

# From monolingual family background to bilingual identification: the case of pre-adolescents in Mataró and the Aragonese border area<sup>1</sup>

Vanessa Bretxa, Llorenç Comajoan i Natxo Sorolla

## Abstract

This article presents data on the relationship between the language used by the pre-adolescent population of Mataró and the Aragonese border area with their parents and their language of identity. In particular, it looks into the way in which different socio-linguistic variables affect this relationship so as to create two linguistic sub-groups: one group which has a mono-lingual Spanish-speaking family background, and identifies with this language, and the other which, despite also having a monolingual family background, exhibits bilingual identification. Various socio-linguistic variables (oral competence, social networking, attitudes, confidence and the language used with siblings) are analysed, with a view to predicting membership of one group or the other. The results show that social networks among peers are the variable giving the best indication of individuals' language of identity. és la variable que millor permet discriminar la llengua d'identificació dels individus.

## Summary

1. Introduction
2. Language of identity and language used with parents
  - 2.1. Language and identity
  - 2.2. Language and identity in Catalonia
3. Methodology of the study
4. Results
  - 4.1. Language used with parents and language of identity
  - 4.2. Descriptive analysis
    - 4.2.1. Oral competence in Catalan
    - 4.2.2. Social networks
    - 4.2.3. Attitudes to Catalan
    - 4.2.4. Linguistic confidence in Catalan
    - 4.2.5. Language used with siblings
  - 4.3. Multivariable analysis: discriminant analysis
5. Discussion and conclusions

## 1. Introduction

This paper studies the relationship between the language spoken by pre-adolescents in the family setting (language spoken with parents) and the language they identify as their own (language of identity). In particular it presents data regarding the pupils in the sixth year of primary education (aged 11 to 12) at schools in Mataró and the Aragonese border area and investigates the way in which different socio-linguistic variables can help to explain the language with which these young respondents identify. The study takes the following observation as its starting point: there are a number of boys and girls who speak to their parents in Spanish (the family background is monolingual) but who do not necessarily identify exclusively with this language (bilingual identification). The article therefore examines variables which might explain this change. In particular, data is provided on oral competence in Catalan, the languages used in the respondents' social networks, attitudes to language, the language used with brothers and sisters and language confidence. The descriptive data for each of these variables provides an overall view in which certain differences can be seen between the socio-linguistic backgrounds of Mataró and the Aragonese border area (for example, the level of competence in Catalan is higher in Mataró than in the Aragonese border area). On the other hand, there are certain similarities in the two areas (for example, the widespread use of Spanish in the pupils' social networks). By

<sup>1</sup> This study has been carried out as part of the project *Resocialisation and languages: the linguistic effects of the transition from primary to secondary education (RESOL)*. The project is financed by the Ministry of Education and Science (HUM 2006-05860/FILO). Head of research: Dr. F. Xavier Vila i Moreno.

means of the statistical technique of discriminant analysis, we determine which of the variables studied (oral competence, social networks, attitudes, confidence and the language used with siblings) allow the best prediction of the language of identity. The results of the study show that the social networks of sixth year primary pupils are the variable which allows us to determine the language of identity most reliably.

## 2. Language of identity and language used with parents

Explaining the way in which multilingual individuals come to use their different languages, and the different levels of competence they have in them, is a complex task, because a multitude of factors intervene in the explanation of language use in plurilingual contexts. Different theoretical models have attempted to make these factors explicit, particularly for those cases in which one of the language varieties in the socio-linguistic environment is weaker. For example, in the classic framework of the reversibility of language change, Fishman (1991, 2001) emphasised the role of the family and language transmission in the family as major factors in saving a threatened language (see Boix and Vila, 1998; Comajoan, 2005). Other theoretical frameworks – more psychological than social – centre on situational and personal factors (for example the Willingness to Communicate concept of MacIntyre et al. 1998). In this paper we study some of the variables of both theoretical frameworks when explaining certain discrepancies in the language used by the family and the language of identity among pre-adolescents in Mataró and the Aragonese border area in 2007.

### 2.1. Language and identity

The language an individual speaks and his/her identity are inseparable (Tabouret-Keller, 1997, p. 315).<sup>2</sup> This inseparability makes the study of the relationship between language and identity particularly complex. On the one hand, as Tabouret-Keller says, identities are not homogeneous since "At any given time a person's identity is a heterogeneous set made up of all the names or identities, given to and taken up by her" (p. 316). On the other hand, because of the heterogeneous nature of identities, it is also possible to (re)negotiate them (Pavlenko & Blackledge, 2004).

Tabouret-Keller gives two reasons to explain the close relationship between language and identity. The first, which is psychological, is motivated by the individual's identification with the use of language through their relationship with their parents, or by one individual's imitation of another's behaviour. The second is related to the explicit and implicit legal factors governing the use and knowledge of languages in specific contexts.<sup>3</sup>

In this paper the psychological relationship is studied from data on the language with which individuals identify and the language used in the family context (language used with parents). Obtaining data on these two variables by means of a written socio-linguistic questionnaire, as was done for this study, is necessarily a problem, as it does not allow us to observe the way in which individuals express the dynamic aspect of their identity. However, the fact that various combinations of languages were available for the possible answers means that this dynamic aspect was allowed for to some extent (see section on methodology).

This study considers two other variables which are psychological in nature and closely related to the psychological dimension of the language of identity: attitudes to language and language confidence. Attitudes are "a disposition to react favourably or unfavourably to a class of objects" (Garrett et al., 2003, p.3); in the case with which we are concerned the "objects" are languages. As Garret et al. point out, attitudes are structured in three parts: they have a cognitive component (beliefs about languages), an affective component (feelings and emotions regarding languages), and a behavioural component (a predisposition to act in a certain way) (Lasagabaster, 2003).

---

<sup>2</sup> The close relationship between language and identity (and ethnicity) has often been denounced or presented as a problem in academic and popular discourse. See May (2001) for an exposition and critique of this problem area.

<sup>3</sup> We will not go further here into whether the identities are chosen or imposed. For a discussion of this question see Tabouret-Keller (1997).

