

# **Ofercat. Indicators of the availability of Catalan: An instrument at the service of language planning**

by *Mercè Romagosa, Pilar López and Albert Fabà*

## **Abstract**

Languages need instruments to enable us to measure their vitality and power, and then go on to design language policies that will promote them. Census studies on language knowledge and surveys of the population to obtain data on the use of these languages provide essential information. After all, the use that citizens make of a language will be determined mainly by three factors: the knowledge they have, their will to use it, and lastly, the actual opportunities they have to use it

These possibilities are conditioned by the presence or availability of the language in a given territory. For that reason it is important to have tools that allow us to gauge and calibrate that presence at any one moment and go on to observe the ebb and flow of the supply over time.

Ofercat is an instrument that measures the presence of the Catalan language in a population. The results obtained by this means are the necessary complement to data on knowledge and use of the language provided by the censuses and voting lists, the questionnaire surveys and other studies. It will be invaluable in the designing of language policies implemented by the different institutions involved in this process.

## **Summary**

1. Introduction
  - 1.1 Stages of the Ofercat Project
  - 1.2 What is Ofercat?
2. Technical features and methodology
  - 2.1 Computer application
  - 2.2 General aspects
  - 2.3 Random sampling
  - 2.4 Visits to organisations
  - 2.5 Telephone calls
  - 2.6 Other observation
  - 2.7 Further information already included in the programme
3. Main results
4. Evaluation and prospects for the future

## **1. Introduction**

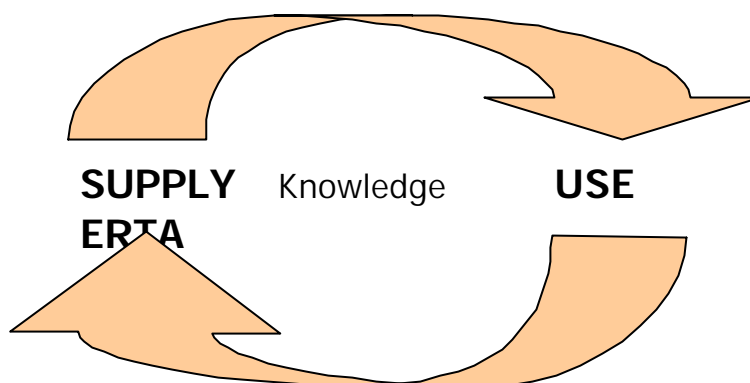
Languages need instruments for measuring their vitality and power, and subsequently shaping language policy that will promote them. For that reason, governments adopt such useful tools as census studies, which usually collect information on knowledge of languages. In Catalonia, what's more, a number of different organisations and institutions, both public and private, carry out surveys among the population to obtain information on language use.

This type of information is in fact essential, since the use made by citizens of a language is largely determined by three factors: the knowledge that they have of them, the individual and collective will to use them, and lastly the actual opportunities they have of using them. And these opportunities are conditioned by the language supply, the languages that these citizens have available in a particular territory.

This, it is important to adopt tools that will permit one to measure this supply at any one moment and be able to observe the trends apparent in this aspect over time.

It is within this context that we present Ofercat, as an instrument that attests the presence or otherwise of the Catalan language in a population. This presence in the public domain can be measured as an

outcome of the language offer inhabitants have access to in their everyday lives, both as users of public or private services, and as customers and tax payers - as citizens in fact. What do we mean by language supply or languages on offer? The answer is multiple, but quite simple: can citizens, walking through the city where they live, read signs and notices in the shops in Catalan? Are they attended in Catalan when they telephone a local government department or a service company? Are the forms to renew their identity card in Catalan? Can they see a movie in Catalan or dubbed into Catalan? In short, do the organisations in their city, public or private, large or small, offer them communication in Catalan?



