

THE USE OF CATALAN BY MILLENNIALS IN CATALONIA: LESS INFLUENCE OF LINGUISTIC ORIGIN

Natxo Sorolla*

Aveli Flors-Mas**

Summary

The purpose of this study is to shed light, by analysing demolinguistic data from the *Enquesta d'usos lingüístics de la població* (Survey on language use of the population) of Catalonia in 2013, on the factors that condition the use of Catalan by young people in Catalonia. The overall results indicate that Catalan millennials (or Generation Y) use Catalan similarly to the population of Catalonia as a whole, that is to say at a rate of around 40%. The results of a multivariate analysis (multiple regression and path analysis) show that the language of identification is the explanatory variable most closely associated with language used by young people in Catalonia. The analysis highlights the explanatory importance of this attitudinal and affective dimension, which is greater than more structural variables, such as first language or declared oral proficiency of Catalan. The article concludes with some comments on the implications of results for policies promoting the use of Catalan in Catalonia.

Keywords: demolinguistics; Catalan language; young people; Catalonia; language use; language of identification.


L'ÚS DEL CATALÀ ENTRE ELS MILLENNIALS DE CATALUNYA: EL PES DILUÏT DE L'ORIGEN LINGÜÍSTIC

Resum

L'objectiu d'aquest treball és dilucidar, mitjançant l'anàlisi de dades demolingüístiques, quins factors condicionen l'ús del català entre la població jove de Catalunya a partir de dades de l'Enquesta d'usos lingüístics de la població a Catalunya de 2013. Els resultats generals indiquen que els membres de la generació Y o millennials catalans fan un ús del català similar al del conjunt de la població de Catalunya, a l'entorn de 0,4 sobre un total d'1. Per la seua banda, els resultats d'una anàlisi multivariable (regressió múltiple i anàlisi de camins) mostren que la llengua d'identificació és la variable més associada amb l'ús lingüístic dels joves de Catalunya. L'anàlisi subratlla, doncs, la importància explicativa de la dimensió actitudinal i afectiva per damunt de variables més estructurals com la llengua inicial o el coneixement oral declarat del català. Clouen l'article algunes reflexions sobre les implicacions dels resultats per a les polítiques de promoció del català a Catalunya.

Paraules clau: demolingüística; llengua catalana; joventut; Catalunya; usos lingüístics; llengua d'identificació.

*Natxo Sorolla, Universitat Rovira i Virgili, CUSC–Research Centre for Sociolinguistics and Communication, University of Barcelona, CRUSCAT–IEC network. natxo.sorolla@urv.cat  0000-0001-7480-0359

** AveHí Flors, CUSC–Research Centre for Sociolinguistics and Communication, University of Barcelona. aforsmas@ub.edu  0000-0003-1689-5956

Article received: 08.04.2020. Blind reviews: 03.05.2020 and 08.05.2020. Final version accepted: 14.05.2020.

Recommended citation: Sorolla, Natxo, & Flors-Mas, AveHí. (2020). The use of Catalan by millennials in Catalonia: less influence of linguistic origin. *Revista de Llengua i Dret, Journal of Language and Law*, 73, 50-68. <https://doi.org/10.2436/rld.i73.2020.3448>

Summary

1 Introduction

2 Theoretical framework

3 Methods

3.1 Study and analysis variables

3.2 Analytical techniques

4 Results

4.1 Use index

4.2 Which variables are associated with use of Catalan by young people in Catalonia?

4.3 How are these variables associated with language use?

5 Conclusions

Reference list

1 Introduction¹

The aim of this study is to provide a response to the following research question: which factors condition the use of Catalan by young people in Catalonia? In order to answer this question we analyse the demolinguistic data from the *Enquesta d'usos lingüístics de la població 2013* (2013 Survey on language use of the population of Catalonia) (EULP13), specifically, the data for cohorts between the ages of 15 and 34. Born in the 1980s and 1990s, this group is known as Generation Y, or, more popularly, as millennials (Howe & Strauss, 2009).

We first present the theoretical framework and the methodology used for the analysis, and the characteristics of the group under scrutiny. We then go on to present the results of a multivariate analysis which helps to establish the sociolinguistic and sociodemographic variables most commonly associated with the use of Catalan by young people in Catalonia. This section concludes with a further investigation of the relationship between the central variables that emerged from that analysis: first language, the language of identification, and knowledge of Catalan, as independent or explanatory variables, and the spoken language use index (IUP) of Catalan as dependent variable.