While the "attitude" construct has been extensively studied in Catalan socio-linguistics, language confidence has not been dealt with in socio-linguistic studies of Catalan. Confidence is defined as "a relative lack of anxiety when using an L2 coupled with the belief in being able to cope linguistically with the L2 situation" (Rubinfeld et al. 2006). According to various studies by Clément and associates, confidence is a key construct in explaining the use of a second language and the identification of an individual with the second language group. Broadly speaking, Clément's studies have shown the following: "contact with or confidence in a second language leads individuals to identify with the L2 community. This process brings about more positive representations of the L2 culture. On a day to day basis, these studies suggest that learning an L2 may have a positive influence on inter-group relations. In the context of learning an L2, we find a greater identification with that community, which consequently makes us feel more positive about the community" (Rubinfeld, et al., 2006, p. 627).

Lastly, this paper includes the use of language in social networks as a variable which could explain identification with specific languages. As Tabouret-Keller (1997) explains, the psychological component stems from both the relationship with one's parents and the way in which one imitates the behaviour of others. The analysis of social networks provides a relational view of social reality (Molina, 2005), i.e. the relations which individuals have with other individuals around them. While the classical approach to social sciences is based on the study of individual attributes such as gender, age or level of education, the analysis of social networks focuses on relations between individuals. The main subject of study when considering social networks is, then, the nature of the interaction between individuals. These include the existence of a relationship, its strength and its characteristics. Nevertheless, the attributive and relational views are complementary in sociological or psychological analysis.

When we turn to the relationship between language and identity on an institutional level (explicit or implicit legal measures), we need to take into consideration the importance of language policy, both in general and in education, in the socio-linguistic context of Catalonia and the Aragonese border region. The most obvious difference between Catalonia and Aragon is that in the former Catalan is the language of instruction in primary education, while in Aragon it has only a token presence. Other differences are shown in Table 1.

**Table 1. Socio-linguistic description of Mataró and the Aragonese border area**

	Catalunya / Mataró	Aragonese border area
Population	7,210,508 / 119,858	47,686
Political status	Autonomous community	Part of the Autonomous Community of Aragon, with no political status of its own
Legal situation of Catalan	Laws on language policy (1983, 1998)	Not an official language. Recognised in the Law on the Aragonese cultural heritage (1999)
Language policy in education	Language immersion: Catalan is the language of instruction in primary and secondary education, the aim being competence in both Catalan and Spanish.	Pupils are entitled to two hours a week of Catalan as an optional subject at school

## 2.2. Language and identity in Catalonia

The relationship between language and identity in socio-linguistics in areas where Catalan is spoken is a topic which has been studied extensively from different theoretical points of view, although quantitative studies predominate (see Baldaquí, 2004; Bastardas, 2007; Fabà, 2005a, 2005b; Fundació Congrés de Cultura Catalana, 2006; Llobera, 2001; Pujolar, 1997; Querol, 2001; and Strubell, 2008, among others). The most recent data for Catalonia and the Aragonese border area is that provided by the corresponding statistical studies on language use (EULC 03, Torres, 2005; EULF 03, Sorolla, 2005; EUL 03, Querol et al., 2007). For example, EULC 03 figures for Catalonia show that the majority first language in Catalonia is Spanish (53.5% of the population, as against 40.4% for Catalan), while the majority language of identity is Catalan (48.8% of the population, as against 44.3% for Spanish) (Fabà, 2005a).

Most studies tend to examine two variables to see how the language of identity relates to another socio-linguistic variable. For example, Vila (2006) studies the relationship between

language of identity and language use and reports three facts: a) Catalan has a certain power of attraction (a considerable number of speakers have Spanish as their first language but identify with Catalan), b) figures for the use of Catalan are explained more readily by the proportion of speakers who identify with Catalan than the proportion of speakers who have Catalan as their first language and c) language transmission between generations in Catalonia favours Catalan (see also Fabà 2005a, 2005b and Torres 2005).

All the studies carried out up to now deal with samples of young individuals and adults (aged 15 and over in the case of EULC 03 and EULF 03). An exception, both because of its methodology and the age range of the sample, is the study by Baldaquí (2004), who researched the perception of ethno-linguistic vitality by young people (n=202, aged 13-14) in l'Alacantí county. The results of the study showed that the ethnolinguistic vitality perceived by respondents depends on both objective vitality (the sociolinguistic environment) and a series of sociolinguistic variables (first language, educational programmes, attitudes to language, and language learning or conservation). Querol (2001) also carried out multi-variable analyses to explain and predict language use in Catalonia by 16-year-old pupils according to three variables: the social representation of the languages, reference groups and social networks. Results for 2000 show that the main variables which predict the use of Catalan are the representation of Spanish, the social network in Catalan, the social network in Spanish, and identity. It should be noted that results for 1993 included the parents' language as one of the main variables but not social networks in Spanish.

A number of sociolinguistic studies have dealt with samples whose ages were similar to those of the subjects of this paper, although they were slightly older. They focus on the study of attitudes to language and its relationship to sociolinguistic variables (see Lasagabaster, 2003; Lasagabaster & Huguet, 2007, for a summary of these studies). The results given by Huguet (2007) for a sample of 309 university students in Girona and Lleida (average age 20) show that there is a significant difference in linguistic attitudes to Catalan according to the first language of the respondents (more favourable attitudes among those whose first language was Catalan), the language background at school (more favourable to Catalan among those educated in Catalan), and the predominant language in the town or city in which they live (more favourable to Catalan the more widely Catalan was used). Other variables, such as gender, social and professional status, and the population of the town or city, did not give significant differences.

Finally, Huguet and Suils (1998) studied attitudes to language among the pupils in the second year of compulsory secondary education (the eighth year of basic general education at the time of the study, 1995-1996, n=257) in the border area between Catalonia and Aragon (Baix Cinca and Baix Segre). They found that pupils in both areas had positive attitudes to Catalan and that the variable which could explain differences in attitude was the family language background, even though the situations observed were reversed: attitudes to Catalan were positive in the Baix Segre, while they were positive to Spanish in the Baix Cinca; however, less positive attitudes to Spanish were recorded in the Baix Segre and to Catalan in the Baix Cinca, depending on the first language of the respondents (for example, a less favourable attitude to Spanish among Catalan speakers in the Baix Segre).