The results obtained by Ofercat are, therefore, a necessary complement to the data on knowledge and use of the language offered by the census, questionnaires and the voting lists or housing registers, and thus can be very useful in the design of language policy by institutions involved in the process. One of the original aspects of Ofercat is that it is not a questionnaire of personal usage with the latter's subjective answers, but rather an observation of public use of the language based on certain previously selected, objective indicators. And it is not just a study, but rather a systematic tool, easy to operate. In a graphic and entertain way it gives mathematic form to aspects that are otherwise difficult to quantify.

### 1.1 Stages of the Ofercat Project

Ofercat was applied for the first time in 1998 and 1999 in the town of Santa Coloma de Gramenet, on the outskirts of Barcelona. In 1999 this project won the runners-up prize of the Badia i Margarit award for empirical sociolinguistic research, and in the same year a presentation was given in Bilbao at the VII International Conference on Minority Languages.

Subsequently, the interest Ofercat generated was gauged, as an electronic tool that can be adapted to different areas and populations, and it was decided to make use of it as a computer application along the lines of Indexplà (see in this issue the article about Indexplà, by Alfred Castells) translated into overall indicators. Furthermore, it was decided that its findings should be displayed in graphs and diagrams that could be used comparatively, comparing graphs across territories, and over time, by sectors and by factors.

Accordingly, in 2000 the Ofercat Commission was set up<sup>1</sup>, consisting of members of the Consorci per a la Normalització Lingüística (Consortium for Language Normalisation) and the Institute of Catalan Sociolinguistics (ISC) of the Directorate General for Language Policy, the latter being responsible for the development of the tool, for the IT applications and for the manuals and handbooks.

---

<sup>1</sup> Ofercat Commission: Alfred Castells, Albert Fabà, Pilar López, Mercè Romagosa, Noemí Ubach (Consortium for Language Normalisation), Joan Solé (Institute of Catalan Sociolinguistics) Alfons Duarte (advisor for IT evaluation). External advisors: Josep M. Aragay and Albert Bastardes (sociolinguists).

In 2001 the tool was piloted to find out how effective it was in different types of settlements. Five districts were chosen based on the linguistic coefficient derived from the 1996 census, the main economic activity of the settlements (tourist areas, rural areas) and the associated population density. The aim was to have represented the diversity of the overall territory, bearing in mind, additionally, that these were settlements for which other data was available (studies of signs and labels, surveys etc.).

Finally the following settlements were chosen: Tarragona, Manresa, Tremp, Lloret de Mar and Barcelona old city (Ciutat Vella).

In 2002 the pilot study results were made use of, there was the first public presentation at the World Congress on Language Policies and the final touches were put to the IT tool, which would permit the application of the consolidated Ofercat product to Lleida, Reus, Olot, Sabadell and Santa Coloma de Gramenet, in 2003.

### 1.2 What is Ofercat?

Ofercat is a methodological, IT tool that gauges the supply of Catalan in any Catalan-speaking population. Using a series of indicators, Ofercat enables us to take a snapshot -at a particular point in time- of the degree and extent of the presence of Catalan in a given territory. This measurement can be taken at intervals when changes in the situation are suspected.

### 1.3 Characteristics

- Using a series of prefixed indicators (124), the programme summarises, in numerical indices and in different graphs, the linguistic supply in a specific population.
- Ofercat evaluates the linguistic usage of the individuals in question as representative of an organisation and not in terms of their private conversations.
- The application engages with the oral and written outputs of both public institutions and private entities as well as of the civil society itself.
- The selected indicators are easily data-collected and can be observed periodically.
- The indicators are grouped by factors according to the type of communication referred to. The written language factors observed are: labelling for identification and for providing information; the oral language factors are: oral identification and the language of oral convergence. These terms will be explained below.
- The population under study is divided into five sectors: Public administration, Economy, Society and leisure, the Media, and Education and health. Each sector is subdivided into subsectors, which in turn are divided, providing a third level of detail we refer to as ambits.