In general, the analysis shows that the language of identification is the variable that has the closest relationship with the use of Catalan by young people in Catalonia, more so than first language and knowledge of Catalan. It therefore shows that existing language policy—particularly in education—may have encouraged a certain “de-ethnicisation” of language use in Catalonia, such that knowledge and use of Catalan may have become detached from transmission of the language within family, as has been postulated by various studies that have focused particularly on Barcelona (Woolard, 1989, 2016; Boix-Fuster, 1993; Pujolar & González, 2013). Our data highlights that in a context in which linguistic origin has a reduced weight as an explanation for language use, elective factors such as identification with the language may have gained influence. At the same time, the data indicates that the current conditions of the linguistic market and current language policy in Catalonia do not encourage the consolidation of extensive use of Catalan among competent users for strictly instrumental purposes, with use resulting from a relationship of affection for and identification with the language.

2 Theoretical framework

The aim of this study is to analyse language use by young people in Catalonia and to establish the factors that condition the use of Catalan from a macrosociolinguistic perspective. The data analysed comes from the 2013 Survey on language use of the population of Catalonia, in which the young people (between the ages of 15 and 34) studied in the survey coincide fully with Generation Y, those born in the 1980s and 1990s, also known as millennials (Howe & Strauss, 2009). This generation was preceded by Generation X (born in the 1960s and 1970s) and followed by Generation Z, also known as Centennials. In general terms, members of Generation Y are characterised by being born at a time when the Soviet empire was imploding, giving way to a unipolar world dominated by the United States. In areas where Catalan is spoken, young people in Generation Y are characterised by having grown up in nuclear families, reaching a higher level of education than their parents with more going on to university, by adopting new ITC technologies while growing up, and by entering the labour market, initially at a time of economic boom and then during a time of crisis following the 2008 crash. These are, moreover, a particularly important generation in Catalonia from a sociolinguistic point of view, as they belong to the first group whose entire education was within the so-called Catalan conjunction model (Arnau & Vila, 2013; Milian, 1984). This means that they are characterised by a high degree of bilingualism and literacy in both Catalan and Castilian.

Language use in Catalonia has been amply studied within various different sociolinguistic research traditions, both from macrosociological and demolinguistic perspectives and in studies of a microsociological, qualitative, and/or ethnographic nature. Adolescents and young people have been scrutinised particularly closely. As Pujolar (2008) points out, this is in part because infants, adolescents and young people are more easily accessible as research respondents in educational institutions. To a great extent, it is also due to a

¹ This study received financial support from the Department of Culture of the Government of Catalonia as part of the *Impuls d'estudis sobre la situació del català a Catalunya 2018* (Support for studies on the situation of Catalan in Catalonia 2018) program carried out by the CUSC—Research Centre for Sociolinguistics and Communication of the University of Barcelona.

“certain sociological common sense that sees young people as the seed of society of the future, and, therefore, as the segment where we can best evaluate the progress of various political projects” (2008: 1). To attempt an exhaustive summary of the researches on language and young people in Catalonia is clearly beyond the scope of this article and also runs the risk of excluding important studies. We therefore refer readers to the thoroughly up-to-date summaries of Flors-Mas (2016), Newman & Trenchs-Parera (2015), and Trenchs-Parera & Newman (2015).

In general, the results of these studies point to a process of sociolinguistic depolarization and to a certain de-ethnicisation of language use among those born in Catalonia, inasmuch as the educational language model has tended to universalise young people’s knowledge of Catalan and Castilian (Vila, 2000), and has therefore facilitated the incorporation of Catalan by people for whom it is not their first language, although this does not happen automatically nor in all cases. This was less common in the years following the reinstatement of the Government of Catalonia, according to a number of studies (Flors-Mas, 2017; Flors-Mas & Vila, 2014; González et al., 2014; Newman et al., 2008; Pujolar & González, 2013; Sorolla, 2012; Vila, 2005; Woolard, 1989, 2016).

The demolinguistic empirical analysis on which our study is based finds that the use of Catalan by young people is grounded in three main foundations: first language, that is to say, the language the informant started to speak at home when very young; knowledge of Catalan, measured in terms of speaking ability on a scale of 0 (doesn’t speak at all) to 10 (speaks it perfectly); and lastly, the language of identification, which is the answer to the question “what is your language?” asked at the beginning of the survey and which gives an idea of speakers’ attitudes towards Catalan.

These three focal points dovetail neatly with the schema proposed by Berger & Luckmann (1966) in their treatise on socialization. For these sociologists, there are three processes in the development of individuals’ relationship with society. Individuals’ behaviour is first “externalised” by means of a process of ritualization, then objectivised, and lastly, subjectivised again by the individuals by means of a process of internalization. This converts society into a human product (externalization) and an objective reality (objectivisation) and, at the same time, converts humans into a social product (subjectivisation) (Berger & Luckmann, 1966, p. 79). What interests us here is the final process, which in sociolinguistics and linguistic anthropology has been studied within the paradigm of language socialization (Garrett & Baquedano-López, 2002; Schieffelin & Ochs, 1986). Linguistic socialisation involves the interiorisation of the different linguistic varieties of one’s communicative repertoires, the social norms needed to appropriate them for use in different socio-cultural contexts (with whom, when, for what purposes they should be used), and the meanings and social values of the different varieties, all the while constructing one’s own social and linguistic identity. At the same time, the differences between individuals in this process help us to understand why, within a similar context, two persons may engage in different socio-linguistic behaviour.

