but it seems unlikely that anybody would embark on a project to develop a new, general Catalan atlas. New initiatives only have a regional scope, such as the *Atles Lingüístic de la Comunitat Valenciana*, a work still in progress, the *Atlas Lingüístic de la Terra Alta* or the *Atlas Lingüístic of the Diòcesi of Tortosa*, which have already been published in paper form.

From what has been said, one can deduce the following: *a*) no new initiatives for creating a general linguistic atlas of Catalan are expected; *b*) hypothetically, the tendency would be to create regional or local atlases or, disregarding cartography, the development of monographs concerning several linguistic aspects of a certain dialectal area; *c*) except in the case of the ALPI, there is no perceived need for an electronic publication of the atlas or the release of an internet version (the general format used is paper); *d*) there is a possibility of computerising the data contained in old atlases, but there are currently no such projects underway, with the exception of the ALPI.

Although to date all the atlases have been published in paper format, no initiative has arisen to publish them electronically and no institution has undertaken to design a program for computerised cartography. For the Catalan dialects, a process of mapping has been applied to two sorts of materials: *a*) old materials, gathered by Antoni M. Alcover approximately one century ago, which form a systematic and complete corpus of the verbal morphology of Catalan at the beginning of the 20th century; *b*) recent materials, taken from the surveys made in the ECOD project. The *Corpus Oral Dialectal* (COD) compiles a systematic set of phonetic and morphological materials, to

which a similar protocol used for Alcover's data has been applied in order to produce electronic maps.

It is necessary at this point to give an explanation of the concept of corpus. In general, this term has been applied to a collection of texts in electronic format that have been collected and systematised according to determined criteria and objectives. They are primarily useful resources for making dictionaries, but they also offer excellent and representative data on other aspects of the language. If we broaden the meaning of this term, corpus can also denote complete collections of linguistic data, particularly those that are included in databases. Thus, the 470,255 entries from Alcover's *La flexió verbal* or the content of the COD included in several databases can be denominated "corpus"; and it is to these corpus that the electronic cartography has been applied.

The consideration of electronic supports is also a response to the high costs of publishing an atlas in paper format, particularly when it includes a large number of data. It is therefore desirable that the preparation and publication of atlases should not take too long. In a linguistically dynamic society such as Catalonia, due to frequent movements among its population or the impact of high levels of immigration, the publication of dialectal data would have to be fast enough to reflect the linguistic reality of the moment. Internet is the most suitable support for an atlas that shows the data from the first moment of processing. These data, once transcribed, can then be updated progressively.

In this long introduction the basic considerations of this paper have been outlined. The main aim is to describe the processes of systematisation and mapping of dialectal data. The paper is structured in five parts: *a*) The corpus of morphological and phonetic data; *b*) Mapping the data; *c*) Using the program; *d*) Sound maps; *e*) Conclusions.

2. The corpus of morphological and phonetic data

The automatic mapping project began in 1999. The first step was to enter into a database (Microsoft ® Access) the materials published in *La flexió verbal en els dialectes catalans*, by Antoni M. Alcover and Francesc B. Moll. Next, the corpus was completed by adding unpublished data from several notebooks. The database included almost 500,000 entries which were classified into several fields. The search program

made it easy to transfer from the phonetic entries to the geographical areas of the future computerised map.

The two programs (database and mapping) were published on a CD-ROM, financed by the Balearic Government. The CD shows the geographical distribution of the complete conjugation of 117 verbs in 149 areas of the Catalan linguistic domain. It is, then, an electronic atlas that allows the design of 6,435 maps. The atlas is consistent with the typology indicated by Francis (1983: 120-126) and contains lists of data and the corresponding maps.