## Ambits, subsectors and sectors observed

Ambit	Subsector	Sector
Local government offices	Local government	Public Administration
Local government town planning		
Central government offices	State administration	
Offices Generalitat (Catalonian government)	Generalitat administration	
Offices of the courts	Courts	
Business, professional and trade union associations	Economy	Economy
Primary and secondary sector companies		
Retail and service businesses		
Public transport		
Shopping centres and hypermarkets		
Outdoor advertising (bill drops, hoardings...)		
Associations	Society	Society and leisure
Cinema and theatre	Leisure	
Public sports events		
Leisure areas (theme parks, aquatic parks, etc.)		
Magazines	Printed media	The media
Newspapers		
Television	Audio-visual media	
Radio		
The Internet		
Nurseries	Education	Education and health
State primary education		
Private or concerted primary education		
State secondary education		
Private or concerted secondary education		
State tertiary education		
Private or concerted tertiary education		
Adult or further education	Health	
Public health service		
Private or concerted health service		

## 2. Technical features and methodology

### 2.1 Computer application

There is an easy to use and clearly useful computer programme which can be utilised both for the organisation of observations and for the introduction and processing of data, as well as for the drafting of reports based on the exportation of graphs manufactured directly by the programme. This is a 32-bit application for Windows constructed from C++ and Visual Basic, which takes up only 20 Mb of space in the hard disc.

### 2.2 General aspects

In general, all Ofercat data is derived from data collection using sociolinguistic methods of observation, and in fact four different techniques are combined: random route observation of commercial premises and service sector companies, visits to organisations, telephone calls and other observation.

Depending on the district in question and its territorial setting, samples were selected of the different organisations and random routes taken for the effecting of observations.

### 2.3 Random sampling

Statistical samples are those constructed by random selection. Strategic samples are obtained when we work in strategic areas or areas with a great concentration of commercial premises. Where necessary, both methods were used in combination. With Ofercat, in the case of the samples of commercial premises and service sector companies, we worked with a margin of error of  $\pm 5$ .

To calculate the total number of routes, the number of commercial premises and service sector companies in the sample for each area were divided, approximately, by 50 shops and companies per route. Along these routes we observed identification labelling and informative labelling in the shops and premises, the advertising in the street, and the signs.

The number of routes was selected in the different areas of the town or city in question proportional to the number of commercial premises per area, where we had this information, and in an even distribution where we lacked such information. Specific routes were also selected going to the shopping centres and the main covered markets.

### 2.4 Visits to organisations

In the organisations visited we observed the two types of labelling: identificational and informative. In the case of identificational notices, the main notice identifying each office or premises was selected. In the case of information notices, the first three notices of this type that were seen were included in the sample.

The number of organisations (institutions, bodies and boards, associations and other organisations) included in the sample, was determined in accordance with table 1.

**Table 1. Number of organisations observed**

To determine the number of offices, customer service points and organisations (other than the shops and companies) that we were to observe, we were guided by the following table:

Total number of offices	Number of offices to observe
from 0 to 30	all
from 31 to 50	30
from 51 to 100	50
more than 100	70

### 2.5 Telephone calls

The initial language of identification and the language of the subsequent response from the interlocutor at the organisation were noted.

*Initial language:* the language used by the person answering the telephone before asked any question is made to them. To avoid ambivalent cases ("Si?" ("Yes"), for example, which is the same in Spanish as in Catalan) or in the cases where nothing is said, the caller would remain silent, to force the person to say something. Where nothing is said, the telephone call was null and void, at least as far as the initial response question was concerned.

*Language of Response:* This is the language used by the interlocutor in response to the question asked of them in Catalan.

Two types of telephone call were made: those made to the sample of shops and service companies, selected at random from the Yellow Pages, and those made to the other organisations. The first were gathered according to the grid for telephone calls to shops and service sector companies. In the second cases observers were supplied with a list featuring the telephone number of the organisations to be telephoned (in most cases, these were the same organisations that we visited personally, with the exception of the schools, which were not telephoned).