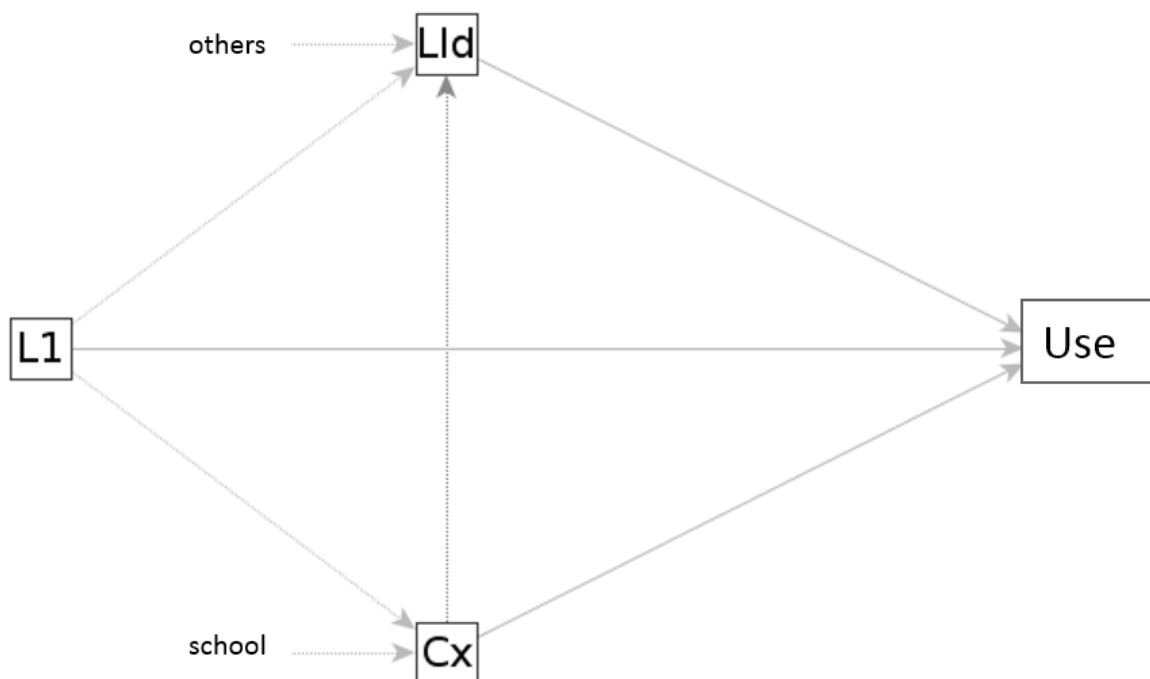
Berger & Luckmann suggest that the stage involving the internalization of society by individuals may be broken down into two sub-stages (1966, p. 149 *et seq.*). Primary socialisation involves a sub-process whereby the individual becomes a member of society during its infancy within the family. And secondary socialization is a later process in which the individual socializes in other spaces, such as educational institutions. Thus it would not be out of place to suggest parallelisms between the first language as an effect of primary socialization, proficiency in Catalan (also) as a product of the process of secondary socialization—at school, especially when the learner has not acquired Catalan as a first language—, and the language of identification as a synthesis of the processes of interiorising, subjectivising and developing an identification with the language or languages learnt during both primary and secondary socialisation.

In order to discover the relative weight of these variables and the primary and secondary socialisation stages in the use of Catalan by millennials in Catalonia, we used regression and path analysis techniques to establish the influence of various independent variables on a dependent variable. Path analysis, moreover, makes it possible to establish if there is mediation and sequencing between different variables, and to what extent any existing theoretical model among variables may correspond to the data. Rubenfeld et al. (2006), who apply this analysis to the influence of sociocultural and identity-related dimensions on learning a second language

in Canadian university students, point out that it is not possible to infer any causal relationship between the variables.²

Graphic 1 sets out the theoretical proposal we developed to answer our research question. Like a hypothetical model, it shows that both the first language and the language of identification and oral proficiency of Catalan are directly associated with Catalan use, with continuous lines linking them. At the same time, the first language, which is learnt during primary socialisation, is correlated with the internalization of the languages (language of identification) and with competence in Catalan. Lastly, the secondary socialisation process, with the intervention of agents such as schools and which reflects, in part, (the distinction between first language and) knowledge of Catalan, may also have an effect on the language that young people identify with (the language of identification).³

Graphic 1. Path analysis of Catalan use (Use) with first language (L1), language of identification (LId) and oral proficiency in Catalan (Cx) variables. Theoretical proposition.