In short, previous studies have demonstrated the possible interrelation between different sociolinguistic variables when explaining those variables which are considered primary (or independent). Depending on the methodology of the studies, what has been considered primary has been the use, identity or attitude studied. Today, therefore, we have a jigsaw of sociolinguistic studies which provide multiple sets of data about samples varying widely, both in age and in sociolinguistic background. This study deals with an age range which has not been extensively studied (pupils in the sixth year of primary education) and presents a descriptive and discriminant analysis, which attempts to explain the relationship between the respondents' language of identity and the language they speak with their parents. The fact that the data comes from two areas which are markedly different from a sociolinguistic point of view also enables us to investigate the impact which the sociolinguistic environment has on the language of identity and the language used with parents.

The study poses the following questions:

1. What is the language of identity of pre-adolescents in the sixth year of primary education in Mataró and the Aragonese border region?

2. What language do young people speak with their parents in Mataró and the Aragonese border region?
3. How do sociolinguistic variables (oral competence, social network, attitude, confidence, language spoken with siblings) relate to the different groups for language of identity and language used with parents?

### 3. Methodology of the study

A total of 1,017 pupils in the sixth year of primary education at state and grant-maintained schools in Mataró and 245 at schools in el Baix Cinca and la Llitera (Aragonese border area)<sup>4</sup> took part in the study. The average age of those participating was 12. The breakdown by gender in Mataró was 506 boys (50.5%) and 495 girls (49.5%); in the Aragonese border area it was 126 boys (51.9%) and 117 girls (48.1%).

To obtain data a sociolinguistic questionnaire was distributed in which questions were asked about the variables being studied, as well as others not dealt with here (for example, media consumption, motivation, etc.). Table 2 includes a sample item for each of the variables studied and the source from which the questions or items have been taken.

**Table 2. Sociolinguistic variables, examples of items and source of items**

Sociolinguistic variable	Sociolinguistic questionnaire	Source
Language of identity (1 item)	Language of identity (1 item)	EULC 03
Language used with parents (2 items)	<u>Question:</u> What language do you speak with your father? What language do you speak with your mother? <u>Possible answers:</u> Only Catalan or mostly Catalan/ Catalan and Spanish equally/ only Spanish or mostly Spanish/ another language	EULC 03
Attitude to language (10 items for Catalan and 10 for Spanish)	<u>Statement:</u> I think it's good for all boys and girls in Catalonia to study Catalan. <u>Possible answers:</u> Yes - No	Huguet & Suïls (1998)
Language competence (4 items, one for each language skill)	<u>Statement:</u> I can speak Catalan. <u>Answer on scale:</u> 0 1 2 3 4 5 6 7 8 9 10	EULC 03
Social network	<u>Instructions:</u> Write the names of the 20 people you have most contact with. What language do you use to speak to this person? <u>Opcions de resposta sobre les llengües:</u> català/ igual en català i castellà/ castellà/ altres llengües	Molina (2005)
Language confidence (6 items)	<u>Statement:</u> I think I know enough Catalan to write it easily . <u>Possible answers:</u> 5-point Likert scale (agreement-disagreement)	Clément & Baker (2001) (adapted)

For each variable indices are obtained from one or more items. For example, in the case of language of identity, the index was obtained from a single item (shown in Table 2), while the index for the language spoken with the respondent's parents is based on two items (language spoken with the father and language spoken with the mother).

The definitions of the variables *language used with parents* and *language of identity* were the same as those used for EULC 03. The term *own language* was not used, as this concept has often been used in a more political and institutional sense. The language of identity is,

<sup>4</sup> Col·legi Rural Agrupat (CRA) La Llitera (Albelda, El Campell, Valldellou, Camporrells, Castellonroi and Estopanyà), Tamarit de la Llitera, CRA El Torricó (Algaió, El Torricó and Vensilló), La Melusa, Saidí, Fraga and CRA Ribera del Cinca (Torrent de Cinca, Vilella de Cinca and Mirasolt).

therefore, taken to be the language which the respondent considers to be "his (or her) language".

The variable *language used with parents* was based on two items in the questionnaire (Table 3). Answers to these two items were coded as shown in the third column of Table 3, the categories "Catalan" and "Spanish" referring to the almost exclusive use of the languages with both parents, while the category "Catalan and Spanish" referred to various possible situations.

**Table 3. Coding of the variable "Language used with parents"**

Item 1: What language do you speak with your father?	Item 2: What language do you speak with your mother?	Coding of the variable "Language used with parents".
Exclusively or mainly in Catalan	Exclusively or mainly in Catalan	Catalan
Exclusively or mainly in Catalan	Both Catalan and Spanish	Catalan and Spanish
Exclusively or mainly in Catalan	Exclusively or mainly in Spanish	
Both Catalan and Spanish	Both Catalan and Spanish	
Both Catalan and Spanish	Exclusively or mainly in Catalan	
Exclusively or mainly in Spanish	Exclusively or mainly in Catalan	Spanish
Exclusively or mainly in Spanish	Exclusively or mainly in Spanish	
In another language	In another language	In another language
In another language	Exclusively or mainly in Catalan	
In another language	Both Catalan and Spanish	
In another language	Exclusively or mainly in Spanish	
Exclusively or mainly in Catalan	In another language	
Both Catalan and Spanish	In another language	
Exclusively or mainly in Spanish	In another language	

The index to assess attitudes to Catalan was based on answers to 10 items, 5 of which indicated a positive attitude to the language while 5 showed negative attitudes. The scale of attitudes goes from -1 to 1. On this scale -1 signifies an unfavourable attitude to Catalan, while 1 indicates a favourable attitude.

The index for confidence combined 12 items, 6 of which were related to language confidence and 6 to anxiety. The index for confidence ranges from -1 to 1. On this scale -1 indicates a lack of confidence in Catalan and 1 indicates confidence when speaking Catalan.