## 2.6 Other observation

Other data was collected by other than the above-mentioned methods. These included:

- Media with the greatest print run or circulation of the various organisations
- The language of the websites
- Specific random routes: factory estates and shopping centres
- Written communications sent to the service industries' customers - water, electricity, gas and telephone.
- Tickets issued by public transport
- Advertisements: shops and surveys, and adventure playgrounds
- Church services and ceremonies
- Local press (magazines, television and radio)
- Language of education

## 2.7 Further information already included in the programme

These documents, from centralised organisations, which are the same throughout Catalan territory, were collected via the Ofercat Commission and, therefore, their results are included directly in the computer programme.

## 3. Main results

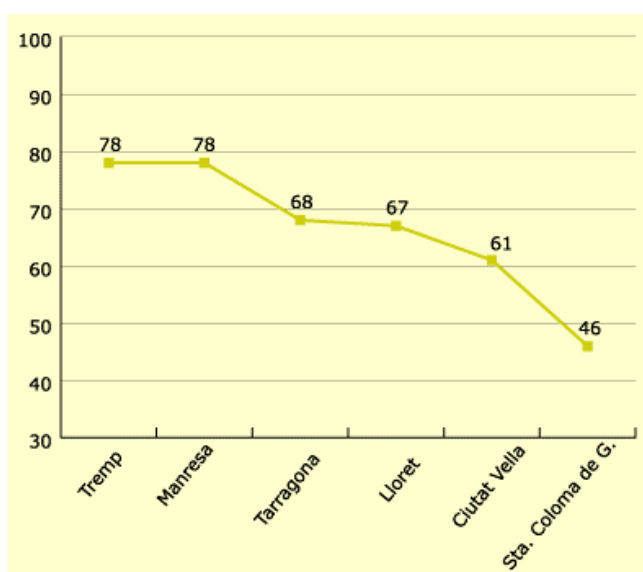
At the present time we have the results of the pilot Ofercat trial carried out in Santa Coloma de Gramenet [suburb with many residents originally from outside Catalonia], and the first cycle of observations, carried out in the towns and cities of Tremp, Manresa, Tarragona and Lloret de Mar, and the Old City district (Ciutat Vella) in the centre of Barcelona, during 2001.

Consequently, the data we offer refer to the five cities and the district of Ciutat Vella was mostly obtained in 2001, with the sole exception of Santa Coloma de Gramenet, which date from 1998. In spite of this time lapse, we preferred retaining the Santa Coloma data, given that it is the only city in the inner ring of satellite towns for which we have data at this moment.

At the same time it should be noted that in 1998, in Santa Coloma, not as many observations were included as in the *standard* version of Ofercat, and accordingly this result has to be taken as provisional. However, the information we now have, gathered in 2003, of the same town, which now includes a sufficient number of observations, indicates that the index we have for Santa Coloma dating from 1998 constitutes a good approximation.

First of all we will present the general indices, and move on to the results of each of the subsectors and, lastly, we will present averaged figures for the main subsectors and ambits, and for the district of Santa Coloma de Gramenet by factors.

