# Uses and social representations of languages in the Balearic Islands

by Ernest Querol

## Abstract

This article presents the data obtained in October 2001, from research on a representative sample of fourth-year ESO (compulsory secondary education) pupils in the Balearic Islands. Firstly, we shall describe the use of Catalan before moving on to the explanation (the behaviour of the main variables that favour use), and we shall finish off with the predictions we can make using these data, given that the validity of the model has already been confirmed by comparing the data obtained in 1993 and 2000 in Catalonia, in 1998 in the Valencian Country and in Andorra at this very moment.

## Summary



## 1. Introduction

One of the main aims of social science research is to be able to make predictions. The model that we have designed allows us to arrive at this point. However, before doing so, we will present the **description** of language uses, the **explanation** of these uses and finally, the **predictions** for future uses that we discuss in this article.

## 2. Objectives

For our description, we need to find out the language uses of fourth-year ESO pupils. As regards the explanation, we want to find out the main variables that favour the use of Catalan, and as for predictions, we will reveal the prospects for use indicated by our theoretical and methodological model. Overall, therefore, our aim is to corroborate this new model again – as we did with Catalonia in 1993 and 2000, the Valencian Country in 1998 and, as we stated above, we are now testing in Andorra.[1]

[1] The complete results for the Community of Valencia can be found in Querol (2000*a*), and other, partial results can be consulted in Querol (2000*b* and 2000*c*). The most complete results for Catalonia can be read in Querol (2001*c*) and other, partial results can be consulted in Querol (2001*a* and 2001*b*). A preliminary comparison between Catalonia, the Balearic Islands and the Valencian Country was presented in Querol (2002).

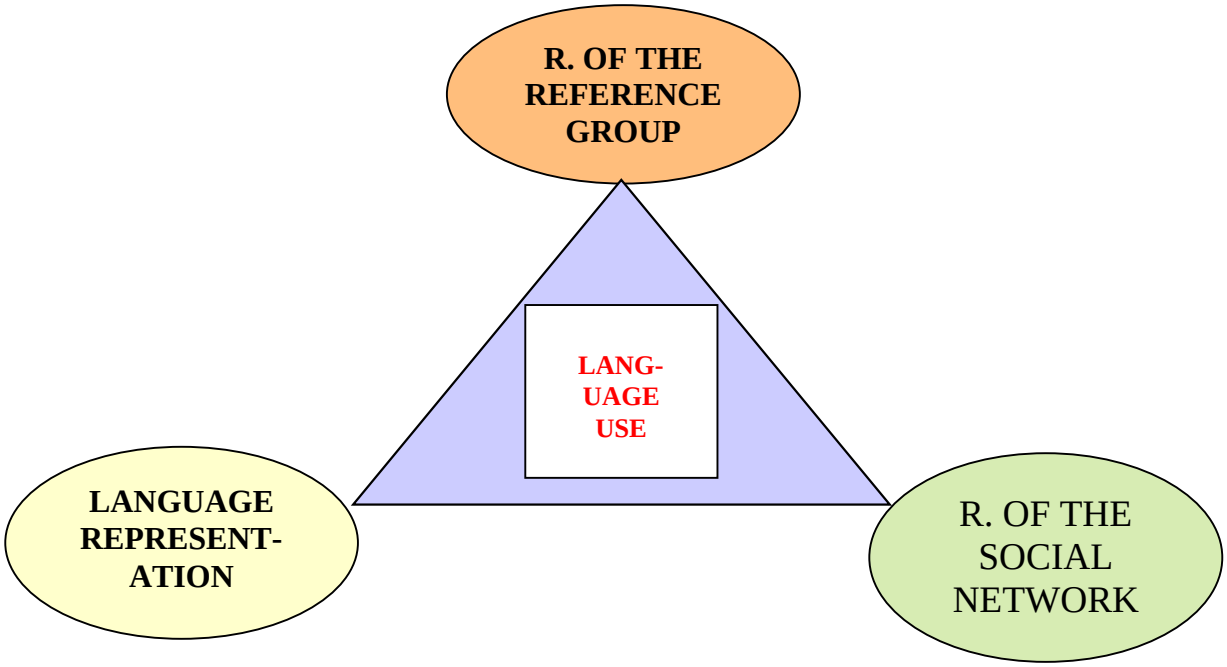
## 3. The new theoretical framework

A new theoretical framework was required to obtain this type of explanatory and predictive data, since the results obtained with the concept of *language attitude* were not statistically significant (Wicker 1969). We have presented this new proposal on a number of occasions, including in this journal (Querol 2001*a*).[2]

Figure 1 summarises our theoretical model for the study of language uses (dependent variable). The independent variables will be:
- social representations of languages,[3]
- social representations of reference groups[4] and,
- social representations of social networks.[5]

**Figure 1. Relations between the independent variables and the dependent variables (use of languages)**

Source: Own work.




All of the independent variables relate to each other and to the dependent variable, i.e. they feed from one another. The hypotheses we put forward initially focused on the relationships between all of these. We shall see these hypotheses now.

