Map names generalization at the Institut Cartogràfic de Catalunya

Blanca Baella, Maria Pla
Institut Cartogràfic de Catalunya
Parc de Montjuïc
E-08038 Barcelona
Telephone 34 93 5671500, FAX 34 93 5671567

Introduction

In 1984 the Institut Cartogràfic de Catalunya (ICC) started the publishing of the Orthophotomap of Catalonia at scale 1:5.000, the first series at large scale that covers all the country, approximately 32.000 km2. The maps included, moreover the orthoimage, map names and marginalia. The map names were compiled using existing data and carrying out a new fieldwork campaign, trough local administration consults or people interviews. In 1991, the total coverage was completed and it included about 350.000 map names. From the beginning, the orthophotoimage was a digital product obtained using software developed at the ICC, but until 1989 the photomontage of the orthophotomaps was based in manual techniques, like stripping for the map names. The implementation of workflows fully digital for map publishing, forced, in particular, the design of a map names model. Moreover the publishing of the maps, the goal was also the generation of map names data for GIS databases.
Very soon, the availability of a total coverage of map names at scale 1:5.000, jointly with the increasing demand of cartographic products at smaller scales, generated the need of map names generalization tools. The existing commercial software was focused on fulfilling the map publishing standards or on covering GIS requirements, but neither took into account both aspects at the same time. The ICC choice was to design a data model and to develop software applications for map names placement and management, suitable to ICC needs.

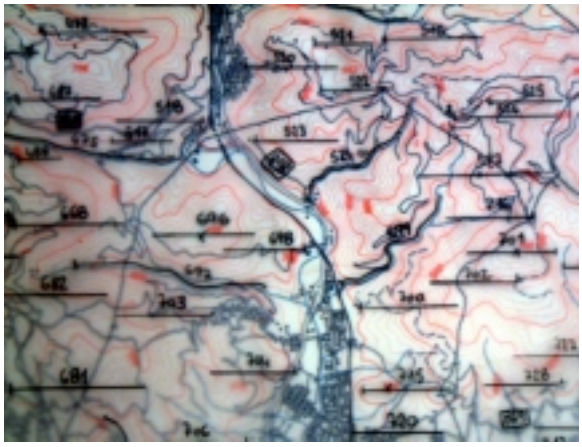
The ICC model for map names

INPUT DATA

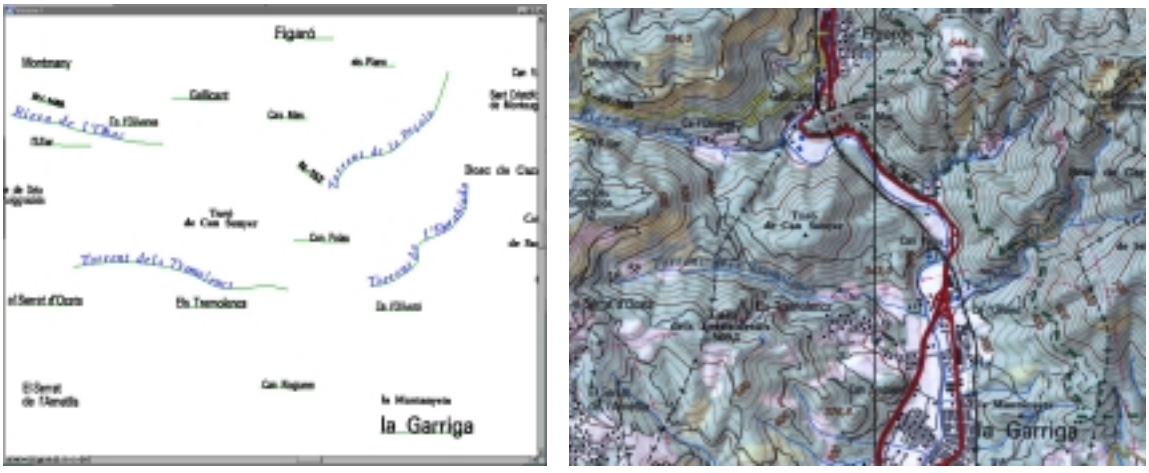
The map names compiled in the fieldwork campaign, after a complete revision process, were introduced in an Excel file, which contained also the attributes for each map name. The georeference was drawn manually on top of an existing paper map of the same area. Both information, the Excel file and the paper map, were used as input data to digitize map names.