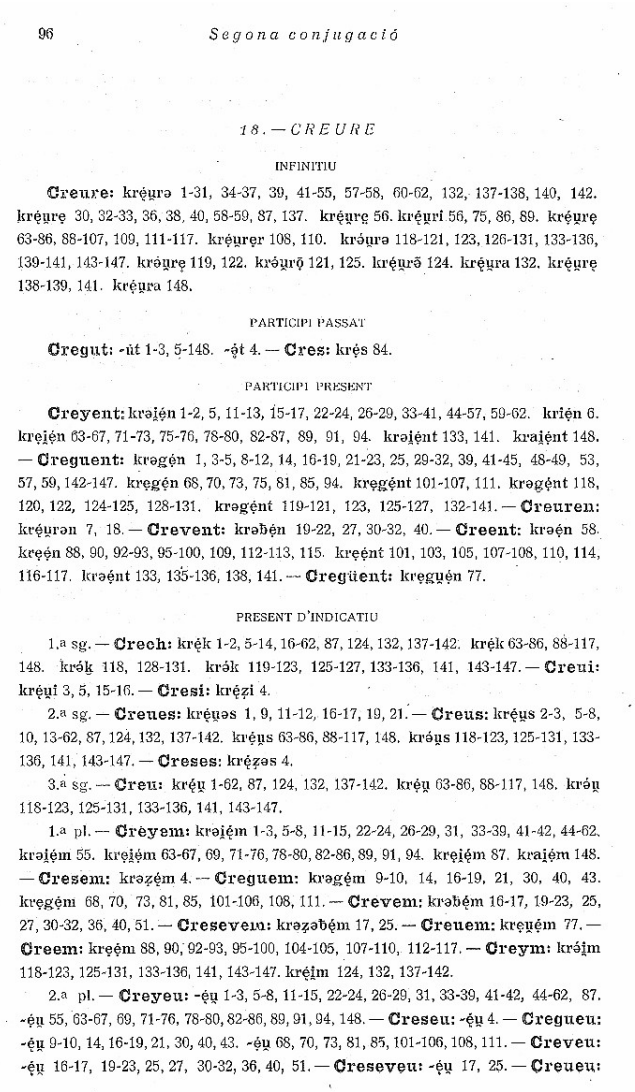
The electronic cartography program was next adapted for the phonetic and morphologic data obtained in relatively recent surveys (1995-1997), as part of the research project ECOD (HUM2004-1504), developed by the Department of Catalan Philology of the University of Barcelona. In this case, several databases compile dialectal material corresponding to phonetic and morphological aspects (articles, possessives, personal pronouns, locatives, demonstratives, pronominal clitics and regular verbs).

For the sake of concision, the presentation of methodological aspects of both projects is omitted. These include the type of questionnaire used, the method of eliciting responses, the selection of informants and localities and the validity of the results, all of which can be found in other papers (cf. Alcover: Perea 2001 a, b; 2002; COD: Lloret & Perea 2002; and Alturo, Boix, & Perea 2002).

2.1. *Alcover's data*

In the original edition of *La flexió verbal*, the sixty-seven verbs recorded were classified by conjugation. Different verb tenses (infinitive, gerund, participle, present indicative, past indicative, preterite, future, conditional, present subjunctive, past subjunctive and imperative) of each verb were shown (see figure 1).

To develop the verb paradigm, each verb form was related to a morphological variant, which helps to determine its dialectal scope, and to a phonetic form. Alongside the phonetic variant appears a list of the localities in which this answer was recorded. The localities were represented by numbers.



96 *Segona conjugació*

18.—CREURE

INFINITIU

Creure: krę̯urə 1-31, 34-37, 39, 41-55, 57-58, 60-62, 132, 137-138, 140, 142. krę̯urę 30, 32-33, 36, 38, 40, 58-59, 87, 137. krę̯urę 56. krę̯urí 56, 75, 86, 89. krę̯urę 63-86, 88-107, 109, 111-117. krę̯urer 108, 110. krə̯urə 118-121, 123, 126-131, 133-136, 139-141, 143-147. krə̯urę 119, 122. krə̯urō 121, 125. krę̯urō 124. krę̯ura 132. krę̯urę 138-139, 141. krę̯ura 148.

PARTICIPI PASSAT

Cregut: -út 1-3, 5-148. -ęt 4. — **Cres:** krę̯s 84.

PARTICIPI PRESENT

Creyent: krəi̯ę́n 1-2, 5, 11-13, 15-17, 22-24, 26-29, 33-41, 44-57, 59-62. kri̯ę́n 6. kre̯i̯ę́n 63-67, 71-73, 75-76, 78-80, 82-87, 89, 91, 94. krəi̯ę́nt 133, 141. krai̯ę́nt 148. — **Creguent:** krəgę́n 1, 3-5, 8-12, 14, 16-19, 21-23, 25, 29-32, 39, 41-45, 48-49, 53, 57, 59, 142-147. kre̯gę́n 68, 70, 73, 75, 81, 85, 94. kre̯gę́nt 101-107, 111. krəgę́nt 118, 120, 122, 124-125, 128-131. krəgę́nt 119-121, 123, 125-127, 132-141. — **Creuren:** krę̯urən 7, 18. — **Crevent:** krəbę́n 19-22, 27, 30-32, 40. — **Creent:** krəę́n 58. kre̯ę́n 88, 90, 92-93, 95-100, 109, 112-113, 115. kre̯ę́nt 101, 103, 105, 107-108, 110, 114, 116-117. krəę́nt 133, 135-136, 138, 141. — **Cregüent:** kre̯gu̯ę́n 77.