[2] For other approximations to the model, see Querol (1997), which deals with catastrophe theory, Querol (1999*b*), which also considers the reversal of processes of language shift and lastly, Querol (2001*b*), where the model is related to data obtained in 1993 and 2000 in Catalonia.
[3] Jodelet (1989: 36) offers this definition of *social representation*: "It is a form of knowledge with a practical vision, created and shared by society, which constructs a reality common to a social group".
[4] Merton and Kitt (1950) characterise this as follows: "It suggests systemising the determinants and consequences of processes of evaluation and self-evaluation, where the individual adopts the values or norms of other individuals and groups as a reference framework".
[5] The ***social network*** is the group of people who share one of their languages and have linguistic interaction with each other.

### 3.1 Hypotheses

#### *3.1.1 General hypothesis*

Types of language behaviour[6] are the result of the interaction of a set of social representations that the subject makes. Interaction between the following three is particularly important:

- representation of each of the languages at issue
- representation of the interpersonal communication network
- representation of the social reference group

#### *3.1.2 Subhypotheses*

***Representation of languages and the social network***

**First Subhypothesis:** There is a positive correlation between the representation of languages in contact and the individual social network: the more compact the social network in Catalan (i.e. the more possibilities of contact with Catalan-speakers),[7] the higher the representation in this language.

**Second Subhypothesis:** The representation of the languages (and social network) are correlated with the use of these languages in contact.

**Third Subhypothesis:** The representation of the Catalan language, as seen by Catalan-speakers, will be greater overall than that of the Catalan language as seen by Spanish-speakers.

***The social reference group***

**Fourth Subhypothesis:** Interaction between the social network and the language representations will be conditioned by the social reference group. In other words, this latter will mark the perception of linguistic groups. Hence, this will produce a desire to converge with or diverge from one's own group. Spanish-speakers' desire for distinction from their group results from the need to form part of the social network of Catalan-speakers.

## 4. Methodology

Our research tool was the survey, mainly because we wanted to compare the results obtained with the concept of language attitude and these latter were always quantitative.

### 4.1 Delimitation of the sample

#### *4.1.1 Basis of the sample*

The basis of the sample are fourth-year ESO pupils from public and private schools across the Balearic Islands.

---

[6] In the simplest of interlinguistic relations, we consider the general possibilities of language behaviour (contact between two languages, although in other research we have pointed out the convenience of always bearing in mind the chosen language for a more accurate analysis) to be:
- L1 monolingualism
- more L1 bilingualism than L2
- L1 / L2 bilingualism
- more L2 bilingualism than L1
- L2 monolingualism

Nonetheless, in the empirical part, we will not consider the possibility of 'equivalent use of both languages', since the different areas have a different social meaning. We do consider this possibility theoretically, as suggested by René Thom.

[7] The terms "Spanish-speaking" and "Catalan-speaking" refer to the use of each language by pupils, as they themselves state.

*4.1.1.1 Selection of the level*

The starting point for our empirical research already focused the field of analysis on secondary education, which was the scope of our work as teachers, with the aim of possibly drawing up different learning proposals for other research. We decided that the population to analyse would be fourth-year ESO pupils. The reasons for this were numerous. The most influential factor, however, was that this is the point at which compulsory secondary education is completed and, therefore, all pupils reach this level. This compulsory point also adds certain special features because it represents a type of crossroads for pupils, since their decisions on which options to take as regards education must begin here: this could be continuing their education with the Baccalaureate, studying on professional training schemes, or entering the world of work. In this last case, there may also be the option of taking further studies courses. Otherwise, at the very least, it could be used to extract sociolinguistic data for analysing future tendencies and uses of the different languages. The complexity of the questionnaire was also a reason for choosing the final year, since a more mature age group was preferred.

*4.1.1.2 Sample Confidence Interval*

The confidence interval we chose for the sample was of two sigmas, i.e. 95.5%; which is habitually used for working with. It indicates that the values of the universe must fall within this quantitative space formed by the average, the percentage obtained ±2 (plus or minus 2%). This **sampling error** is merely the maximum statistical error of the sample. It is also generic, because it is valid for the set of all the diverse samples of the same size that could be taken from the same population.

*4.1.1.3 Sample Unit*

As we said earlier, this is made up of fourth-year ESO pupils from public and private schools across the Balearic Islands. On the basis of the population of pupils across the Balearic Islands in the fourth year of ESO during the 2001-2002 academic year, we established what the population had to be to ensure a representative sample.

*4.1.1.4 Sample Design*

The population chosen for study were, therefore, fourth-year ESO pupils from schools across the Balearic Islands. To ensure greater representativeness of the sample population, we defined diverse strata to control, in accordance with a set of variables considered to be essential for our research. These were:

1. Territorial division: the Islands.
2. Type of population agglomeration:
   - towns with over 30,000 inhabitants in Mallorca and Ibiza, and 20,000 in Menorca.
   - towns with less than 30,000 inhabitants in Mallorca and Ibiza, and 20,000 in Menorca.
3. Type of ownership of the school: public / private.
4. Type of secondary school in terms of the language used:
   - more Catalan teaching;
   - more Spanish teaching.