Language competence takes only one item into account: the respondent's self-assessment of his/her ability to speak Catalan on a scale from 0 to 10, where 0 indicates that the respondent cannot speak Catalan at all and 10 indicates complete fluency in Catalan.

Finally, to assess the social network the index was based on the number of friends and classmates given by each respondent and the language in which they spoke to each other. The index is a scale from 0 to 1 on which 0 signifies that the person never uses Catalan with friends and classmates and 1 indicates that Catalan is the only language used with them.

To analyse the data descriptive techniques (percentages) and multivariable discriminant analysis were used. Discriminant analysis enables us to verify and confirm the consistency of the distinction between different groups of subjects. The fundamental objective of this technique is to analyse and explain the differences between groups in such a way that it is possible to predict the likelihood of a person's belonging to a particular group.

## 4. Results

### 4.1. Language used with parents and language of identity

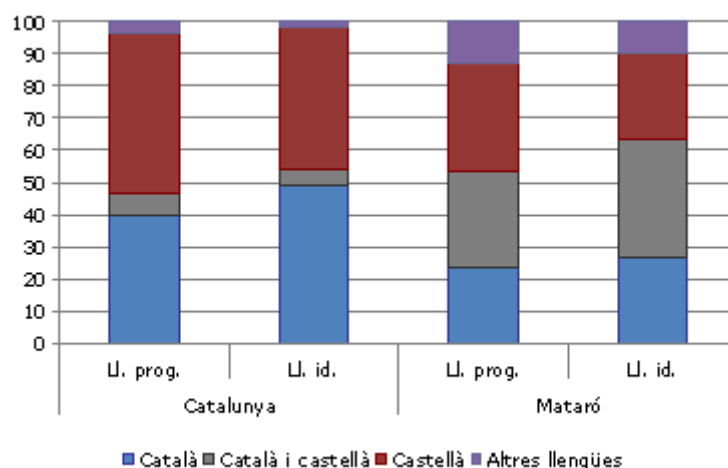
Table 4 (figure 1) compares data on the language used with parents and the language of identity in the EULC 03 study (population of Catalonia aged over 15) with data for the sample from Mataró in this study (pupils aged 11-12).

**Table 4. Language used with parents and language of identity in Catalonia and Mataró**

	Language used with parents				Language of identity			
	Catalonia		Mataró		Catalonia		Mataró	
	N	%	N	%	N	%	N	%
Catalan	2113	39,5	242	23,8	2670	48,8	268	26,6
Catalan and Spanish	371	6,9	298	29,4	283	5,2	366	36,3
Spanish	2689	50,3	343	33,8	2425	44,3	274	27,2
Other languages	174	3,3	132	13,0	92	1,7	101	10,0
TOTAL	5347	100	1015	100	5470	100	1009	100

Note: Data for Catalonia are from the EULC 03 study. Data for Mataró is drawn from this study.

**Graphic 1. Language used with parents and language of identity in Catalonia (EULC 03) and sixth-year primary pupils in Mataró**



If we compare data for all of Catalonia and Mataró on the language used with parents, it will be seen that most of the population of Catalonia uses either Spanish (50.3%) or Catalan (39.5%). In Mataró, on the other hand, in the case of pupils aged 11-12, it will be observed that a considerable percentage speak to them in both Catalan and Spanish (29.4%, as against 6.9% for the whole of Catalonia).

A similar trend can be observed if we compare the language of identity in Catalonia and in Mataró. While only 5.2% of respondents among the population of Catalonia identify with both Catalan and Spanish, 36.3% of the Mataró pupils identify with both languages.

If we compare data for language used with parents and language of identity in Table 4, three trends can be observed.<sup>5</sup> First, the percentage of respondents who report Spanish as the language they use with their parents is higher than the percentage who identify with Spanish (33.8% and 27.2% respectively). Secondly, the percentage of pupils who identify with both Catalan and Spanish is higher than the proportion who report that they use both languages with their parents (36.3% and 29.4% respectively). Lastly, the difference between those who

<sup>5</sup> This paper focuses on the categories for Catalan, Spanish and combinations of these two, and does not deal with other languages.

identify with Catalan and speak Catalan with their parents is not so marked as in the other cases mentioned (26.6% and 23.8%, respectively).

Table 5 (Figure 2) allows us to compare the total population of the Aragonese border area with data for pre-adolescents in this area. Of the whole population, the majority use Catalan when speaking to their parents (71.6%), followed by Spanish (19.6%), while only a small percentage (8.1%) use both Catalan and Spanish. However, if we consider the pre-adolescent population, the percentage who use both Catalan and Spanish with their parents is 29.3%.

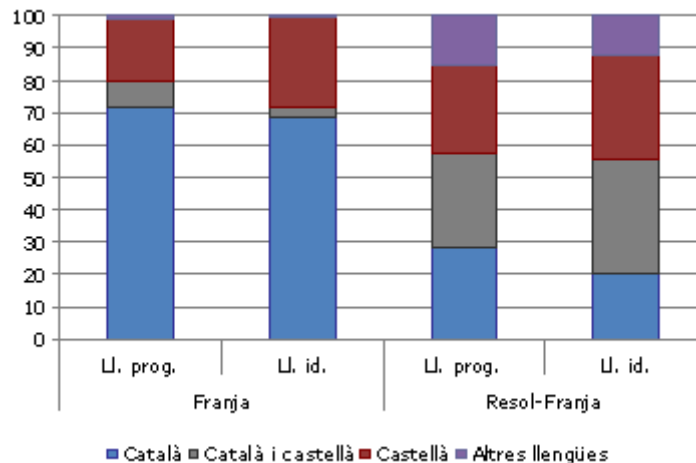
Regarding the language of identity, the figures show that 35.5% of the pre-adolescent population identify with both Catalan and Spanish, while only 3.1% of the whole population of the Aragonese border area identified with both languages.