PRESENT D'INDICATIU

1.ª sg. — **Crech:** krę́k 1-2, 5-14, 16-62, 87, 124, 132, 137-142. krę́k 63-86, 88-117, 148. krə́k 118, 128-131. krə́k 119-123, 125-127, 133-136, 141, 143-147. — **Creui:** krę̯uí 3, 5, 15-16. — **Cresi:** krę́zi 4.

2.ª sg. — **Creues:** krę̯ués 1, 9, 11-12, 16-17, 19, 21. — **Creus:** krę̯us 2-3, 5-8, 10, 13-62, 87, 124, 132, 137-142. krę̯us 63-86, 88-117, 148. krə̯us 118-123, 125-131, 133-136, 141, 143-147. — **Creses:** krę́zəs 4.

3.ª sg. — **Creu:** krę̯u 1-62, 87, 124, 132, 137-142. krę̯u 63-86, 88-117, 148. krə̯u 118-123, 125-131, 133-136, 141, 143-147.

1.ª pl. — **Creyem:** krəi̯ę́m 1-3, 5-8, 11-15, 22-24, 26-29, 31, 33-39, 41-42, 44-62. krəi̯ę́m 55. kre̯i̯ę́m 63-67, 69, 71-76, 78-80, 82-86, 89, 91, 94. kre̯i̯ę́m 87. krai̯ę́m 148. — **Cresem:** krəzę́m 4. — **Creguem:** krəgę́m 9-10, 14, 16-19, 21, 30, 40, 43. kre̯gę́m 68, 70, 73, 81, 85, 101-106, 108, 111. — **Crevem:** krəbę́m 16-17, 19-23, 25, 27, 30-32, 36, 40, 51. — **Cresevem:** krəzəbę́m 17, 25. — **Creuem:** kre̯u̯ę́m 77. — **Creem:** kre̯ę́m 88, 90, 92-93, 95-100, 104-105, 107-110, 112-117. — **Creym:** krə́i̯m 118-123, 125-131, 133-136, 141, 143-147. kre̯i̯m 124, 132, 137-142.

2.ª pl. — **Creyeu:** -ę́u 1-3, 5-8, 11-15, 22-24, 26-29, 31, 33-39, 41-42, 44-62, 87. -ę́u 55, 63-67, 69, 71-76, 78-80, 82-86, 89, 91, 94, 148. — **Creseu:** -ę́u 4. — **Creguеu:** -ę́u 9-10, 14, 16-19, 21, 30, 40, 43. -ę́u 68, 70, 73, 81, 85, 101-106, 108, 111. — **Creveu:** -ę́u 16-17, 19-23, 25, 27, 30-32, 36, 40, 51. — **Cresevеu:** -ę́u 17, 25. — **Creueu:**

Figure 1. A detail of the edition of "La flexió verbal en els dialectes catalans"

In general, only the 1st person singular of the present indicative was transcribed, since this form has a stressed root. The remaining roots were not transcribed – only the endings. This format provides a concise representation of a large number of verb forms, although it remains a simplification, and both the consultation and reconstruction of the verb forms were restricted.

We surmised that problems generated by incomplete verb forms could be overcome with a computerised methodology. As forms were entered into the database, all the verb paradigms were developed. The computerised system also allowed the correction of a number of errata and made it possible to add to the corpus of verb forms by incorporating the unpublished list of irregular verbs that was included in Alcover's original notebooks, but not in the final published work.