Codi seqüencial	Codi geogràfic	Codi tipogràfic	Longitud (mm)	Topònim
696	50203	56	0	Turó de Can Sunyer
697	60303	48	0	Torrent dels Tremolencs
698	10301	16	0	Can Palau
699	60303	48	0	Torrent de l'Enrabiada
700	10301	16	0	Ca l'Oliveró
701	50203	56	0	Turó de Santa Margarida
702	10301	16	0	Can Parades
703	10105	17	0	els Tremolencs

Detail of an Excel file with the compiled map names and their attributes.



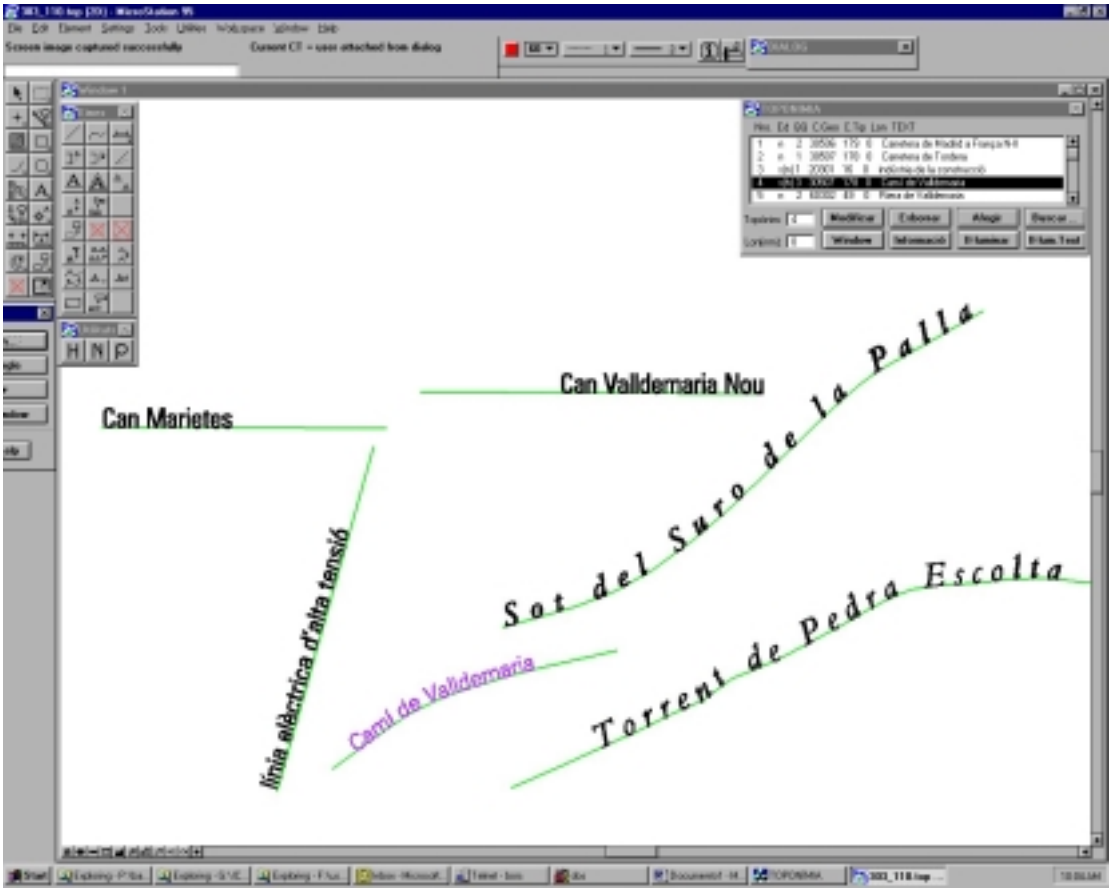
Detail of a paper map with the map names manually georeferenced.