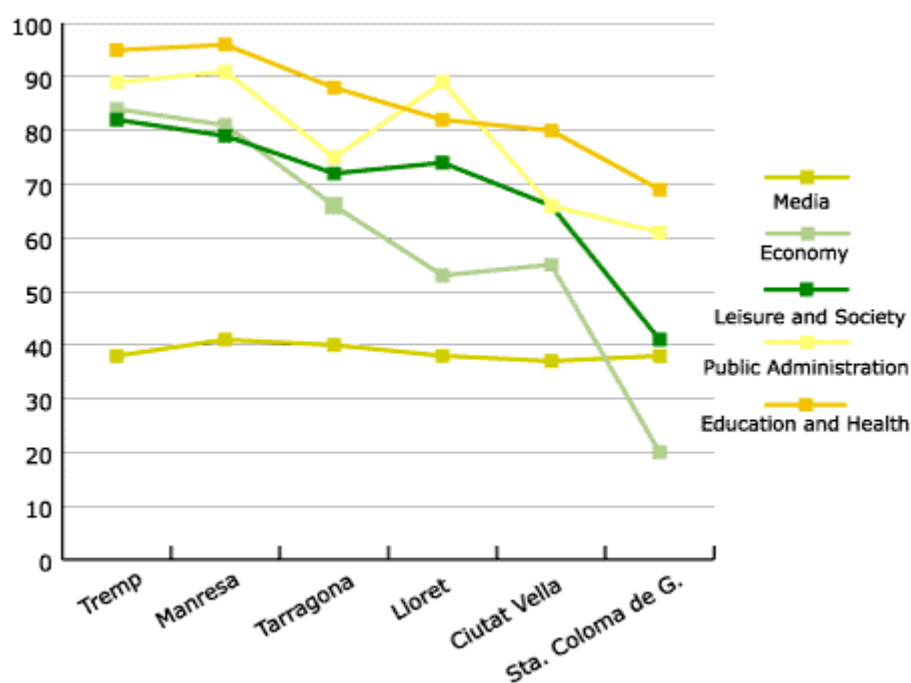
Figure 1. Ofercat Results by towns



As the graph shows, the availability of Catalan goes from a maximum value of 78 (in the case of Tremp and Manresa), to a minimum of 46 (in the case of the Barcelona suburb Santa Coloma de Gramenet). What we have here, therefore, is a rather marked step or range of 32, seen in the distance that exists in terms of oral and written competence in Catalan, between one place and another, according to the 1996 linguistic census. Thus, between the 89.36% of the population who claim to speak Catalan in Tremp, and the 52.2% who claim to do so in Santa Coloma de Gramenet, there is gap of 37.16 %; and between the 55.71% in the Manresa sample who said they knew how to write it, and the 32.61% once again in Santa Coloma, there is 23.10 % difference.

Even so, the gradient is not even, but rather hints at different taxonomies, all related to the Catalan availability factor. Clearly, the number of observations or sightings taken to date is still very limited, and it is not possible to draw definite conclusions. We need to wait, therefore, to get the results of the twenty-five towns and cities in the overall Ofercat cycle. Nonetheless, what we have here is a group of populations where the availability of Catalan is majority level (Tremp and Manresa), a group which is quite homogenous; an intermediate group (Tarragona, Lloret and Ciutat Vella), definitely more disperse; and a final downturn, to a "group" where we just have, at present, the results for Santa Coloma de Gramenet.

Figure 2. Ofercat results by sectors



But it is necessary to go into finer detail, and observe the data by sectors as shown in figure 2.

If we calculate tendencies, the sectors which, in most cases, have highest values (score highest) are Public Administration and Education and Health; there is a second block with Leisure and Society, and Economy; and lastly the Media. The only exception to this, in this scheme, is that in Santa Coloma de Gramenet the Economy index or score is lower than the Media one, although, as we indicated above, we have to some extent to take these data as provisional.

What is very clearly observable, once again, is the trend of the lines between sectors, which show a pronounced down curve from Tremp to Santa Coloma, except in the case of the Media, which are much the same everywhere. This is logical and predictable, given that both in audio-visual media and the press, national level distribution predominates over the local or county level.