The database included the following fields:

1. Verb identification	9. Edition (published work / Alcover's notebooks)
2. Locality	10. Infinitive form
3. Number of locality	11. Standard form
4. Verb tense	12. Conjugation: regular and special conjugations
5. Person: 1st, 2nd, 3rd singular plural and	13. Dialectal areas and subareas: Catalan Northern,
6. Morphological variant	Catalan Eastern, Catalan Western, Valencian, Balearic
7. Phonetic form	(Majorca, Minorca and Ibiza), and Algueres (Sardinia)
8. Year of inquiry	14. Remarks (sociolinguistic aspects)

The process of ordering and completing the materials generated a morphological corpus of 470,255 entries (adapted to the IPA alphabet) comprising of all the verbs that appear in the published work, the verb forms of certain tenses (i.e. future and conditional) that in *La flexió verbal* are referred to the pattern of the verb *cantar* ('to sing'), verbs with special conjugations, the new verbs taken from Alcover's original notebooks and the verb forms from new localities that were not included in the published work.

2.2. *COD data*

The data for the phonetic and morphological corpus were obtained through several surveys made in the capitals – or equivalent towns – of all the counties in the Catalan linguistic domain. Surveys were carried out in 86 towns and a minimum of three informants were interviewed.

The answers of the questionnaire were orthographically and phonetically transcribed and then typologically grouped in several databases. The databases, together with the number of entries, are the following:

	Entries
1. Phonetics	49,812
2. Articles	5,049
3. Demonstratives	1,512
4. Possessives	2,006
5. Locatives	803
6. Personal pronouns	1,608
7. Pronominal clitics	21,591
8. Regular verbs	68,152

The structure of each database varies according to its typology. In short, the database containing regular verbs differs in part from that used to systematise Alcover's *La flexió verbal*. A general modification is the inclusion in all COD databases of the morphological segmentation of each phonetic form. The verbal database includes the following fields:

1. Dialect	9. Phonetic form
2. Locality	10-17: Morphological segmentation: 10. Root; 11.
3. Informant	Epenthesis-1; 12. Extension; 13. Thematic vowel; 14.
4. Conjugation	Mode; 15. Mode/Tense; 16. Epenthesis-2; 17.
5. Verb identification	Number/Person
6. Tense	18: Alternative forms
7. Person	19. Remarks
8. Standard form	20. Phonetic remarks

3. Mapping the data

Once systematised and completed, the materials of *La flexió verbal* formed a suitable corpus for the creation of a computerised linguistic atlas. The eight COD databases, once revised, were also ready for mapping.

The data mapping process consisted in transferring the linear structure of the original printed data to a database structure and to a cartographic representation of the entries included. Using available technology, researchers are able to design an *à la carte* set of maps – either of the entire Catalan linguistic domain or of specific dialectal areas

– which will progressively increase their own linguistic atlas if saved to the hard disk of their computers.

The linguistic atlas created by the computer program is a collection of potential maps that can be updated as the user wishes. Each map places the phonetically transcribed dates in the various localities surveyed, represented by points. The maps are defined by the outline of the area to which the printed data refers, following the tradition of the *Atlas Linguistique de la France* (ALF) by Jules Gilliéron, the *Sprach-und Sachatlas Italiens und der Südschweiz* [*Atlante Italo-Svizzero*] (AIS) by Karl Jaberg and Jacob Jud, *the Linguistic Atlas New of England* (LANE) by Hans Kurath and other national and regional atlases. Users may choose between this presentation and an alternative format in which all the phonetic results are associated to symbols. The symbols clearly highlight the geographic areas in which particular linguistic uses coincide.

The result is a linguistic atlas that accomplishes three goals: *a*) it presents a synchronic, morphological and phonetic description; *b*) it shows the formation of different linguistic areas through the distribution of coinciding forms; and *c*) it provides representative material for subsequent study or interpretation of the data.

The mapping of Alcover's data and the COD databases was carried out in different phases. The COD data were supported by sound entries, unlike the data from *La flexió verbal*, for which we have only the phonetic and orthographic transcriptions. For this reason, it is desirable for the electronic maps generated with the COD materials to incorporate the corresponding sound sequences.

There is a minor problem related to the selection of the most representative data. Since the COD project includes the answers – not always corresponding – of three informants, from a practical point of view it was considered preferable to select only one answer for each town. The phases implemented in this project to finally produce the computerised cartography of the various databases are described below.

3.1. *Selecting the COD entries*

Given the variation present in some answers, we considered that associating more than one answer with the corresponding sound would make it difficult to obtain the correct visualisation on the map. For this reason, once phonetically transcribed, the

materials were selected by a semiautomatic procedure. A program selected a form at random when the answers of the three informants corresponded exactly. When two answers matched and one differed, the program selected the most frequent entry (one of the two corresponding answers). Finally, when all three entries differed, the program marked the results so that the choice could be made manually. The researcher, comparing other similar results from the survey, chose the form that was most representative of the town in question.