Right image is a detail of a published map using the digital map names showed in the left.

DATA MODEL

The model structure is based in two types of information, graphical and alphanumerical, linked by an identifier. The georeference, the text shape and the character distribution compose the graphical part of the structure, which is stored in DGN files, the file format of the software Microstation from Bentley. For each map name placement, exists an associated line or curve that defines their situation.



Display of the graphic information associated to the map names, using Microstation.

The alphanumerical information contains data about the geographical, administrative and typographical properties of the name. For each map name, the following parameters are stored: a numerical identifier, a geographical code that indicates the topographic class, an administrative code, a typographic code that assigns the typographic characteristics, a bit mask that indicates if the digital map name has been modified, the length in millimeters of the text and the number of placements. For each map name placement, the parameters are: a numerical identifier, the X and Y coordinates, the length of the text if it is different from the assigned previously to the map name, a code that indicates the method of placement as horizontal, along a curve, etc.
Both data are generated and managed by a nucleus of software developed at the ICC.

MAP NAME DATA								MAP NAME PLACEMENT DATA			
Number	Edition mask	Geog. code	Typog. Code	Adm. Code	Map Name	Lon	Number of placements	Coordinates	Lon	Mode	Number
4	101000001	30507	178	0	Camí de Valldemaria	0	3				
								47845909,461895148	12	2005	11
								47805086,461781038	12	2005	27
								47881309,461978812	0	2005	29
19	10000000	60303	49	0	Torrent de Pedra Escolta	0	2				
								47899420,461965108	40	2005	30
								47993264,461978513	43	2005	33
21	10000000	10301	16	0	Can Valldemaria Nou	0	1				
								47899811,461990399	0	3002	6
22	10000000	50304	39	0	Sot del Suro de la Palla	0	1				
								47907283,461983657	45	2005	31
23	10000000	10301	16	0	Can Marietes	0	1				
								47821739,461971301	0	1002	1
40	10000000	30605	16	0	línia elèctrica d'alta tensió	0	2				
								47857292,461959104	0	1004	5
								47811913,461807371	0	1004	26

Display of the alphanumerical information associated to the map names.

Each project has a particular set of specifications that are managed by a new application, which is designed ad hoc and uses the previous nucleus of software. Specifications include constraints over alphanumerical values or special representation methods.

The pressure to start up the digital publishing workflow impeded a deep analysis and development for the GIS requirements. The graphic software presented also at that time restrictions in the text management and in the facilities for customized tools implementation.
At the present, ten years after the original design and after the exhaustive use of the model and the applications, limitations are clear. Some of them have been minimized by the continuous software improvements. Other ones, as the data continuity, are being developed at the moment. But some limitations cannot be solved without changing the data model or the graphic software Microstation. The design of the new data model requires a detailed analysis of the problems related with links between topographical objects and map names, multiscale storage and representation, versioning, multiple typographies for a single map name placement, etc. Otherwise, it seems that themes as the improvement in the kerning of text placement have to be provided by the graphic software to be used.

ICC Map Names databases

As has been mentioned before, the first total coverage of map names for 1:5.000 cartographic products was completed in 1991. From 1989, the information was digitized observing the data model designed at the ICC, and organized sheet by sheet, penalizing the global continuity of the data: multiple placement for an object name inside the same sheet are occurrences of a single map name, but multiple placements in different sheets, become a different map name for each sheet. Although this, the digital information was called Map Names Database at scale 1:5.000. This database was completed in 1992 and it is updated simultaneously with the two existing series at this scale, the Orthophotomap and the Topographic Database of Catalonia 1:5.000.