3 Methods

3.1 Study and analysis variables

The analysis presented in this article is based on the 2013 Survey on language use of the population of Catalonia (EULP13) (Generalitat de Catalunya, 2015; Direcció General de Política Lingüística (Coord.), et al., 2018). A total of 7,255 people were surveyed, 1,869 of which were young people between the ages of 15 and 34, born between 1979 and 1998. They represent 25.8% of the total and are the group analysed here. When interpreting the results, it is important to take into account that, despite the high quality of EULP data in terms of demoscopic research standards, this data is produced by declarations made by the respondents themselves, and thus, must be understood as representations of their own uses, identifications and competence.

² “Because the data are usually collected using a correlational design, it would, however, be inappropriate to infer a causal relationship between aspects of the proposed model” (Rubinfeld et al., 2006: 621).

³ Although it may be considered as a reciprocal relationship, we prefer to consider that greater language proficiency may increase identification with the language. This is supported, for example, by the socio-contextual model of second language learning, which shows that there is an association between confidence in one’s abilities in a second language and identification with speakers of that language (Rubinfeld et al., 2006: 610-612).

The key variable of the analysis is use of Catalan, which is measured using a spoken language use index (IUP).⁴ The variable measures respondents' use of Catalan on an index of 0 to 1 in ten domains of language use:

A. Use outside the home

1. Use of Catalan with friends
2. Use of Catalan with neighbours
3. Use of Catalan with fellow students for people who are currently studying
4. Use of Catalan with workmates (with all people worked with, that is to say, now and in the past)
5. Use of Catalan in shops
6. Use of Catalan in shopping centres
7. Use of Catalan in banks
8. Use of Catalan with doctors

B. Use with family different from the family's linguistic origin; which may occur at home or outside the home

9. Use of Catalan with the partner
10. Use of Catalan with one's own children

The IUP results from an average for use of Catalan in all domains analysed based on the numeric value assigned to each level of use in Table 1, where Lx in this study stands for Catalan.

Table 1. IUP of Lx in a given domain and value assigned to each response

Response in questionnaire	Value
Only in Lx	1.00
More in Lx than in Ly	0.75
Same in Lx as in Ly	0.50
Both in Lx and in Ly and in Lz	0.33
Above all, in languages other than Lx	0.25
Only in Ly or combinations excluding Lx	0.00

For the first part of the analysis, we incorporated various sociolinguistic and sociodemographic factors as independent variables in a multiple linear regression analysis (Domínguez & Simó, 2003; Wooldridge, 2015). The results of this initial operation, which is covered exhaustively in the (unpublished) paper entitled *Juventut i ús del català a Catalunya*, drawn up by the authors at the CUSC–Research Centre for Sociolinguistics and Communication of the University of Barcelona with the support of the Directorate General for Language Policy of the Government of Catalonia,⁵ indicated that sociolinguistic variables are more closely associated with IUP than sociodemographic variables. Moreover, it showed that the three variables with most explanatory power were the language of identification, oral proficiency in Catalan, and first language, which are the variables focused on in this study. The report also included an analysis of sociolinguistic clusters and variables such as young peoples' families' geographic origin or their place of residence.

The variables measured on an interval scale were directly incorporated into the models, while in the general models created initially the categorical variables were treated as different dichotomous (dummy) variables,

⁴ For a more detailed description of the construction of the variable, see Sorolla & Vila (2018), which describes the development of the index, called *IUPI0* due to the 10 variables on which it was based.

⁵ Readers may contact the authors of this article in order to gain access to the full report.

with each category of the variable treated as a different variable in the regression model. Knowledge of Catalan was based on scalar variables—resulting from self-evaluation of the four different skills (oral comprehension, speaking, reading and writing) by the participants on a scale of 0 to 10—and not on dichotomous variables (yes or no questions on the four skills).⁶