The selection program was called the "corrector" (Figure 2) and was used to adjust the alternative forms that appear in the phonetic entries, with the aim of facilitating the computerised cartography. The maps included the phonetic entry and the sound sequence. The selection procedure made it easy to apply other statistical methods of quantitative analysis to the results, such as dialectometry.

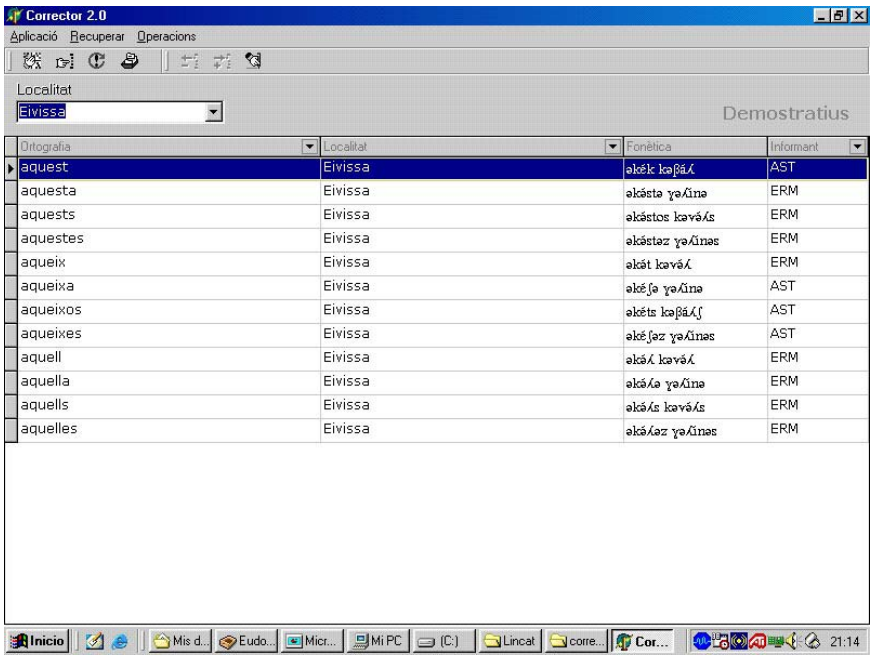


Figure 2. The main screen of the programme "Corrector"

It is not necessary for the selected results to refer to the same informant; what is important is that the data are genuinely representative of certain localities.

Furthermore, the use of a single entry for each answer in a given locality does not mean losing the information not selected. The complete set of entries is available in a general database that has been compiled with the ORACLE application. The total

amount of data can be consulted in order to obtain more exhaustive information, with the aim of developing a complete descriptive work.

3.2. *The segmentation of sound entries*

The segmentation mechanisms of the full sound sequences and the association of the text (phonetic and orthographic) with the corresponding sound are extremely laborious. In order to make these easier, a computerised program called "Dialectical cuttings" (Figure 3) has been developed, with which the following steps have been carried out:

1. The selection of a locality.
2. The selection of a database, from the eight available on ORACLE.
3. The association of the sound with the phonetic transcription of the selected locality, together with other relevant information.