Example of a map of the Orthophotomap of Catalonia at scale 1:5.000, 2001.

Detail of an orthophotomap, showing several types of map name placements. It shows an exploitation example of the Map Names Database at scale 1:5.000.

In 1991, a new topographical series at scale 1:50.000 was started. The first maps to be published belong to areas where map names data didn't exist yet in digital format and any exploitation of the Map Names Database at scale 1:5.000 was possible. Realizing a selection for a 1:50.000 scale of the compiled data, a new database was created, the Map Names Database at scale 1:50.000, which was completed in 1995.

Detail of the Mapa Comarcal de Catalunya 1:50.000. Alta Ribagorça, 1995. It shows an exploitation example of the Map Names Database at scale 1:50.000.

Map names generalization

The availability of the map name databases with a total coverage of the country, one at large scale, 1:5.000, and another at medium scale, 1:50.000, obliged to find specific tools for data exploitation in order to obtain cartographic products at smaller scales. That is to say, map name generalization tools. At early 90's, several generalization applications appeared in the market, but none of them take into account the map name generalization requirements. At present, some functionalities, as selection or scaling, can be found in the commercial software. These tools are GIS oriented and have serious limitations to obtain high quality cartographic products. In front of this situation, the ICC decided to develop map name generalization tools that take advantage of the existing data model. The previous experience in the generalization of topographic objects, recommended the development as of automatic tools as of assisted editing tools.

AUTOMATIC GENERALIZATION

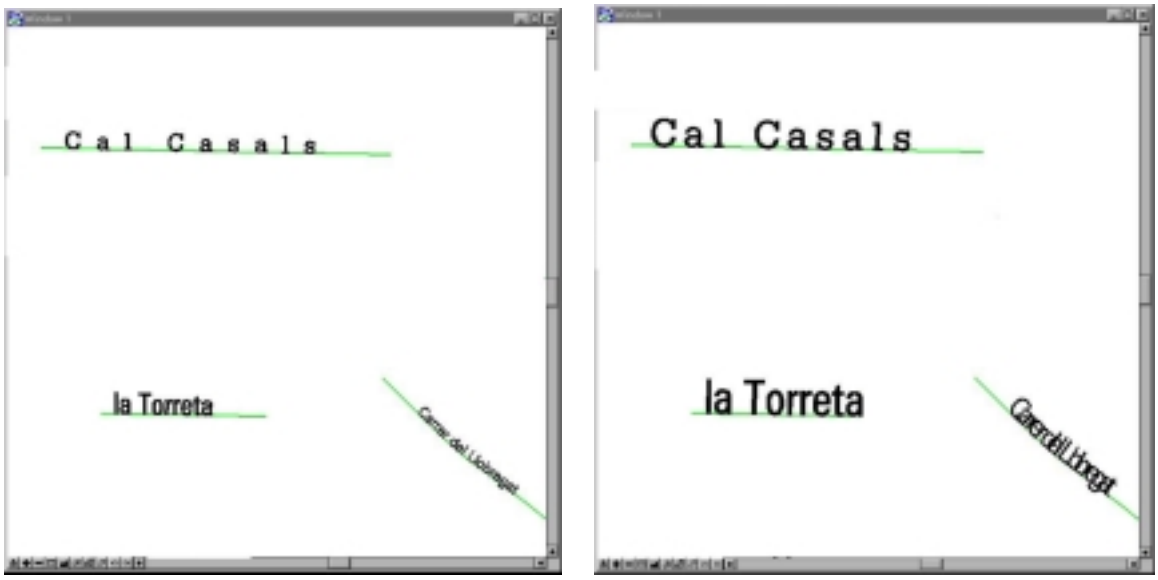
The automatic generalization tools include selection, cartographic scaling and merge of map name data. The parameters for selection are geographical code, typographical code or a combination of both. The cartographic scaling consists in a new map name

placement with a new typographical code, a new intercharacter spacing and a new line spacing adapted to the smaller scale.