As “class group” was considered to be a unit of study, we used a mixed strategy when designing the sample. There were two discernible stages:

Stage One. Using data provided by the Council for Education and Culture of the Government of the Balearic Islands, the total population of fourth-year ESO pupils in schools across the Balearic Islands was calculated at 3,897 pupils in private centres and 4,939 in public schools (2000-2001 academic year).

On the basis of these criteria, we calculated that the minimum sample size had to be 384 pupils surveyed. Once we had calculated the minimum number of pupils to be surveyed, hypothesising that each class group contained 22, we calculated that the total number of class groups that had to take the survey (conglomerates, “cluster sampling”) was 18 sample groups; these would be divided according to geographical situation, as shown in the following table.

We observed that 1.047% of the population were in the fourth year of ESO, and that the percentage of pupils surveyed was 4.47%.

Stage Two. At this stage, the decisions were made as to the distribution of these 18 groups on the basis of the strata. To do so, information on the variables to control was obtained and analysed.

Table 1 illustrates the sample obtained from applying the criteria discussed above, and how many surveys were obtained in each school.

**Table 1. Sample obtained using all strata**

| PRIVATE SCHOOLS | Island and town/city | Agglome-ration type | Ownership of centre | Linguistic orien-tation | Social strata | Number of surveys | % of total |
|---|---|---|---|---|---|---|---|
| Lluís Vives | Mallorca Palma | Palma Centre | Private, not grant-aided | Catalan | High | 26 | 5.8% |
| Pius XII | Mallorca Palma | Palma Centre | Private, grant-aided | Spanish | Middle | 28 | 6.3% |
| La Salle | Mallorca Palma | Palma Outskirts | Private, grant-aided | Spanish | High-middle | 31 | 7% |
| Sant Vicenç de Paül (2) | Mallorca Palma | Palma Outskirts | Private, grant-aided | Spanish | Lower-middle | 25 | 5.6% |
| St. Francesc d'Asís | Mallorca Manacor | Urban Inland | Private, grant-aided | Spanish | Middle | 24 | 5.4% |
| Nostra Senyora de la Conso-lació | Mallorca Alcúdia | Rural Coastal | Private, grant-aided | Spanish | Lower-middle | 28 | 6.3% |
| Cor de Maria | Menorca Maó | Urban | Private, grant-aided | Spanish | High-middle | 24 | 5.4% |
| Virgen de las Nieves | Ibiza Sant Josep | Rural | Private, grant-aided | Spanish | High-middle | 17 | 3.8% |

**Table 1. Sample obtained using all strata (cont.)**

| PUBLIC SCHOOLS | Island and town/city | Agglome-ration type | Owner-ship of centre | Linguistic orientation | Social strata | Total surveys | % of sample |
|---|---|---|---|---|---|---|---|
| Biel Martí | Menorca Ferreries | Rural | Public | Catalan | Middle | 15 | 3.4% |
| Sa Colomina | Ibiza Ibiza | Urban | Public | Spanish | Middle | 28 | 6.3% |
| Joan Alcover | Mallorca Palma | Palma Centre | Public | Equal | High | 21 | 4.7% |
| Francesc de Borja | Mallorca Palma | Palma Outskirts | Public | Spanish | Low | 20 | 4.5% |
| Ramon Llull | Mallorca Palma | Palma Centre | Public | Spanish | High-middle | 28 | 6.3% |
| Arxiduc Lluís | Mallorca Palma | Palma Centre | Public | Spanish | Middle | 22 | 4.9% |
| Josep Sureda i Blanes | Mallorca Palma | Palma Outskirts | Public | Equal | Lower-middle | 23 | 5.2% |
| Capdepera | Mallorca Capdepera | Rural Coastal | Public | Equal | Middle | 35 | 7.9% |
| Guillem Colom | Mallorca Sóller | Rural Inland | Public | Equal | Middle | 25 | 5.6% |
| Llucmajor | Mallorca Llucmajor | Urban | Public | Catalan | Middle | 25 | 5.6% |

Statistics were processed using a package commonly used in social science research: the Statistical Package of Social Science (SPSS), Version 8.0 (registered software). We applied the analyses of frequencies and percentages of each variable (univariable analysis), the intersection of two variables (bivariable analysis) and the multivariable analysis (discriminating and multiple regression analyses) to the fieldwork results. We also used a new method of graph-based induction techniques called SIPINA: Interactive System for the Processes of Non-Arborescent Interrogation No-Arborescent, which will see in Section 5.2.2.

We have divided our account of the results according to their purpose: descriptive, explanatory and predictive.

## 5. Results

### 5.1. Descriptive Results

The reliability coefficient ranges from .9282 for the scale on identity to .9808 for the scale on social representations of languages.

As regards **languages learned**, the first language of pupils is Spanish in 56.7% of cases and Catalan in 39.5%. However, we were unable to classify 0.9% of pupils because they did not answer the survey correctly. The second language learned is Catalan in 50.9% of cases and Spanish in 40.8%. A total of 8.3% pupils learned a different language. The first language of 37.5% of fathers was Catalan – two percentage points lower than for pupils. Spanish was the first language of 55.5% of fathers whereas 7% had learnt another language. The Catalan-learning percentages for pupils' mothers were higher than those of the fathers: 38.8%. However, the majority – 52.6% – learnt Spanish as their first language, and 8.6% of mothers learnt a different language. We also observe that only 38.3% of fathers and 39.2% of mothers learn Catalan as a second language. Finally, 12.2% of fathers and 17.9% of mothers do not learn any language other than their first.