3.2 Analytical techniques

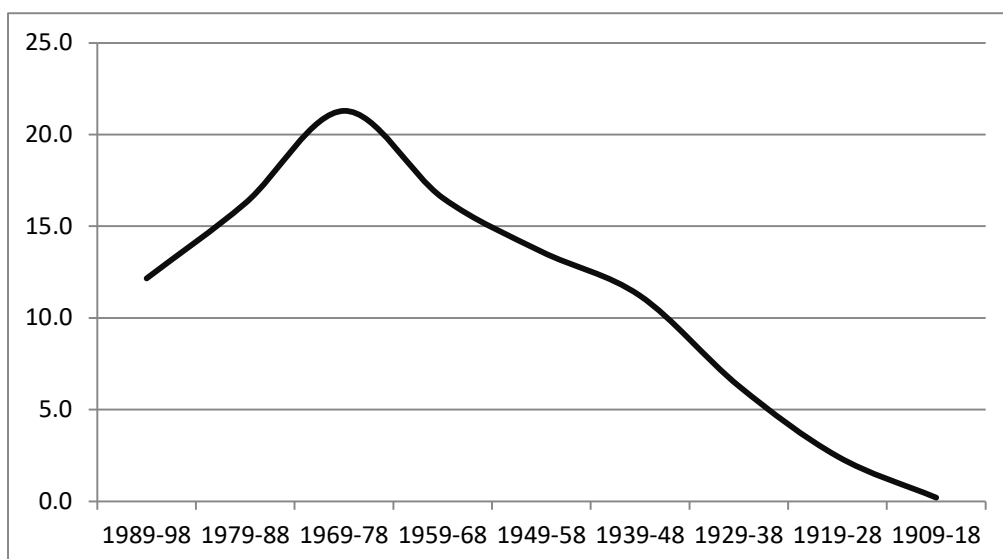
The basic analytical technique employed for this analysis is multiple regression analysis. This multivariate technique makes it possible to measure the association between a dependent variable—in this study, the IUP, the spoken language use index—and multiple independent variables, while controlling the effect on each another. Moreover, it also makes it possible to optimise models that, by selecting certain independent variables, are designed to improve the adjustment on the predicted values for the dependent variable with the minimum number of variables possible. This predictive ability is calculated by measuring the adjusted r squared value. And it uses a stepwise method that starts with the independent variable most closely associated with the dependent variable and then incorporates new variables step by step according to their capacity to improve the model. Priority is given to models that have the better predictive capacity and fewer variables in the model. The technique also makes it possible to measure and compare the intensity of the relationship between each of the independent variables and the dependent variable in each model using standardised beta coefficients. Lastly, a path analysis model was created, and employed to describe the relationships and dependencies between a set of variables. This type of analysis compares the theoretical proposal regarding how a set of variables is related, as mentioned above, with empirical evidence from different multiple regressions (Ho, 2006; Rodríguez & C. Bosch, unpublished).

4 Results

When analysing data on language use of young cohorts (15 to 34 years) and assessing their weight in the Catalan population as a whole, it is important to remember that these groups represent a decline in the population pyramid. Graphic 2 demonstrates the volume of each age band according to the decade in which they were born using data from EULP13. The youngest groups analysed in this report were born between 1979 and 1998 and were between 15 and 34 years old at the time of the survey.

⁶ The Catalan use and language proficiency variables were taken to be interval variables, as is habitual with scale variables ordered in multiple categories. The first language and language of identification variables were treated as interval variables, despite being divided into three ordered categories. There is a certain debate over whether it is more appropriate to treat them as dichotomous (dummy) or interval variables (Boyle, 1970), but, regardless of whether the two variables are measured as ordinals or as dummies, or whether the associations between the four variables are measured as ordinals or intervals, the order of the correlations remains consistent and the only variations are in magnitude: in all cases, language of identification is the variable most closely associated with use, followed by first language and lastly, language proficiency. At the same time, in all cases, proficiency is less closely associated with language of identification, and is particularly distant from first language. Of the six possible associations between the four variables, only one of the associations presents an important variation due to the statistical assumptions made. This is the association between first language and language of identification. When both variables are measured as ordinals (Catalan, Catalan and another language, or one other) this is the most important association of all six, whilst if they are treated as dummy variables (with two categories: Catalan and Not Catalan), the relationship reduces (only) the importance of the relationship between the two. This change is logical if one takes into account that a large proportion of the information in the two cases is lost if they are treated as dummy variables, whereas if they are treated with the three ordered categories, a wealth of nuances is maintained. The results are consistent both when treating the relationship between the four variables with ordinal statistics (Kendall tau-b or Spearman rho) and when assuming that the separation between the three categories is the same interval, and measuring the association with the Spearman correlation. The change only affects this relationship between the two variables mentioned, which is not a central association in the research question. Due to the consistency of the results, regardless of the statistical assumption made, the study assumes that first language and language of identification are measured with three interval categories and not as dummy variables (Boyle, 1970). For practical purposes, the associations between the two ordinal variables are measured as if they were interval categories.