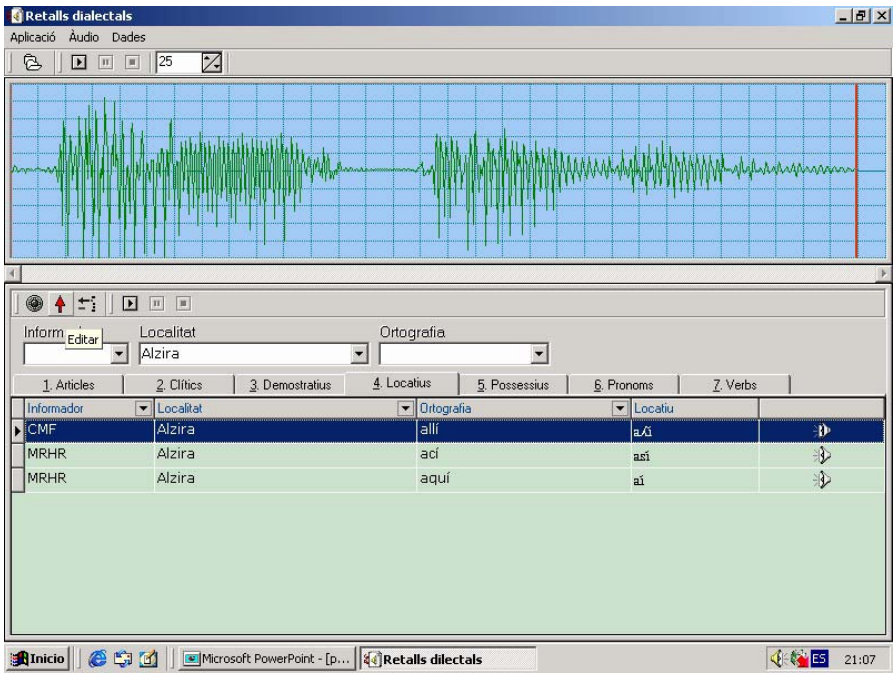


Figure 3. The main screen of the programme "Dialectal cuttings"

4. Using the program

With the mapping program, users can design maps on screen, save them on the hard disk and print them.

The data obtained through the search screen can be visualised in two ways: lists and maps. Maps also have two options:

a) A general map of the Catalan linguistic domain (figure 4).

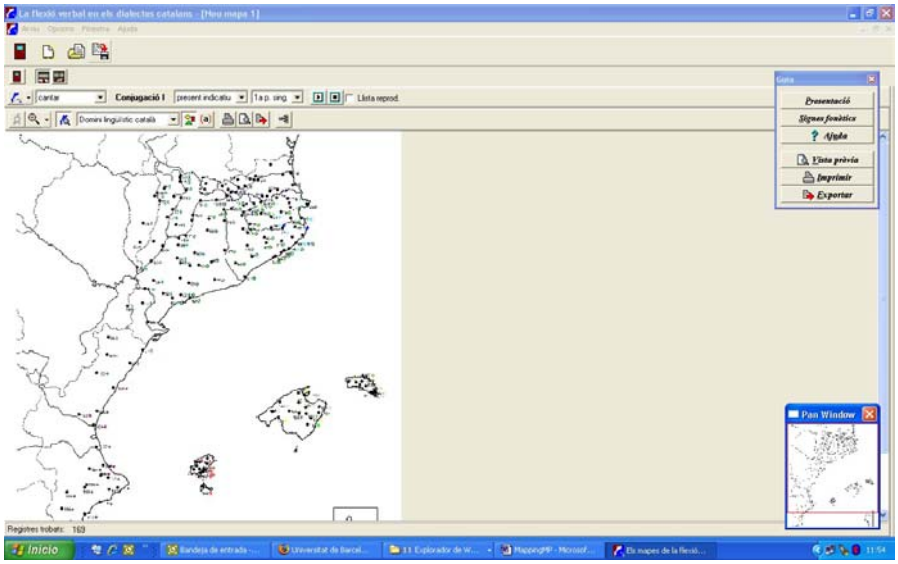


Figure 4.

b) More detailed maps of each of the six main dialectal areas into which the Catalan linguistic domain is divided.