**Intergenerational Language Transmission**

General data on the language spoken by parents of the pupils that we surveyed reveal that, when both parents were Catalan-speaking, Spanish was only transmitted in three cases, representing 2.4%, and 97.6% transmitted Catalan. When both were Spanish-speaking, 4.2% spoke in Catalan with their children. Therefore, 95.2% transmitted Spanish. In 0.5% (one case) a different language was transmitted.

**Table 2. First language learnt by the pupil-parents' language**

| | | | Language of parents | | | | Total |
|---|---|---|---|---|---|---|---|
| | | | Catalan | Catalan-Spanish | Spanish-Catalan | Spanish | |
| First language learnt | Catalan | Breakd | 120 | 16 | 21 | 8 | 165 |
| | | % of first language learnt | 72,7% | 9,7% | 12,7% | 4,8% | 100,0% |
| | | % of language of parents | 97,6% | 45,7% | 47,7% | 4,2% | 42,2% |
| | | % of total | 30,7% | 4,1% | 5,4% | 2,0% | 42,2% |
| | Spanish | Breakd | 3 | 19 | 23 | 180 | 225 |
| | | % of first language learnt | 1,3% | 8,4% | 10,2% | 80,0% | 100,0% |
| | | % of language of parents | 2,4% | 54,3% | 52,3% | 95,2% | 57,5% |
| | | % of total | ,8% | 4,9% | 5,9% | 46,0% | 57,5% |
| | Other | Breakd | | | | 1 | 1 |
| | | % of first language learnt | | | | 100,0% | 100,0% |
| | | % of language of parents | | | | ,5% | ,3% |
| | | % of total | | | | ,3% | ,3% |
| Total | | Breakd | 123 | 35 | 44 | 189 | 391 |
| | | % of first language learnt | 31,5% | 9,0% | 11,3% | 48,3% | 100,0% |
| | | % of language of parents | 100,0% | 100,0% | 100,0% | 100,0% | 100,0% |
| | | % of total | 31,5% | 9,0% | 11,3% | 48,3% | 100,0% |

The transmission of languages is fairly similar in **linguistically-mixed marriages**: Catalan-speaking fathers and Spanish-speaking mothers transmitted Catalan in 45.7% of cases, and Spanish in 54.3%. When the father was Spanish-speaking and the mother was Catalan-speaking, 47.7% transmitted Catalan, whereas 52.3% transmitted Spanish. Thus, **the tendency is to transmit Spanish more**.

The overall data on intergenerational language transmission, as revealed in Table 2, are: 57.5% transmit Spanish, whereas Catalan is transmitted in 42.2% of cases. Thus, we can conclude that **Spanish is the language that tends to be transmitted more**, regardless of whether both parents use the same language or whether they speak different languages.

Pupils chose to **answer the survey** in Spanish in 62.4% of cases and in Catalan in 31.5%; the remaining 6% chose to respond in both languages (one language in each part).

## 5.2 Explanatory Results

### *5.2.1 Correlations, Discriminatory analysis and Multiple Regression Analysis*

In the Objectives section, we saw that the aim of the research was to re-define the main variables that favour the use of languages. Using the proposed theoretical model, the following results were obtained:

**Table 3. Summary of Explanatory Data**

| BALEARIC ISLANDS 2001 | |
|---|---|
| **Sample**<br><br>**Number of surveys** | Representative of the fourth year of ESO across the Balearic Islands<br>447 |
| **Reliability of the survey scales** | Up to 0.9808 |
| **Main correlations with the use of Catalan** | - representation of Spanish -.826<br>- representation of Catalan .817<br>- social network in Catalan .801<br>- social network in Spanish -.757 |
| **Main variables of the function of discriminatory analysis as regards use:**<br><br>**Cases well-classified using discriminatory analysis:** | - representation of Spanish<br>- representation of Catalan<br>- social network in Catalan<br>- social network in Spanish<br><br>60% |
| **Main variables that predict the use of Catalan according to multiple regression analysis:** | - representation of Spanish<br>- social network in Catalan<br>- social network in Spanish<br>- reference group |

We shall now comment on the main correlations between the variables proposed by our model and the use of Catalan:

- The **representation** of Spanish (the image that speakers have of this language) is the main factor that correlates with the use of Catalan (-.826).
- The representation of Catalan co-varies .817 with use of this language.
- Social **network** in Catalan (i.e. the group of people with whom the speaker normally relates) is .801 related to use of Catalan.
- Social network in Spanish is much less associated with use of Catalan: .-757.