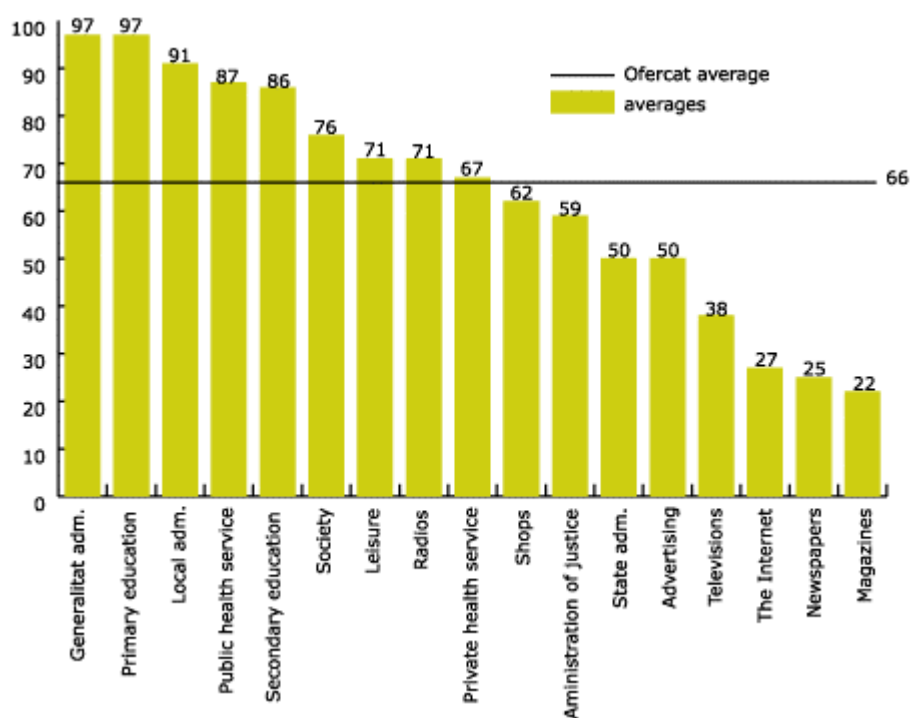
In the same way, the difference between the sector with the highest score, and the lowest (the media excepted) is more accentuated as we go down. Thus, in Tremp, the gap between Education and Health

on the one hand, at 95% and Society and Leisure, at 82% is 13%; in contrast, in Santa Coloma, between the 69% scored by Education and health, and the 20% for Economy there is a gap of...49%!

Differentiating itself somewhat from these general tendencies, the coastal resort of Lloret de Mar shows an upturn for Public Administration, at a level above Education, and one in the opposite direction, owing to the more pronounced downturn for Economy. If we look at the Lloret de Mar data by subsectors or, by ambits, we can appreciate the dual nature of this locality in Girona province: it behaves like a very Catalanised town, on the one hand (with above average scores, for example for the Central Government Administration); but with quite low scores in different subsectors or ambits connected with the principal economic activity of the town, that of tourism, on the other.

To take a closer look at this, we present figure 3, which shows the average indices in the different localities for the principal subsectors and ambits. We have added the Ofercat mean index for the six localities.

**Figure 3. Mean indices by subsectors and ambits, and Ofercat mean score.**



Here the range of scores is even wider, going from 97% for the Generalitat, to the 22 % for the non-daily press, a yawning difference of 75 percent.

Local government offices and those of the Generalitat have the highest scores for presence of Catalan, greater in every case, than the Ofercat mean. The radio stations provide the most paradigmatic instance of this. They bolster the lower mean scores for the rest of the media; they are precisely the part of the media regulated by the Generalitat, while the rest of the media are regulated by central government (the Spanish state).