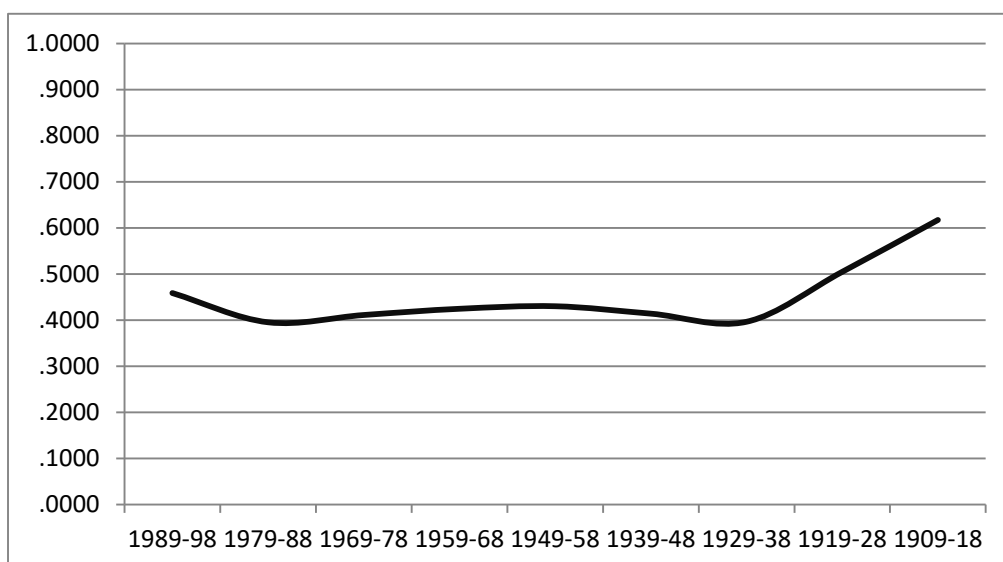
Graphic 2. Population pyramid, 10 year age bands, Catalonia, 2013 (EULP13). In percentage



4.1 Use index

Looking for a moment at the population of Catalonia as a whole, Graphic 3 shows that the spoken language use index (IUP) of Catalan is fairly similar in all the cohorts born from the 1930s onwards, at around 0.4. It should be mentioned that IUP is expressed on a scale between 0, indicating no use of Catalan in the 10 contexts analysed, and 1, indicating an exclusive use of Catalan in all the contexts analysed. 0.4 therefore indicates that Catalan is largely used by a minority in all age cohorts. Setting aside the cohorts born before 1929, the graph represents a proportionally small increase in use of Catalan in the youngest generation surveyed, those born between 1989 and 1998. This is possibly related to the fact that this is a cohort that has not yet incorporated the population born abroad; this tends to happen from the age of 18 onwards.

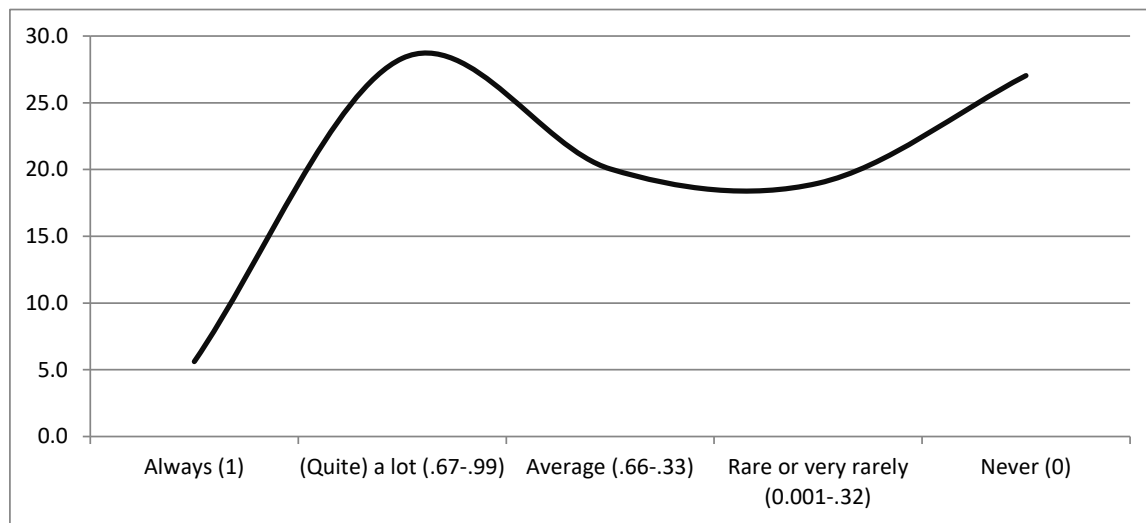
Graphic 3. Spoken language use index (IUP) as a function of age. Catalonia, 2013 (EULP13)



With regard to the IUP among young people, Graphic 4 indicates that there are two peaks in the distribution: groups of young people who use Catalan a lot or quite a lot—IUP between 0.67 and 0.99—and those who never use it—an IUP of 0. These groups each represent between 25% and 30% of all Catalan young people. Those who use Catalan an average amount or rarely—IUP from 0.33 to 0.66 or from 0.001 to 0.32,

respectively—each represent around 20% of the population as a whole. The smallest group is the group that speaks Catalan at all times, which is scarcely 5% of the total.