- The results for the **variables** influencing the use of Catalan are quite similar, regardless of the year and the analysis used. This is illustrated in the following table:

**Table 4. Main Predictive Variables by Analysis Type**

| PEARSON'S CORRELATIONS | REGRESSION ANALYSIS | DISCRIMINATORY ANALYSIS |
|---|---|---|
| - Representation of Spanish | - Representation of Spanish | - Representation of Spanish |
| - representation of Catalan | - Social network in Catalan | - representation of Catalan |
| - Social network in Catalan | - Social network in Spanish | - Social network in Catalan |
| - Social network in Spanish | - Reference group | - Social network in Spanish |

We therefore see that the 'representation of Spanish' variable always takes first place, followed by representation of Catalan, even though it does not appear in regression analysis. Social network in Catalan takes second place in regression analysis and third in the other two. Social network in the other language, Spanish, also appears in all three techniques: it takes third place in regression analysis and comes fourth in the other two. Finally, reference group only appears in regression analysis; this variable also formed part of our model, the outline of which is illustrated in Figure 1.

*5.2.2 Graph-induction techniques (SIPINA)*

In addition to analyses of data we have termed "explanatory", such as those we have just seen (correlations, discriminatory and multiple regression analyses), we have used another technique, proposed by Rakotomalala (1997 and 2000). This will be very useful because it will enable us to explain a qualitative variable using explanatory variables, which are themselves discretized, qualitative variables. Here, we are referring to knowledge through graph-induction. This method involves constructing a predictive function in the form of a graph and a decision tree; it allows us to explain or predict the value taken by a particular "endogenous" variable, from a series of "exogenous" variables. We discussed the features of this technique elsewhere in this journal (Querol, 2001*a*).

We will now characterise the four linguistic groups that we have defined:

-Exclusive speakers of Catalan (ECAT);
-More Catalan-speaking (+CAT);
-More Spanish-speaking (+CAS);
-Exclusive speakers of Spanish (ECAS);

using some of the variables that we have introduced into the research:

Represent = Representation
Transmission L. = Language transmitted by the parents
Ident. = Identity
Ins. Mot. cat. = Instrumental motivations regarding Catalan
Int. Mot. = integrative motivations
Net = Social network
Fear = Fear of being assimilated
Bil / U = Bilingual / Unilingual
Cat. , Ct. = Catalan
Span., Sp. = Spanish
Eur = European

The procedure is actually similar to that of discriminatory analysis. Figure 2 illustrates the overall results using this new technique.

The following figure reveals the segments that will allow us to identify each group better. The first is representation of Spanish. The next two factors dividing the two preceding groups are: transmission of the language, i.e. whether or not the parents spoke to the pupil in

Catalan, and the representation of Catalan. Once more, we see that the variable differentiating groups the most is representation, which bears out our hypotheses.

The third factors dividing these subgroups are: Spanish identity, European identity and, on two other occasions, language transmission. At the fourth level, the following come into play: instrumental motivations in Catalan; social network in Spanish; social network in Catalan; representation of Spanish, and the fear of being assimilated by Spanish-speakers. At the fifth level, we find the instrumental motivations in Spanish and Spanish identity. If we narrow down further the Catalan-speaking group, we see that the following are still involved: identity with language (chosen from a scale of which the two extremes were: Catalan language/Spanish language), bilingual / unilingual identity. Finally, we see the fear of being assimilated by Spanish-speakers and integrative motivations in Spanish.

Thus, we see that, of the variables of our hypotheses (excluding reference group, which was not a numerical variable), both representation of languages and social network are always found at the fourth level or lower. That is, we have discovered a new way of confirming our proposal by the use of induction graphs. This new knowledge technique enables us to outline the features of each group. We shall see this in Table 5, which reveals the segments that best allow us to identify each group. We will comment on the segments containing a sufficient number of pupils, allowing us to best isolate the group we want to describe (these were marked in Figure 2). The number of pupils analysed is 357, the percentages shown relate to each of the groups and all scales used are from 1 to 10.

**Table 5. Description of the four linguistic groups using induction graphs**

| | |
|---|---|
| **Exclusively Catalan-speaking** | Pupils with:<br>- A representation of Spanish of less than 5.95<br>- A Catalan- speaking parent<br>- A Spanish identity of less than 2.5<br>- Instrumental motivations in Catalan of over 6.25 |
| **More Catalan-speaking** | Pupils with:<br>- A representation of Spanish of less than 5.95<br>- A Catalan-speaking parent<br>- A Spanish identity of over 2.5<br>- A Spanish network of less than 6.9<br>- Instrumental motivations of Spanish of over 6.75 |
| **More Spanish-speaking** | Pupils with:<br>- A representation of Spanish of over 5.95<br>- A representation of Catalan of over 4.35<br>- Both parents who are Spanish-speaking, or with a Spanish- speaking father and a Catalan-speaking mother |
| **Exclusively Spanish-speaking** | Pupils with:<br>- A representation of Spanish of over 5.95<br>- A representation of Catalan of less than 4.35<br>- Both parents who are Spanish-speaking, or with a Spanish-speaking father and a Catalan-speaking mother<br>- A fear of being assimilated by Spanish-speakers of less than 1.5 |

If we compare these results with our hypotheses, we see that even the order that we put forward coincides with the induction graph results. Let us take a closer look at this: apart from the language spoken by parents to the pupils, which we did not take into consideration because of its obvious character, the 'representation' variable is the most helpful in defining the groups. Therefore, the 'exclusively Catalan-speaking' and the 'more Catalan-speaking' groups are determined by the low representation of Spanish. Representation is also the factor that best distinguishes the two groups that speak more Spanish from the Catalan-speaking groups. The second variable for Catalan-speakers is having a Catalan-speaking parent whereas, for the two Spanish-speaking groups, it is the level of Catalan

representation that differentiates them; parents are the third variable differentiating them from the Catalan-speaking block.