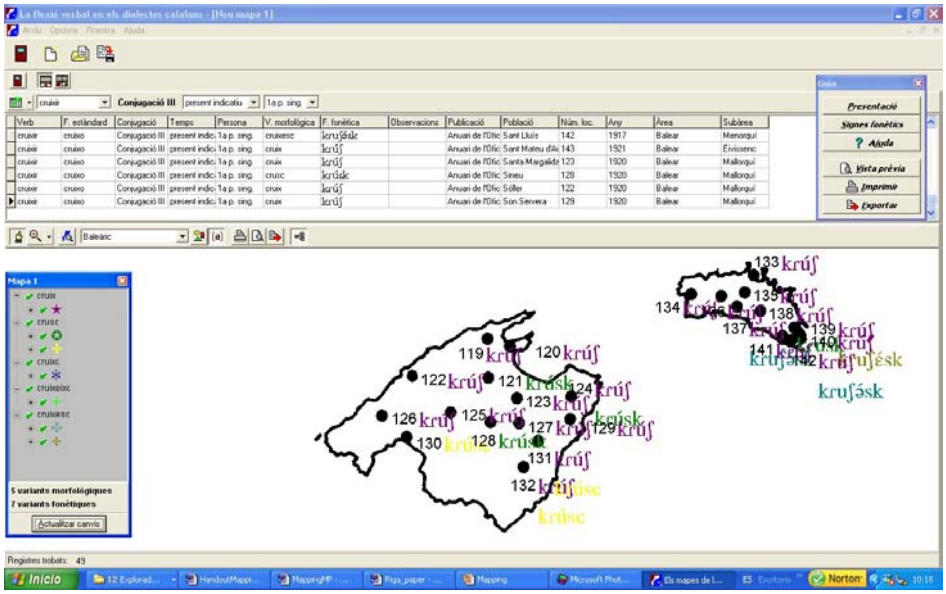


Figure 5. Majorca and Minorca verbal variation

Both types of descriptive maps include the phonetic transcription or a series of symbols to represent the data. In the case of Alcover's data, the inclusion of symbols is particularly useful for areas in which the results are hard to read because of the high concentration of localities.

In addition to the explicit distribution of the results through phonetic transcription or symbolic representation, maps include an implicit outline of isoglosses, which show the morphophonological variation present in either the entire Catalan domain or in a given dialectal area.

In order to compare the results, two maps of different areas can be contrasted. A comparison can also be made between the results of the general map and each map of the different dialectal areas.

Each map, as well as the list generated by applying specific search criteria, can be saved on the hard disk for later consultation. Maps and lists can also be printed.

5. The sound maps

The development of computerised procedures, the processing of linguistic data and dialectal mapping resources have generated the so-called "sound atlas", which includes not only the phonetic transcription but also the corresponding sound entry. In some cases, the sound may be accompanied by its spectrographic representation. Various works by Goebl represent some of the pioneering initiatives in the field (1992: 397-412; 1994: 158-168; 2004: 89-115).

With respect to the materials of the COD, the database corresponding to the locatives is a sample form of sound atlas, which configures different sound maps. The chosen format for publication is CD-ROM.

For Alcover's data, a project to develop a synthesised voice is now in progress in order to create sound maps. Ultimately, each synthesised voice entry will be associated to each phonetic representation. This project is still in its initial stages but preliminary results have been obtained.

6. Conclusions

The mapping program was created by applying IT technology to the methods of traditional linguistic geography in the processing of dialectal data and the production of linguistic atlases. The creation of computerised maps reduces expense, simplifies resources and minimises the distortion in the cartographic representation of the data.

By mapping the data from *La flexió verbal* and the COD we obtained an atlas of the Catalan linguistic domain that presents a complete panorama of morphology and phonetics covering almost a century. Although the dates of the two corpora are different, the numerous correspondences will allow the development of both synchronic studies and also diachronic analysis through comparison with current data.

Maps provide a visual representation of the data, aiding researchers in the comparison of results and the creation of dialectal borders. In addition, this program illustrates that the application of IT to the treatment of dialectal corpora and the production of interpretative maps have become essential components of modern dialectology. The methodology that has been applied could be extrapolated to other types of data – not only materials taken from old atlases but also from corpus of phonetic, morphological or lexical typology.

The concept of the linguistic atlas, traditionally understood a book of maps, is now changing. If the main purpose of the linguistic atlas is to offer an approximate representation of the dialectal reality in a certain period, the most suitable format of edition is through computerised procedures. Publishing the results on CD-ROM has a number of advantages, particularly when the data is part of a closed corpus, taken from existing materials that do not need to be updated. However, this type of publication also has its inconveniences: in few years new versions of Windows – if this platform is used – may no longer run publications on CD-ROM that were perfectly compatible with previous versions. In this case it would be necessary to re-release the work so that it remains available for a number of years. The obvious question is how long this process of updating can remain viable.

The use of maps available on the internet is another possible and desirable format for the publication of dialectal materials. This would be a suitable platform for corpus and atlases that are in the process of transcription and systematisation. The web design would allow the data to be modified from the first steps of the process, and it would be

possible to update them periodically. It would therefore be necessary to create standards for the web edition of the linguistic atlas, with the possible incorporation of the corresponding sound entries. These standards would have to be sufficiently flexible to adapt to the characteristics of each project and the linguistic variety used. They would also have to allow maps to be drawn quickly and easily. Perhaps not long from now this proposal will be a reality.