**Table 5. Language used with parents and language of identity for the whole population of the Aragonese border area and among pupils in the sixth year of primary education in certain towns and villages in the area.**

	Language used with parents				Language of identity			
	Aragonese border area		Aragonese border area - RESOL		Aragonese border area		Aragonese border area - RESOL	
	N	%	N	%	N	%	N	%
Catalan	622	71,6	68	28,1	596	68,6	49	20,2
Catalan and Spanish	70	8,1	71	29,3	27	3,1	86	35,5
Spanish	170	19,6	66	27,3	243	28,0	78	32,2
Other languages	7	0,8	37	15,3	3	0,3	29	12,0
TOTAL	869	100	242	100	869	100	242	100

Note: Data for the Aragonese border area is from the EULF 03 study. Data for the Aragonese border area - RESOL is that used in the present study.

**Figura 2. Llengua amb els progenitors i d'identificació a la totalitat de la Franja (EULF 03) i a la població escolar de sisè a poblacions específiques de la Franja**



To explain the differences between the language used with parents and the language of identity, it is useful to examine the following specific case: the language of identity of pupils who speak Spanish with their parents (Table 6). The figures in the table show that in both Mataró and the Aragonese border area the majority of those who use Spanish with their parents identify with this language, although the proportion is much higher in the Aragonese border area than in Mataró (72.7% and 58.3% respectively), while those who do not identify exclusively with Spanish identify with both languages, especially in Mataró (37.3%).



**Table 6. Language of identity of respondents who use Spanish with their parents.**

	Mataró		Aragonese border area	
	N	%	N	%
Catalan	11	3,2	6	9,1
Catalan and Spanish	128	37,3	11	16,7
Spanish	200	58,3	48	72,7
Other languages	4	1,2	1	1,5
TOTAL	343	100	66	100

We can summarise the data in Tables 4 – 6 regarding language of identity and language used with one's parents as follows:

1. The samples of sixth year primary pupils in Mataró and the Aragonese border area are similar in one important aspect: bilingual identification is the most frequently occurring value in both populations (36.3% in Mataró and 35.5% in the Aragonese border area) in contrast to the total populations of Catalonia and the Aragonese border area, in which the majority identify with one language, be it Catalan or Spanish.
2. Generally speaking, in the preadolescent population, both in Mataró and in the Aragonese border area, the language of identity and the language used with parents do not coincide, the only exception being that of Catalan in Mataró. In particular, we can observe the following trends:
  - In Mataró the proportion having Catalan as their language of identity is greater than the percentage using it with their parents, while in the case of Spanish the percentage using the language with their parents is greater than the percentage who give it as their language of identity. In the Aragonese border area these trends are reversed.
  - There is a percentage of pupils with both Catalan and Spanish as languages of identity who do not use both with their parents: in Mataró 29.4% report that they use both languages with their parents, while 36.9% identify with both languages. In the Aragonese border area 29.3% report that they use the two languages with their parents, but 35.5% identify with both languages.
3. In the case of respondents who use Spanish with their parents, the lower proportion giving it as their language of identity is largely accounted for by the percentage identifying with both Catalan and Spanish.

#### 4.2. Descriptive analysis

For a closer study of the relationship between language of identity and the language used with parents we will study the case of pupils who use Spanish with their parents to determine the factors which may account for their language of identity. In particular the use of Spanish with parents and the language given as language of identity are studied with regard to five sociolinguistic variables which have been related to language and identity in previous studies: language competence, social networks, attitudes, confidence and the language spoken with siblings.

##### 4.2.1. Oral competence in Catalan

All the pupils in Mataró give themselves high scores (average 9 on a scale out of 10) for oral competence in Catalan. The average for pupils who use Spanish with their parents is only 0.4 points less than the average for all the pupils in the study. If we divide the pupils who use Spanish with their parents into those who identify with both Catalan and Spanish and those

who identify only with Spanish, the former have scores which are much closer to the average for all the pupils than the latter.

**Table 7. Competence in spoken Catalan in Mataró and the Aragonese border area**

	Mataró		Aragonese border area	
	N	Average	N	Average
All pupils	990	9,0	205	7,3
Pupils with Spanish as LP	334	8,6	55	5,8
Pupils with Spanish as LP and LI	194	8,3	39	5,1
Pupils with Spanish as LP and both Catalan and Spanish as LI	125	8,8	10	8,1

Note: LP = language used with parents. LI = language of identity

Figures for language competence in the Aragonese border area exhibit three differences from those for Mataró. Firstly, the overall figure for competence is lower (an average of 7.3). Secondly, those pupils who identify to some extent with Catalan have the highest scores for competence in Catalan (8.1 for those who identify with both Catalan and Spanish). Lastly, pupils who identify with Spanish have the lowest score.

#### 4.2.2. Social networks

An analysis of the social networks of pupils in Mataró shows that on the whole they use Spanish in their social networks. The average result is 0.3 on a scale from 0 (Catalan is never used) to 1 (only Catalan is used). Youngsters in Mataró use Catalan with 3 out of every 10 of their friends. The use of Spanish is even more pronounced in the case of pupils who use Spanish with their parents and identify with Spanish.

**Table 8. Social networks and the use of Catalan in Mataró and the Aragonese border area**

	Mataró		Aragonese border area	
	N	Average	N	Average
All pupils	643	0,3	159	0,2
Pupils with Spanish as LP	207	0,1	43	0,07
Pupils with Spanish as LP and LI	123	0,08	28	0,02
Pupils with Spanish as LP and both Catalan and Spanish as LI	74	0,1	28	0,2

Data for social networks in the Aragonese border area shows a similar index to that for Mataró if we consider all the pupils (0.2). It will be seen that pupils who identify with Catalan and Spanish have networks in which Catalan is used more extensively, while those who use Spanish with their parents hardly use Catalan at all in their social networks.

#### 4.2.3. Attitudes to Catalan

In Mataró pupils' attitudes to Catalan are very favourable (0.7 on a scale from -1 to 1). There is little difference between sub-groups.

**Table 9. Attitudes to Catalan in Mataró and the Aragonese border area.**

	Mataró		Aragonese border area	
	N	Average	N	Average
All pupils	882	0,7	198	0,6
Pupils with Spanish as LP	304	0,6	53	0,5
Pupils with Spanish as LP and LI	181	0,6	38	0,4
Pupils with Spanish as LP and both Catalan and Spanish as LI	108	0,7	10	0,7

Attitudes to Catalan in the Aragonese border area are generally less favourable than in Mataró (0.6). As observed elsewhere, pupils who include Catalan as a language of identity have more favourable attitudes, while the attitudes of those whose only language of identity is Spanish are less favourable.