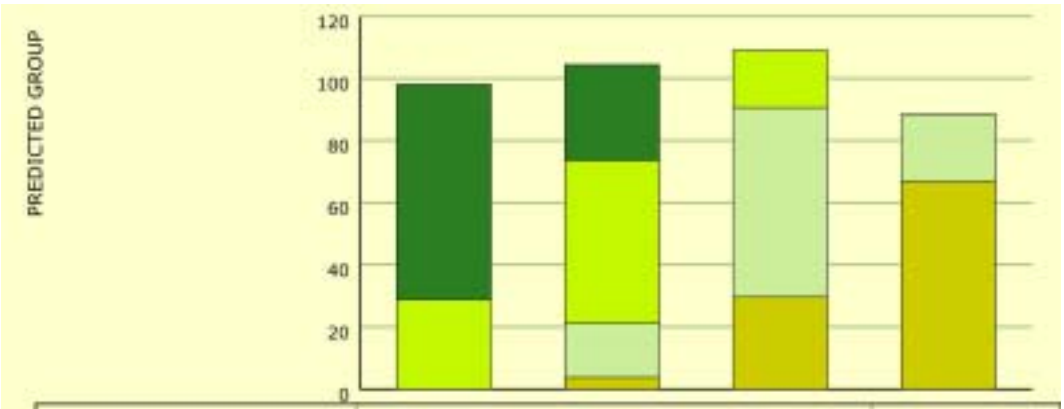
As we have seen then, the importance of the 'representation' variable is clear, and it was this variable that featured in all of our subhypotheses. The other variable that appeared in our first subhypothesis was 'social network': in the description of linguistic groups using induction graphs, this is the fourth variable that delimits the group of more Spanish-speaking pupils.
The variables that we did not want to form part of the hypotheses, although they were included in the questionnaire, were those referring to the concept of "identity"; in this analysis, we see that this factor does differentiate the two more Catalan-speaking groups – at the third level. Finally, the other variable that delimits the exclusively Spanish-speaking group at the third level is the fear of being assimilated.

## 5.3 Predictive

### *5.3.1 Predictions of Belonging to a Linguistic Group*

The model that we have used allows us to predict which group a pupil belongs to without the information that has led us to make this classification. Discriminatory analysis can be used to classify and assign pupils to a linguistic group. The level of correct results will give us the measure of interaction between the variables involved in the functions. Without eliminating the group in which pupils categorised themselves, i.e. the group to which pupils belong, the percentage of pupils correctly classified by our model is 70.1%. However, we believe that this variable was too close to the classification that we are trying to establish and have therefore eliminated it from this analysis. Thus, the data revealed below do not take into consideration the self-categorised group.

**Figure 3. Correct results of the models**




| | Exclusively Catalan-speaking | More Catalan-speaking | More Spanish-speaking | Exclusively Spanish-speaking |
|---|---|---|---|---|
| Exclusively Catalan-speaking | 69,20% | 30,80% | 0% | 0% |
| More Catalan-speaking | 28,80% | 52,50% | 18,60% | 0% |
| More Spanish-speaking | 0% | 17,40% | 60,90% | 21,70% |
| Exclusively Spanish-speaking | 0% | 3,70% | 29,60% | 66,70% |

REAL GROUP (whole column)

When the representation of Spanish and the Catalan network are taken into consideration, our theoretical proposal is able to correctly group 60.1% of classified cases. If we take a detailed look at the group, we see that, in the first group (exclusively Catalan-speaking), our theoretical variables proposal predicted 69.2% of cases correctly, in the second (more Catalan-speaking), it obtained 52.5% correct results, and in the third (more Spanish-speaking), it correctly grouped 60.9%. Lastly, for the fourth group (only Spanish-speaking), it obtained 66.7% correct results.

This shows that the most successful predictions were obtained in the two groups at either extreme, whereas the least successful were obtained in the middle groups. Analysis of the correct results of the groups at either extreme reveals that not a single error is made in either case by classifying pupils in the two groups that are not immediately adjacent. Therefore, our proposal demarcates the groups well and only makes slight errors.

Another, similar parallelism is observed in the two central groups: in no case is the wrong group at the extreme of the other language assigned, i.e. if we are classifying pupils who speak more Catalan, they are never mistakenly grouped with those who only speak Spanish. Or the other way round: if we are classifying pupils who speak more Spanish, they are never assigned the group that only speaks Catalan.

The best-classified group is that of pupils who only speak Catalan, which is assigned correctly in 69.2% of cases. The percentage of grouping errors for those who speak more Catalan is 30.8%. The second best-classified group is that of pupils who only speak Spanish (66.7% correct results). In this case, the number of pupils classified in the adjacent group drops: 29.6% of errors. Here, one pupil is assigned (3.7%) to the group of more Catalan-speaking pupils.