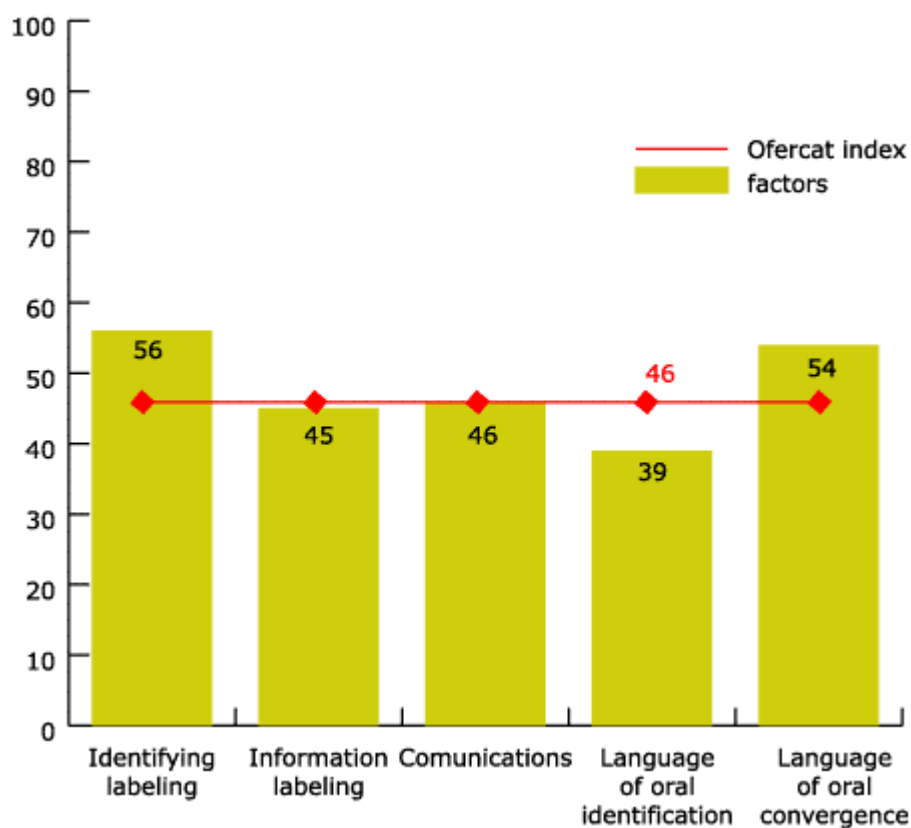
However, it is important to note that Society also scores highly (76, to be exact).

Turning to the subsectors and ambits that score below the mean, these show us where efforts should be concentrated in the future: Economy, Administration of justice and State administration, and the media with the exception of the radio.



Lastly, in figure 4, we show the results for the Barcelona satellite town of Santa Coloma de Gramenet, by factors.

**Figure 4. Ofercat 1998 results by factors at Santa Coloma de Gramenet**



Here it is worthwhile noting the contrast between the language of oral identification and the language of oral convergence. In the first case, as we said above, a note was taken of the language in which the different organisations addressed their customers during the initial turn, while in the second case what was noted was the language used once the observer addressed them in Catalan. The fifteen percent difference between the first score and the second, in the instance of Santa Coloma, is only a nine or ten percent difference in the other areas, but always there is a substantial difference of this sort between the two figures: here, then we have the possibility of growth in this area, as important as it is for oral language use.

#### 4. Evaluation and prospects for the future

As we have seen, the territorial, multisectorial and multilevel object of Ofercat's object of analysis is innovative and is sufficiently flexible in its applicability to be applied to populations great and small, with very different socio-economic features, as the experimentation already carried out has shown. It enables us to analyse which are the stronger and weaker points of the supply of linguistic products and services in a given population and allows us to observe its evolution where Ofercat is used at regular intervals.

One of its most notable formal features is the simplicity of the presentation of the results thus obtained: an index (or score) and a graph for each population gives us the information at a glance, clearly and specifically; but it also makes it possible for us to obtain 100 indices and graphs that enable us to observe a single ambit (commerce, or private healthcare, for example), or a subsector (leisure) or the index for identification labelling for the entire population. That is, we obtain with the exploitation of the results, amenable working material for linguistic planning, light in form, but solid in its contents.

It is for that reason that Ofercat has the characteristics which make it suitable to be included in a stable programme of indicators which, looking to the future, will make it possible to obtain data in series as in

the case of census data and voting lists. Thus we will be able to look at the evolution of knowledge of a language over time, and monitor the trends in the supply of Catalan in our towns and cities.