#### 4.2.4. Linguistic confidence in Catalan

In Mataró the pupils have a medium index of confidence in Catalan (0.5 on a scale from -1 to 1) As in the case of attitude, the differences between sub-groups in Mataró are minimal.

**Table 10. Confidence in Catalan in Mataró and the Aragonese border area.**

	Mataró		Aragonese border area	
	N	Average	N	Average
All pupils	889	0,5	205	0,3
Pupils with Spanish as LP	304	0,4	56	0,1
Pupils with Spanish as LP and LI	184	0,4	39	0,06
Pupils with Spanish as LP and both Catalan and Spanish as LI	107	0,5	10	0,3

In the case of the Aragonese border area, the differences are more pronounced. Overall the pupils in the Aragonese border area report a lower level of confidence in Catalan than the pupils in Mataró. This is especially true of those pupils who identify only with Spanish, where the figure is 0.06.

#### 4.2.5. Language used with siblings

Figures for the use of Spanish and Catalan with siblings indicate that in Mataró the language used most by the pupils with their siblings is Spanish (32%), followed by those who use Catalan (28.7%) and Spanish and Catalan (27.7%). However, the figures are different if we consider only pupils who use Spanish with their parents: of these, 68.9% also use Spanish with their siblings, while 26.2% use Catalan and Spanish (the percentages for pupils having Catalan and Spanish as language of identity are very similar).

**Table 11. Language used with siblings in Mataró**

	Catalan		Catalan and Spanish		Spanish		Other languages	
	N	%	N	%	N	%	N	%
All pupils	253	28,7	245	27,7	283	32,0	102	11,5
Pupils with Spanish as LP	5	1,6	80	26,2	210	68,9	10	3,3
Pupils with Spanish as LP and LI	4	2,2	36	19,9	137	75,7	4	2,2
Pupils with Spanish as LP and both Catalan and Spanish as LI	1	0,9	28	25,2	76	68,5	6	5,4

Pupils in the Aragonese border area, like those in Mataró, use Spanish most often with their siblings. (36.7%) The second largest group (28%) use Catalan, while a third group (23.2%) use both Catalan and Spanish. Of those who use Spanish with their parents, the great majority (84.7%) speak Spanish to their siblings. Results for language of identity show that pupils who include Catalan when speaking to their siblings tend to identify with both Catalan and Spanish.

**Table 12. Language used with siblings in the Aragonese border area.**

	Catalan		Catalan and Spanish		Spanish		Other languages	
	N	%	N	%	N	%	N	%
All pupils	58	28,0	48	23,3	76	36,7	25	12,1
Pupils with Spanish as LP	0	0	9	15,8	48	84,2	0	0
Pupils with Spanish as LI	0	0	5	11,6	38	88,4	0	0
Pupils with both Catalan and Spanish as LI	0	0	3	42,9	4	57,1	0	0

#### 4.3. Multivariable analysis: discriminant analysis

Descriptive analysis allows us to identify certain tendencies regarding the relationship between language of identity and language used with parents but it does not tell us which of the variables studied provides the best explanation for the figures. In this section the technique of discriminant analysis is applied to the data in the study, focusing on those pupils who use Spanish with both parents. The object of this analysis is to determine which of the variables studied best explains the language of identity of pupils who use Spanish with their parents. The pupils were therefore divided into two groups (dependent variable: those who identify only with Spanish and those who identify with both Catalan and Spanish).

Canonical correlation determines the predictive capacity of the independent variables (oral competence, social networks, attitudes, confidence and language used with siblings) included in the analysis. The value for canonical correlation ranges from 0 to 1. Values of around 0.7 are considered to indicate a good model for distinguishing between the two groups.

Discriminant analysis of the data for the Aragonese border area shows that the five sociolinguistic variables permit quite good discrimination for the two groups of language of identity<sup>6</sup> (canonical correlation higher than 0.7). A large number of respondents could not be included in this first analysis, as many of the pupils left one or more questions unanswered and discriminant analysis excludes them completely, even if they have answered all the other questions. Only 22 individuals could be considered, those who had answered all the questions. To improve the analysis a neutral value was assigned to missing cases (they totalled 59). The value assigned was the average obtained from the other pupils who spoke Spanish with their parents and who had given a valid response. In this second analysis the value for canonical correlation was lower (0.548), as had been expected, since the missing values were replaced by values which neutralised the initial differences between cases.

The matrix of correlations between discriminant variables and functions shows that of the five independent variables two are fundamental in determining the function which gives the best discrimination between those Spanish speakers who identify with Spanish and those who identify with both Spanish and Catalan: the presence of Catalan in the social network of friends (0.67) and knowledge of spoken Catalan (0.642). The predictive value of the other three variables - confidence in Catalan (0.443), attitude to Catalan (0.427) and language spoken with siblings (0.416) - is noticeably lower. These three variables can be excluded from the model with hardly any reduction in prediction of the dependent variable.

**Table 13. Matrix of correlation between discriminant variables and canonical discriminant functions for the Aragonese border area.**

Independent variables	Value of the canonical discriminant functions
Presence of Catalan in the social network of friends	0,670
Knowledge of Catalan: speaking	0,642
Confidence in Catalan	0,443
Attitude to Catalan	0,427
Presence of Catalan in interaction with siblings	0,416

Discriminant analysis of the data from Mataró shows that the power of prediction of the five independent variables is not so great as in the case of the Aragonese border area. In the analysis without substitution of missing values (141 valid cases) a canonical correlation of 0.332 is obtained. If the missing values are replaced by the average for those who speak Spanish with their parents (324 cases), an even lower canonical correlation is obtained (0.258).