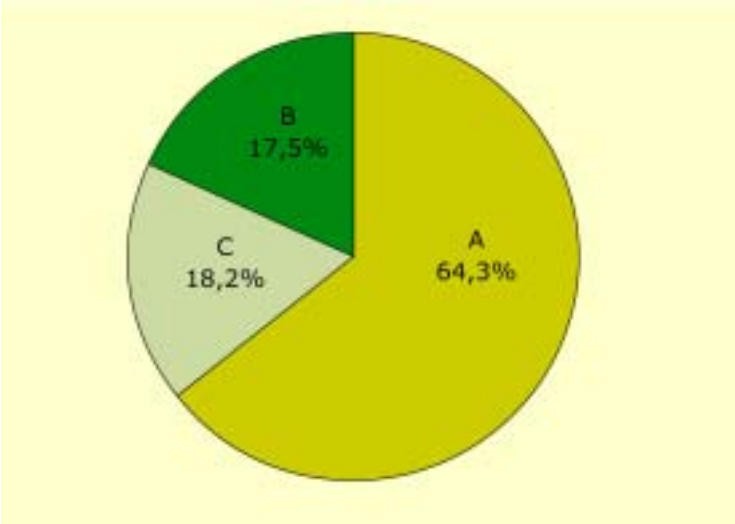
We can therefore conclude that the percentage of correct results is very high and that the errors almost always involve adjacent groups, except for one case which 'jumps' a group.

### *5.2.3 Predictions of Evolutions of Linguistic Groups*

On the basis of the data that we have obtained, our model can predict the evolution of linguistic groups by comparing the linguistic group to which speakers belong and that of reference. We have already confirmed that this prediction is successful with data obtained in Catalonia in 1993 (when we made proposals for future evolution) and 2000. Our present prediction can be seen in Figure 4.

**Figure 4. Relation between the group to which speakers belong and that of reference**

A: Coincidence between the group to which speakers belong and the desired group;
B: The desired group is more "Spanish" than that to which speakers belong;
C: The desired group is more "Catalan" than that to which speakers belong;




- Research into the increase in the Catalan-speaking group.

This research into the increase in the Catalan-speaking group by relating the group to which the subject belongs and the desired group (of reference) reveals that the percentage of speakers that could vary is 0.7%, if we subtract those that will switch to Spanish and those that will switch to Catalan.

## 6. Conclusions

Our research on the situation of the use of Catalan in the Balearic Islands and the variables influencing this use have, once again, confirmed conclusions that we believe can be extended to other contexts and languages (having already been confirmed in Catalonia and the Community of Valencia). Firstly, that the use of language is linked to different factors: as we have seen, the variable that correlates most with use in the Balearic Islands is the representation of the Spanish language, whereas, in Catalonia, it was the representation of the Catalan language and, in the Valencian Country, it was the social network of Catalan. Table 6 summarises the results of these three investigations using the same structure as we saw in Table 3.[8]

[8] For reasons of space, this table only includes the results that we obtained in research carried out in Catalonia in 1993. For the data from 2000, see our comparative article elsewhere in this journal (Querol 2001*a*).

**Table 6. Comparison of the main results from Catalonia, Valencian Country and the Balearic Islands**

| CATALONIA 1993 | VALENCIAN COUNTRY 1998 | BALEARIC ISLANDS 2001 |
|---|---|---|
| Representative of second year BUP across Catalonia<br>432 | Representative of fourth year ESO across the Valencian Country<br>452 | Representative of fourth year ESO across the Balearic Islands<br>447 |
| 0,9751 (121), 0,9770 (112), 0,9613 (84), 0,9653 (82) | 0,9762 (123), 0,9495 (15), 0,9642 (44) , 0,9352 (49) | Up to 0.9808 |
| - representation of Spanish .-846<br>- representation of Catalan .824<br>- social network in Catalan .783<br>- social network in Spanish .-536 | - social network in Valencian .796<br>- representation of Spanish .-768<br>- representation of Valencian .713<br>- social network in Spanish .710 | - representation of Spanish -.826<br>- representation of Catalan .817<br>- social network in Catalan .801<br>- social network in Spanish -.757 |
| - representation of Spanish<br>- social network in Catalan<br>- social network in Spanish<br>- representation of Catalan<br>77.4% | - social network in Valencian<br>- social network in Spanish<br>- group to which speakers belong<br>64.4% | **- representation of Spanish**<br>- representation of Catalan<br>- social network in Catalan<br>- social network in Spanish<br>60% |
| - representation of Spanish<br>- social network in Catalan<br>- social network in Spanish<br>- representation of Catalan<br>- reference group | - social network in Valencian<br>- social network in Spanish<br>- representation of Spanish | **- representation of Spanish**<br>**- social network in Catalan**<br>- social network in Spanish<br>- reference group |

We may wonder what these results mean:

1. If use of the language is linked to the social network and this latter is progressively reduced in generational transmission, the number of speakers will also be reduced.