It is also important to note that, if the project is obtaining samples that are sufficiently representative of Catalonia, we can obtain an overall picture of the supply of Catalan over the major part of the territory.

Accordingly, we could propose a study involving 25 towns and cities. The towns and cities chosen would need to comply with the following criteria:

1. The universe is confined, in principle, to settlements of 25,000 habitants or more.
2. We add 3 settlements more, two chosen from those of 10,000 inhabitants or more, of which 10% or more are foreigners, and a third, Tremp, which formed part of the first round of field work.
3. We distribute these populations in six blocks:
  - more than 100,000 inhabitants (A)
  - between 50,000 and 100,000<sup>2</sup>, with a low or medium low sociolinguistic index (B)
  - between 50,000 and 100,000, with a high or medium high sociolinguistic index (C)
  - between 25,000 and 50,000, with a low or medium low sociolinguistic score (D)
  - between 25,000 and 50,000, with a high or medium high sociolinguistic score (E)
  - the small towns described in point 2 (F)

To make the selection of towns and cities we bore in mind the demographic, territorial and sociolinguistic index factors, and the inclusion of the settlements that formed part of the original pilot study.

As a result of the application of these criteria the following selection emerged:

	POPULATION	INHABITANTS	% FOREIGN SETTLERS	Index
More than 100,000	1 Santa Coloma de Gramen	117.127	2,2	42,41
	2 L'Hospitalet de Llobregat	241.782	2,8	46,38
	3 Badalona	208.944	2,0	49,09
	4 Terrassa	171.794	2,6	55,68
	5 Sabadell	183.727	1,7	57,18
	6 Mataró	104.659	3,6	57,62
	7 Tarragona	114.097	2,1	59,92
	8 Barcelona	1.496.266	4,3	61,60
	9 Lleida	112.194	2,0	65,75
Between 50,000 and 100,000	10 Cornellà	80.998	2,4	43,05
	11 El Prat	62.956	2,6	46,75
	12 Sant Boi	79.337	2,3	47,96
	13 Granollers	52.423	3,1	62,90
	14 Reus	89.179	2,4	63,36
	15 Girona	73.637	3,3	69,82
	16 Manresa	63.742	2,2	70,05
	17 Ripollet	29.877	2,0	46,41
Between 25,000 and 50,000	18 Sant Joan Despl	27.627	2,3	49,42
	19 Castelldefels	45.091	8,6	50,16
	20 Tortosa	29.481	1,8	66,15
	21 Olot	27.967	3,0	74,99
	22 Vic	31.533	4,4	76,88
Other cases	23 Lloret de Mar	20.045	12,0	57,27
	24 Sitges	19.448	10,2	60,37
	25 Tremp	5.377	2,4	68,11

Source: Voting lists for the year 2000. Percentage of foreign settlers according to the 1999 voting lists.

<sup>2</sup> The sociolinguistic index which we are using is calculated by obtaining the overall average score for knowing how to speak, and how to write in accordance with the data from the 1996 linguistic census. to define the groups we added the highest and lowest score divided it by two, and used the resulting figure as the sociolinguistic index in question.

Ofercat is such that it can be easily adapted to other territories where Catalan is spoken, with sufficient supervision. Another possibility to be considered, would be to take its development one step further and make Ofercat a tool of analysis for the servicing of a language other than Catalan, in another part of the world where more than one language coexist. people of different nationalities have contacted the Ofercat Commission to find out more about this possibility. This is a new and exciting challenge.

**Mercè Romagosa** [mromagosa@cpnl.cat](mailto:mromagosa@cpnl.cat)

**Pilar López** [plopez@cpnl.cat](mailto:plopez@cpnl.cat)

**Albert Fabà** [afaba@cpnl.cat](mailto:afaba@cpnl.cat)

Consortium for Language Normalisation