Data from the correlation structure matrix for discriminant variables and functions shows, as in the case of the Aragonese border area, that it is the presence of Catalan in social networks that best discriminates between the two groups of pupils who speak Spanish with their parents (those who identify with Spanish and those who identify with both Catalan and Spanish). Moderate correlation is given by confidence and attitude to Catalan. However, oral

<sup>6</sup> The analysis did not include respondents who used Spanish with their parents but identified with Catalan. There were 6 such cases in the Aragonese border area and 11 in Mataró.

competence and the use of Catalan with siblings show relatively little correlation with the discriminant function and the model is not weakened if they are excluded.

**Table 14. Matrix of correlation between discriminant variables and canonical discriminant functions for Mataró.**

Independent variables	Value of the canonical discriminant function
Presence of Catalan in the social network of friends	0,648
Confidence and anxiety when using Catalan	0,644
Attitude to Catalan	0,528
Knowledge of Catalan: speaking	0,485
Presence of Catalan in interaction with siblings	0,154

## 5. Discussion and conclusions

The results of this study show that in the population of sixth-year primary pupils in the two areas studied, bilingual identification (Catalan and Spanish) is much more common than in the adult population. It also shows that the language which these pre-adolescents use with their parents does not necessarily coincide with the language with which they identify. Discriminant analysis of the population who use Spanish with their parents has revealed that it is the presence of Catalan in social networks which leads some pupils to include Catalan as one of their languages of identity, both in Mataró and in the Aragonese border area. It was also observed that in the Aragonese border area, as well as social networks, oral competence in Catalan discriminates the two groups, while in Mataró discrimination depends on confidence and the attitude to Catalan.

Two main differences can be observed between the populations of Mataró and the Aragonese border area. Firstly, in Mataró a greater number of Spanish speakers report that they identify with both Catalan and Spanish. In Mataró this category accounts for 40% of all those who speak Spanish with their parents (Table 4), while in the Aragonese border area the figure is just over a quarter (Table 5). In addition to this major difference we also find two distinct situations regarding oral competence in Catalan. While in Catalonia self-assessment of competence is very high in the two groups studied (Spanish as language of identity vs. both Catalan and Spanish), with very little difference between them, in the Aragonese border area the figure is high only for those who identify with both languages. These low levels of competence for many Spanish speakers in the Aragonese border area may be a factor impeding greater identification with both languages. However, as shown by data from Mataró, once high levels of oral competence are generalised in all groups, this factor does not normally enable us to discriminate between the two groups studied and the differences are best explained by the social network and other variables (confidence and attitude).

The second difference between the two populations studied is precisely the type of variable that discriminates between the two groups. While the first variable (social networks) is the same, those which follow are not. In the Aragonese border area it is competence which largely explains identification with a single language (Spanish in the majority of cases) or two. The fact that in the Aragonese border area competence in Catalan is not so widespread may make it a differentiating factor for some speakers, affecting whether they identify with the language or not. In Mataró, on the other hand, we find that the majority of respondents report that they have a high level of competence in Catalan (as a result of the education system). This makes it easy for a larger number of young people to identify with Catalan. It is, however, possible that competence in both languages may reach a maximum level, after which this factor will no longer favour identification with the two languages, and will therefore cease to be a discriminating factor. Other factors, such as confidence and attitude, are then likely to be those which allow us to discriminate between the two groups. The data thus shows that language competence may be a factor which explains identification with a language but that, once competence is guaranteed, there are other factors (less linguistic and more social) which help to explain the language identification of Spanish speakers.

These results have implications for educational and language policies to support Catalan. The fact that in both Mataró and the Aragonese border area it is social networks with friends which allow us to best discriminate between the two groups suggests that the presence of Catalan in social networks is sufficient for some pupils who use Spanish with their parents to

include Catalan as a language of identity. Measures to promote Catalan should, therefore, focus on fostering Catalan and ensuring its use in peer networks. It is not an easy task, as it is much easier to take measures on a micro-sociolinguistic level than to intervene in interpersonal use. There are different areas in which it is possible to influence social networks. The school is one but extra-curricular activities should not be forgotten: these provide informal contexts which can help to foster the use of Catalan. In short, it is not a matter of influencing the individual behaviour of young people so that they use Catalan but of at least guaranteeing the free development of social networks in Catalan at crucial moments of resocialisation and the recreation of social networks (for example, at the beginning of the school year or at the transition from primary school to secondary school).

Future studies should look more deeply into the ways in which social networks may be influenced and how they are configured from the linguistic and dynamic points of view (the transition from primary school to secondary, for example). Such studies should also include other multivariable analyses to reconsider the relationship between language of identity and other variables which may be related to it, such as first language, motivation or media consumption. Lastly, further improvements in statistical models are called for, so that we can go beyond correlational analysis (indicating a correlation rather than cause and effect) and determine the reasons which make an individual identify with one language or more than one (cf. Torrijos, 2008). In conclusion, the results of this study are yet another demonstration that the relationship between language and identity cannot be explained by a single factor; the complexity of this relationship calls for the joint study of a series of factors in more than one place and at different times, so that we can appreciate the dynamics of individual identification and that of young people in particular.