CG 10000
-250 276
-1000 278
-5000 280
-10000 282
-50000 284
-100000 286
-9999999 288
CG 10102
17 18
18 20
19 98
20 100
* 18
CG 10104
17 18
18 20
19 98
20 100
* 18
CG 60302
* *
CG ?????
56 58

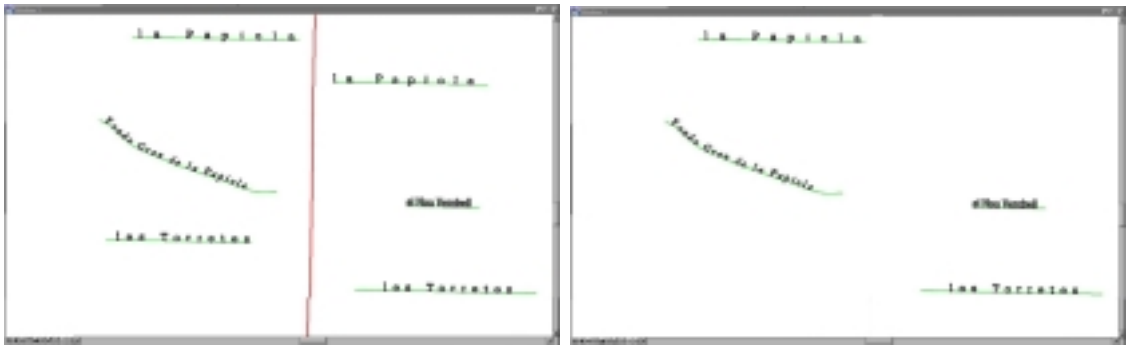
Example of parameters for map names generalization. In the eight first lines, a new typographical code for population map names, geographical code (CG) 10000, is assigned depending on inhabitants. In the following lines, new typographical codes are assigned depending on the geographical codes. Combination of codes that don't appear in the list, are not selected.

The automatic generalization tools developed at the ICC don't solve all the situations. Main problem appears in the map names associated to curves, almost the 70% of the total number of map names. When the map names are associated to a line and they have the default intercharacter spacing, the generalization tool extends the line to achieve the new placement, recalculating all the text parameters. When they are associated to a curve or to a line with a specific intercharacter spacing, the generalization tool can decide neither the extension of the curve nor the new intercharacter spacing. In this case, the text is situated with the generalized parameters but maintaining the original intercharacter spacing. Correct intercharacter spacing should be assigned by cartographers using interactive editing tools.



Left image shows the original data, right image shows the automatic generalized data. Bottom horizontal text has been generalized with the correct typographical parameters: type, character size and intercharacter spacing. Original top horizontal text has a specific intercharacter spacing, although it has not been changed after the generalization, the result is acceptable. Bottom right text, originally placed along a curve, maintains its intercharacter spacing but the result is incorrect.

Another problem is derived from the limitation on the global continuity in the model. When the map generalization requires a merge of several sheets, multiple placements for the same object map name can became different map names, if the placements came from different sheets. In the generalized data, two problems appear. The first one disturbs the model structure, because all the placements should be occurrences of the same map name. Tools for solving automatically this problem are been developed at the moment. The second one can be the density of occurrences for each map name. The generalized data, at smaller scale, don't need all the multiple map name placements present in the original data. By the moment, this problem only can be solved by manual selection helped by interactive editing tools.