2. If use of the language correlates mainly with the representation of this latter, and it is very much linked to the social network, as the representation increases, so too will the number of new speakers. In other words, as in the case of the Balearic Islands, if the representation of Spanish (which is correlative to that of Catalan), drops, the use of Catalan will increase.

The consequences of these results are clear: to reverse a process of language shift, we must ensure:

1. That the variable that best correlates with use is social representation;
2. That generational transmission is not negative for the language in decline or in danger of decline.

We will not discuss here the measures that need to be taken to promote a change in the current trend. To summarise, we will simply affirm that the validation of our theoretical and methodological model in the context of the Balearic Islands provides us with a very productive tool for the study of processes of language shift and for the design of strategies for reversing these latter. Moreover, these advances in sociology of language could also be very useful in other areas of social science.

All in all, these fruits guarantee the model that we proposed. That is to say, we can provide a very good explanation of the linguistic choices of speakers. As we know, there is always room for improvement, but we are in no doubt about the fact that we have made an

important step in discovering the reasons for the use of Catalan in the Balearic Islands, which was our main aim.

## 7. Bibliography


MERTON, R. K. and KITT, A. S. (1950): "Contributions to the theory of reference group behavior", in Merton, R. K. I Lazarsfeld, P. F. *Continuities in social research: studies in the scope and method of the American soldier*, Glencoe, Ill, Free Press.

JODELET, D. (1989): "Représentations sociales: un domaine en expansion", in Jodelet, D.: *Les représentations sociales*. Paris: Presses Universitaires de France.

QUEROL, E. (2002*a*): "Els valencians i el valencià: comparació amb els catalans i amb els balears", framework conference for the Tercer Congrés de l'Escola Valenciana, València [at press].

QUEROL, E. (2002b): *Els catalans i el català. Usos i representacions socials.* Barcelona: Ed. Empúries [at press].

QUEROL, E. (2001*a*): "Evolució dels usos i de les representacions socials de les llengües a Catalunya (1993-2000)" *Noves SL*, Summer 2001.
http://www.gencat.cat/llengua/noves

QUEROL, E. (2001*b*): "A new model for the Evaluation of Language Planning. A case Study: Catalonia (1993-2000)" *Sociolinguistica* [at press].

QUEROL, E. (2001*c*): "Un exemple d'inversió del procés de substitució lingüística: Catalunya 1993-2000", II Congrés Europeu sobre Planificació Lingüística [at press].

QUEROL, E. (2000*a*): *Els valencians i el valencià. Usos i representacions socials.* València: Ed. Denes (Premi d'Investigació Francesc Ferrer Pastor, 7).

QUEROL, E. (2000*b*): "Variables que influeixen l'ús de les llengües al País Valencià", speech presented at the I Congrés Internacional Llengua, societat i ensenyament [at press].

QUEROL, E. (2000*c*): "Resultats d'un nou model per determinar les variables que influeixen sobre l'ús de les llengües (País Valencià i Catalunya)". Congrés de Catalanística of Münster "Catalunya –un cas excepcional a Europa?".Section: Multilingüisme, llengües en contacte i conflicte lingüístic. Münster [at press].

QUEROL, E. (1999*a*): *Cap a un nou marc teòric per a l'estudi de les variables dels processos de substitució lingüística.* Barcelona: Publicacions Universitat de Barcelona (micro-archived doctoral theses, 3568).

QUEROL, E. (1999*b*): "A new Theoretical Approach to the Study of Reversing Language Shift: Catastrophe Theory". VII *International Conference on Minority Languages*, Bilbao [at press].

QUEROL, E. (1997): "Un nou model per a l'estudi del comportament lingüístic: la teoria de les catàstrofes". *Caplletra*. [València], Nº. 21, Autumn, pp. 161-184.

QUEROL, E. (1995): "Les representacions socials de les relacions interlingüístiques i l'ensenyament de les llengües". In Iʳ Congrés de l'Escola Valenciana. València: Federació Escola Valenciana, 1995, pp. 347-361.

QUEROL, E. (1994) "Els processos de substitució lingüística: cap a una metodologia d'anàlisi". In Comunicacions i Ponències. Simposi del Professorat de Valencià de l'Ensenyament Mitjà i de l'Educació Secundària Obligatòria. València: Generalitat Valenciana, pp. 23-33.

QUEROL, E. (1993): “Models en contacte: anàlisi dels estudis sobre les relacions interlingüístiques”. In Actes del IX Col·loqui Internacional de Llengua i Literatura Catalanes. Alacant-Elx, pp. 87-96.

QUEROL, E. (1990): “El procés de substitució lingüística: la comarca dels Ports com a exemple”. In *Miscel·lània 89,* València: Generalitat Valenciana, pp.85-196.

RAKOTOMALALA, R. (1997): *Graphes d'induction*. Doctoral thesis presented at the Claude Bernard - Lyon I University, 13th December.

RAKOTOMALALA, R. and ZIGHED, D. (2000): *Les graphes d'induction*. Paris: Ed. Hermes.

**Ernest Querol**
equerolp@uoc.edu
UOC-Open University of Catalonia