## 6. References

- BALDAQUÍ, J. M. La percepció de la vitalitat etnolingüística pels jòvens de l'Alacantí. Barcelona: Institut Interuniversitari de Filologia Valenciana, Publicacions de l'Abadia de Montserrat, 2004.
- BASTARDAS, A. Les polítiques de la llengua i la identitat a l'era "glocal". Barcelona: Institut d'Estudis Autònoms, 2007.
- BOIX, E. i VILA, F. X. Sociolingüística de la llengua catalana. Barcelona: Ariel, 1998.
- COMAJUAN, L. La escala de disrupció intergeneracional para el estudio de la inversión de la sustitución lingüística: el caso del quechua en el Perú. A: S. Coronel-Molina & L. Grabner-Coronel (eds.) *Lenguas e identidades en los Andes*. Quito: Abya Yala, 2005, p. 83-114.
- FABÀ, A. 2005a «Llengua inicial i llengua d'identificació» A: *Estadística sobre els usos lingüístics a Catalunya 2003. Llengua i societat a Catalunya en els inicis del segle XXI*. Generalitat de Catalunya, Secretaria de Política Lingüística. Barcelona, 2005.
- FABÀ, A. 2005b «Usos lingüístics i grups identitaris. Preferències i opinions lingüístiques» A: *Estadística sobre els usos lingüístics a Catalunya 2003. Llengua i societat a Catalunya en els inicis del segle XXI*. Generalitat de Catalunya, Secretaria de Política Lingüística. Barcelona, 2005.
- FISHMAN, J. Reversing language shift. Clevedon: Multilingual Matters, 1991.
- FISHMAN, J. Can threatened languages be saved? Clevedon: Multilingual Matters, 2001.
- FUNDACIÓ CONGRÉS DE CULTURA CATALANA. Llengua i identitat. VII Jornades de debat Llengua i identitat. Cardona, 25 i 26 de novembre de 2005. Barcelona: Fundació Congrés de Cultura Catalana, 2006.
- GARRETT, P, COUPLAND, N. i WILLIAMS, A. Investigating language attitudes. Cardiff: University of Wales Press, 2003.
- HUGUET, À. Language use and language attitudes in Catalonia. A: Lasagabaster, D. i Huguet, À. (eds.). *Multilingualism in European bilingual contexts*. Clevedon: Multilingual Matters, 2007, p. 17-39.
- HUGUET, À. i SUÏLS, J. Llengües en contacte i actituds lingüístiques. El cas de la frontera catalano-aragonesa. Barcelona: Horsori, 1998.
- LASAGABASTER, D. Trilingüismo en la enseñanza. Actitudes hacia la lengua minoritaria, la mayoritaria y la extranjera. Lleida: Milenio, 2003.
- LASAGABASTER, D. i HUGUET, À. (eds.). *Multilingualism in European bilingual contexts*. Clevedon: Multilingual Matters, 2007.
- LLOBERA, J. R. La força de la llengua en la identitat nacional. L'experiència de l'Europa occidental. A: T. Mollà (ed.) *Ideologia i conflicte lingüístic*. Alzira: Bromera, 2001, p. 57-90.

- MACINTYRE, P. D., CLÉMENT, R., DÖRNYEI, Z., i NOELS, K.A. Conceptualising willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journal*, (1998), núm. 82, p. 545-562.
- MAY, S. *Language and minority rights. Ethnicity, nationalism and the politics of language*. Londres: Pearson, 2001.
- MOLINA GONZÁLEZ, J. L. «El estudio de las redes personales: contribuciones, métodos y perspectivas». *Empiria, Revista de Metodología de Ciencias Sociales*. [Madrid] (2005), núm. 10, p. 71-105.
- PAVLENKO, A. i BLACKLEDGE, A. Introduction: new theoretical approaches to the study of negotiation of identities in multilingual contexts. A: A. Pavlenko i A. Blackledge (eds.) *Negotiation of identities in multilingual contexts*. Clevedon: Multilingual Matters, 2004, p. 1-33.
- PUJOLAR, J. *De què vas, tio?* Barcelona: Empúries, 1997.
- QUEROL, E. "Evolució dels usos i de les representacions socials de les llengües a Catalunya (1993-2000). *Noves SL Revista de Sociolingüística*. [Barcelona] (2008), estiu 2001.
- QUEROL, E. (coord); CHESSA, E.; SOROLLA, N.; TORRES, J.; VILLAVARDE, J.A. *Llengua i societat als territoris de parla catalana a l'inici del segle XXI. L'Alguer, Andorra, Catalunya, Catalunya Nord, Comunitat Valenciana, la Franja i les Illes Balears*. Barcelona, Generalitat de Catalunya, 2007.
- RUBENFELD, S., CLÉMENT, R., LUSSIER, D., LEBRUN, M., i AUGER, R. Second language learning and cultural representations: beyond competence and identity. *Language Learning*. [Ann Arbor] (2006) núm. 56, p. 609-632.
- STRUBELL, M. *Llengua i identitat a Catalunya 2008*. Presentació Powerpoint. [http://www.uoc.edu/portal/catala/la\\_universitat/sala\\_de\\_prensa/noticies/2008/noticia\\_053.html](http://www.uoc.edu/portal/catala/la_universitat/sala_de_prensa/noticies/2008/noticia_053.html)
- SOROLLA VIDAL, N. "El català a la Franja. Any 2003". *Revista Ripacurtia* [Benavarri: Centre d'Estudis Ribagorçans (CERIB)] (2005), núm. 3, p. 239-253.
- TABOURET-KELLER, A. *Language and identity*. A: Coulmas, Florian (ed.). *The Handbook of Sociolinguistics*. Oxford: Blackwell, 1997, p. 315 - 326
- TORRES, J. (COORD.) i altres. *Estadística sobre els usos lingüístics a Catalunya 2003. Llengua i societat a Catalunya en els inicis del segle XXI*. Generalitat de Catalunya, Secretaria de Política Lingüística. Barcelona, 2005.
- TORRIJOS, A. "Caracterització sociolingüística dels joves de Catalunya: alguns factors que expliquen els usos lingüístics". *Noves SL Revista de Sociolingüística*. [Barcelona] (2008), hivern 2008, p. 1-17.
- VILA, F. X. Identificació lingüística i usos a Catalunya: algunes dades recents. A: Fundació Congrés de Cultura Catalana. *Llengua i identitat. VII Jornades de debat Llengua i identitat*. Cardona, 25 i 26 de novembre de 2005. Barcelona: Fundació Congrés de Cultura Catalana, 2006, p. 59-68.

**Vanessa Bretxa**

**Centre Universitari de Sociolingüística i Comunicació (UB)**

[vanessa.bretxa@ub.edu](mailto:vanessa.bretxa@ub.edu)

**Llorenç Comajoan**

**Centre Universitari de Sociolingüística i Comunicació (UB) i Universitat de Vic**

[llorenc.comajoan@uvic.cat](mailto:llorenc.comajoan@uvic.cat)

**Natxo Sorolla**

**Centre Universitari de Sociolingüística i Comunicació (UB) i Xarxa CRUSCAT – IEC**

[xosorolla@gmail.com](mailto:xosorolla@gmail.com)