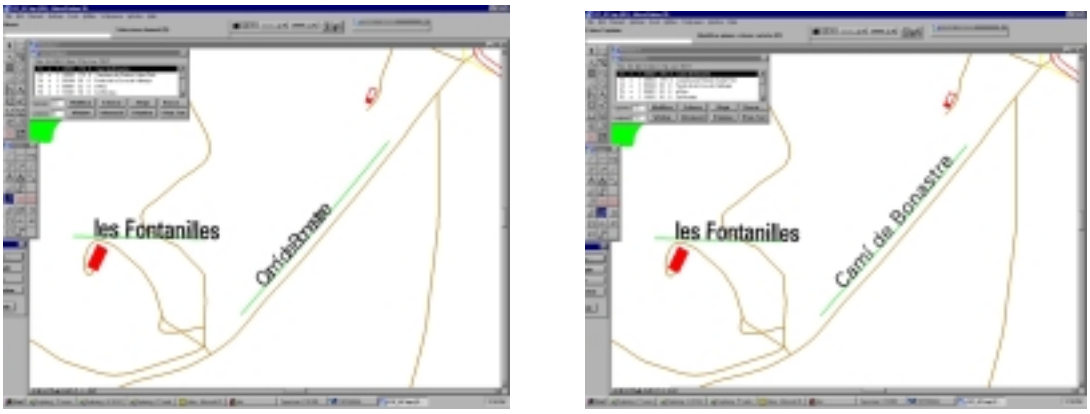


Left image shows too many occurrences for the same map name in the generalized data (red line indicates the boundaries of the adjacent original sheets). In the right image is showed the result after the manual selection.

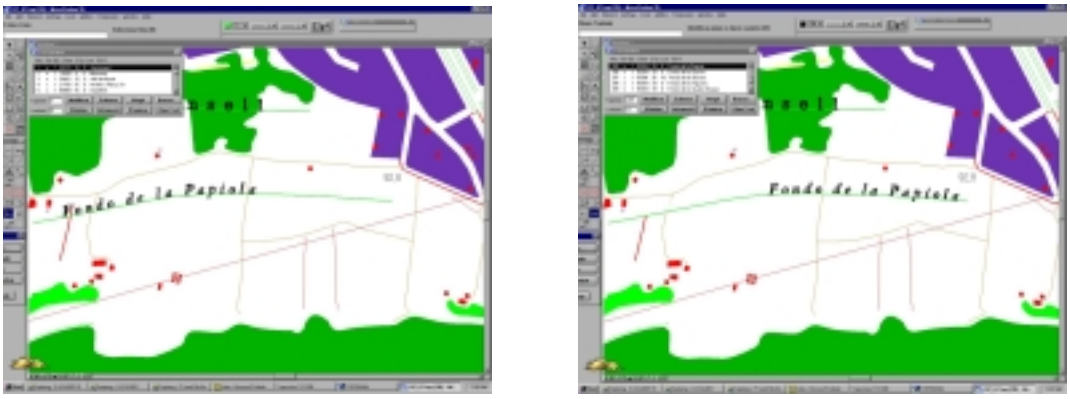
INTERACTIVE TOOLS

A set of interactive tools has been developed to help cartographers in the manual editing tasks, managing together the graphical and the alphanumerical components of the model. These tools are used to improve the result of the automatic generalization and also to realize the aesthetical refinements required for cartographic publishing:

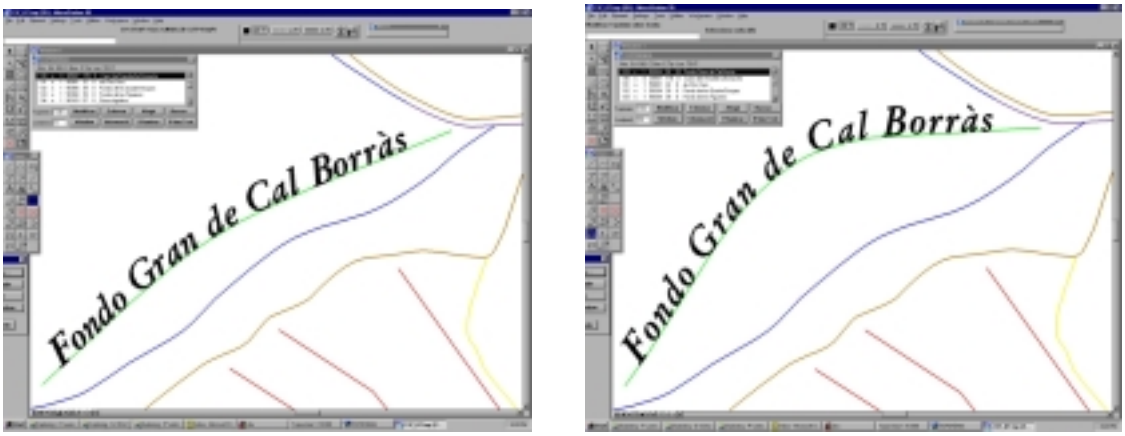
- Deleting tools allow improving the automatic selection, eliminating single map name placements.
- Stretching tools for intercharacter or line spacing, help to the exaggeration operations.
- Moving and rotating tools allow solving overlapping conflicts maintaining the shape.
- Movements of the text along the associated curve or modification on the curve, allow modifying the position changing the shape.
- Changes in the geographical or typographical codes allow adapting a map name to the particularities of the generalized product.



Example of the stretching tool. Left image shows the result after the automatic generalization. Right image shows the stretched map name.



Example of movements along the associated curve. Left image shows the result after the automatic generalization. Right image shows the new map name shape after the movement.



Example of modifications in the associated curve. Left image shows the result after the automatic generalization. Right image shows the new map name shape after the modification.

Moreover, some tools have been developed to show, analyze and modify the database contents:

- Addition of a single map name manually or a new compilation set of map names from an Excel file.
- Display of map names attributes.
- Display and queue of map names that fulfill queries on some attributes: geographical code, typographical code, character strings, etc.
- Verification tools for ensure the consistence of the database.

The example shows, in blue, multiples occurrences of a map name.

AN EXEMPLE: THE TOPOGRAPHIC MAP OF CATALONIA AT 1:10.000

The Institut Cartogràfic de Catalunya has been producing since 1999 the Topographic Map of Catalonia at scale 1:10.000 from the Topographic Base at 1:5.000 using generalization tools. The series has 1.070 sheets covering all the country and until now 230 has been published. The workflow includes, in the map names process, the generalization tools mentioned above.
The following table shows the average time spent in the whole process for one sheet. The map name generalization timing is indicated in red.

GENERALIZATION PROCESS		**TIME**	
Automatic processes			
	Automatic map names generalization	1‘	
	Other automatic generalization	6‘	
Total automatic			**7‘**
Manual assisted editing			
	Manual map names editing	5 h	
	Manual spot height selection	7 h	
	Manual conflict resolution	4 h	
	Manual aesthetic refinement	4 h	
Total manual			**20h**
Data management			**1h 30'**
Total per map sheet 1:10.000			**21h 37’**

As it is usual in the generalization workflows, the time spent in the automatic process is negligible. The saved time is not easy to evaluate. An estimation, based in the first map names digitalization, could be that the map names generalization, which spends now 5h 1’, would spend about 20h to be done manually. That is to say, a 75% of the time is saved.

On the left, detail of the Topographic Map of Catalonia at 1:5.000. On the right the same area of the generalized map at 1:10.000.

Conclusions

The efforts devoted to the research and the development for map names generalization have been until now very low compared with the efforts invested in topographical data generalization.

The generalization tools have to take into account as the GIS needs as the map publishing requirements.

The map names generalization could be more automatic if the data model will take into account other aspects than the typographical or placement parameters. Links to the topographical object, continuity, versioning and multiscale storage and representation, are essential to achieve an optimum management and exploitation of the database, and also to improve the results of automatic generalization processes.

Institut Cartogràfic de Catalunya
31st January 2001