Graphic 4. Distribution of IUP (Spoken Language use index). Catalonia, young people, 2013 (EULP13). In percentage



4.2 Which variables are associated with use of Catalan by young people in Catalonia?

The following sections set out the results of an analysis which uses different regression models to establish the variables most commonly associated with Catalan use by young people in Catalonia. The results are presented in summary form and come from previous and wider levels of analysis that incorporated the effect of other sociolinguistic and sociodemographic variables on the spoken language use index (IUP) by young people, as has been indicated in the methods section above. The regression models mentioned took into consideration only sociolinguistic factors on some occasions, exclusively sociodemographic factors on others, both sets of factors on yet other occasions, and, lastly, segregated analyses according to families' geographic origin. In all cases, the analyses mentioned indicated that there are three key variables involved in the use of Catalan by young people in Catalonia: first language, oral proficiency in Catalan, and language of identification. These four variables, if one includes the dependent variable, coincide with the four key theoretical dimensions proposed: the first language as the language learnt during the process of primary socialisation; oral proficiency in Catalan as a result of primary and secondary socialisation—in the case of native speakers of Catalan or native bilinguals—or only of secondary socialisation in the rest of cases; the language of identification as a result of language subjectification; and lastly, the IUP as a result of language practices of the subjects.

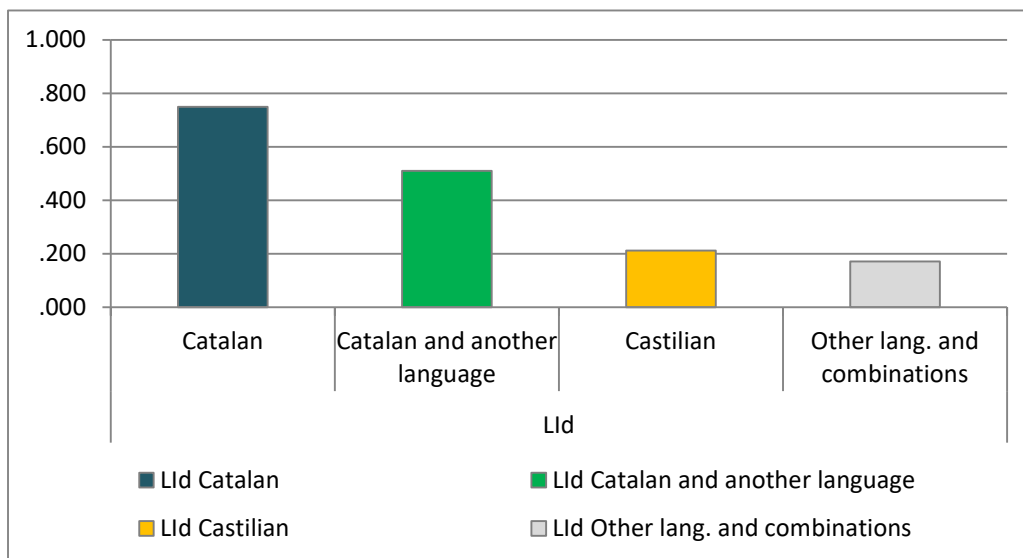
For this reason, we first measure the relationship between these four variables, and then we go on to analyse each of the four relationships in more depth. Table 2 demonstrates that Catalan use is closely associated with language of identification (0.77), first language (0.71) and oral proficiency (0.63). Language of identification has a very consistent association with the first language of the subjects surveyed (0.82) whilst, by contrast, oral proficiency is less closely associated with language of identification (0.48), and even less with first language (0.43).

Table 2. Bivariant correlation (Pearson r) between Catalan use (IUP), first language, language of identification and oral proficiency. Catalonia, young people, 2013 (EULP13).

	First language	Language of identification	Oral proficiency in Catalan
Spoken language use index	-0.71	-0.77	0.63
First language	–	0.82	-0.43
Language of identification	–	–	-0.48

The variable that is most closely associated with young people’s IUP is therefore the language of identification. This variable is more closely associated with use of Catalan than any other sociolinguistic or sociodemographic variable tested for in the wider analyses. In fact, in the regression models this variable alone is able to adjust predictions of Catalan use by 59.5%. Graphic 5 verifies this relationship, as those who select Catalan as a language of identification obtain an IUP of 0.75, and those who select Catalan and another language achieve a rate of 0.51, whilst those who select Castilian or other languages and combinations of languages as language of identification achieve a considerably lower rate, with rates of 0.21 and 0.17, respectively.⁷

Graphic 5. Spoken language use index (IUP) according to language of identification (LId) Catalonia, young people, 2013 (EULP13).