7. References

ALCOVER, Antoni M. & F. de B. MOLL (1929-1933) *La flexió verbal en els dialectes catalans, Barcelona*, Anuari de l'Oficina Romànica (v. II (1929) [73] 1- [184] 112, v. III (1930) [73] 1- [168] 96, v. IV (1931) [9] 1- [104] 96, v. V (1932) [9] 2 - [72] 64).

ALTURO, Núria, Emili BOIX & Maria-Pilar PEREA (2002) "Corpus de Català Contemporani de la Universitat de Barcelona (CUB). A general presentation", in C. D. Pusch & W. Raible (eds.), *Romanistische Korpuslinguistik. Korpora und gesprochene Sprache*, Tübingen: Günter Narr, 155-170**.**

BADIA, Antoni M., Lídia PONS i GRIERA & Joan VENY (1993), *Atles Lingüístic del Domini Català. Qüestionari*, Barcelona: Institut d'Estudis Catalans.

FRANCIS, W. N. (1983) *Dialectology. An introduction*, London & New York: Longman.

GIMENO BETÍ, Lluís (1997), *Atles Lingüístic de la Diòcesi de Tortosa*, Barcelona: Institut d'Estudis Catalans.

GOEBL, Hans (1992) "L'atlas parlant dans le cadre de l'Atlas linguistique du ladin central et des dialectes limitrophes (ALD)", in *Actes du Congrès Internacional de Dialectologie/Nazioarteko Dialektologia Biltzarra*, Agiriak (Bilbo, 1991), Academie de la Langue Basque, Bilbo, 397-412.

GOEBL, Hans (1994) "L'Atlas Linguistique du Ladin Central et des dialectes limitrophes (première partie, ALD I)", in García Mouton, P. (ed.), *Geolingüística. Trabajos europeos*, Madrid, Consejo Superior de Investigaciones Científicas, 155-168.

GOEBL, Hans (2003) "Regards dialectométriques sur les donnés de l'Atlas Linguistique de la France (ALF): Relations quantitatives et structures de profondeur", *Estudis Romànics*, XXV, 61-117.

GRIERA, Antoni (1923-1964), *Atlas Lingüístic de Catalunya*, Barcelona: Institut d'Estudis Catalans/Ediciones Polígrafa.

LLORET, Maria-Rosa & Maria-Pilar PEREA (2002) "A report on 'The Corpus Oral Dialectal del Català Actual (COD)", *Dialectologia et geolinguistica*, 10, 59-76.

NAVARRO GÓMEZ, Pere (1996) *Atles Lingüístic de la Terra Alta,* Tarragona: Diputació de Tarragona.

NAVARRO TOMÀS, Tomàs (1962) *Atlas Lingüístico de la Península Ibérica*, 1, Madrid: Consejo Superior de Investigaciones Científicas.

PEREA, Maria-Pilar (1999) *Compleció i ordenació de* "*La flexió verbal en els dialectes catalans" d'A. M. Alcover i F. de B. Moll*, Barcelona, Institut d'Estudis Catalans, 1999 (& CD-ROM).

PEREA, Maria-Pilar (2001a) "La metodologia de l'enquesta dialectal d'A. M. Alcover i F. de B. Moll aplicada a *La flexió verbal en els dialectes catalans*", *Zeitschrift für Katalanistik***,** 14, 98-127**.**

PEREA, Maria-Pilar (2001b) «Les dades de *La flexió verbal en els dialectes catalans* d'A. M. Alcover i F. de B. Moll. Una aproximació sociolingüística», *Catalan Review. International Journal of Catalan Culture*, 53-77.

PEREA, Maria-Pilar (2001c) *La flexió verbal en els dialectes catalans d'A. M. Alcover i F. de B. Moll. Les dades i els mapes*, Palma de Mallorca: Conselleria d'Educació i Cultura. Govern de les Illes Balears, CD-ROM edition.

PEREA, Maria-Pilar (2004) "New Techniques and Old Corpora: La flexió verbal en els dialectes catalans (ALcover-Moll, 1929-1932). Systematisation and Mapping of a Morphological Corpus", *Dialectologia et geolinguistica*, 12, 25-45.

PEREA, Maria-Pilar (2005) *Dades dialectals. Antoni M. Alcover*, Palma de Mallorca: Conselleria d'Educació i Cultura. Govern de les Illes Balears, CD-ROM edition.

VENY, Joan & Lídia PONS I GRIERA (2001, 2004, 2006), *Atles Lingüístic del Domini Català*, vol. I, II y III, Barcelona: Institut d'Estudis Catalans.