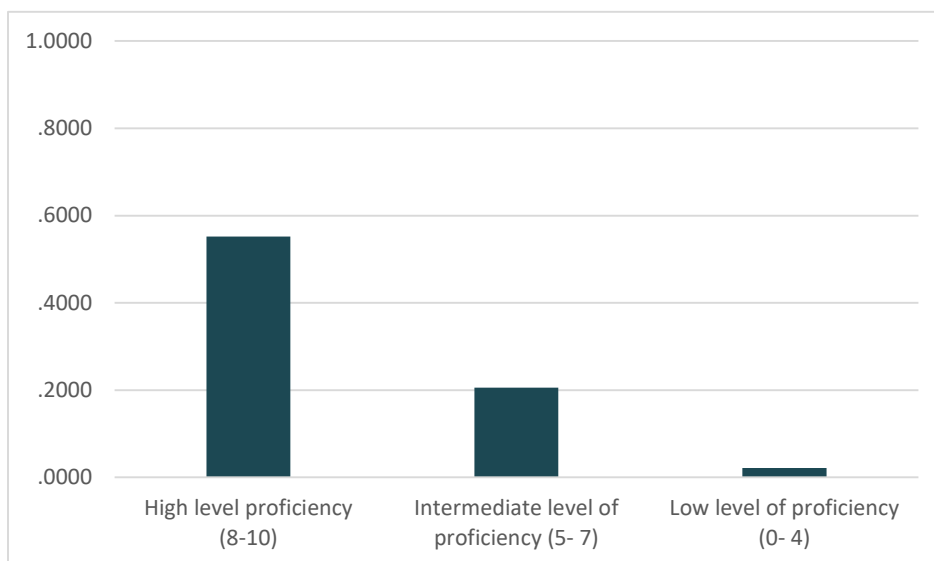


After language of identification, the variable that comes second in its impact on the use of Catalan by young people is declared oral proficiency, measured on a scale of 0 to 10, where 0 indicates nil proficiency and 10, full proficiency. The incorporation of this variable shifts the predicted values up 8.3 percentage points, to 67.8%.⁸ As can be seen in Graphic 6, young people who say they have high oral proficiency—between 8 and 10—obtain a spoken language use index rate of 0.55, whilst those who consider that they have average proficiency—from 5 to 7—or low proficiency—from 0 to 4—obtain a much lower IUP rate, of 0.21 and 0.02, respectively.

⁷ The first language and language of identification variables were treated in the regressions and correlation in an ordinal format, with three ordered categories of reply (Catalan, Catalan and another language, one other and other combinations). Nevertheless, in the wider analysis shown in Graphic 5 and those following, in which the variables were cross-referenced in order to bring out the nuances, the categorical format was applied, with four categories of non-ordered reply: Catalan, Catalan and another language, Castilian, and other languages and combinations.

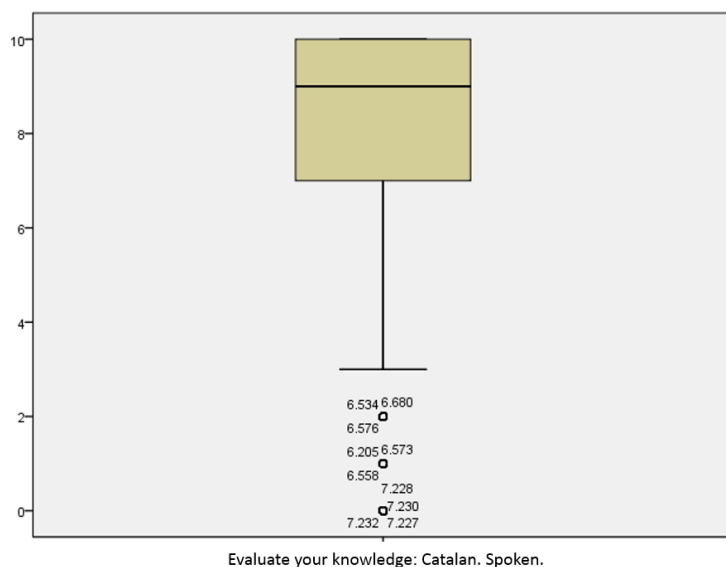
⁸ Oral proficiency in Catalan alone achieved a 39.2% adjustment in the prediction of the values of the use of Catalan.

Graphic 6. Spoken language use index (IUP) according to oral proficiency in Catalan (grouped scalar analysis). Catalonia, young people, 2013 (EULP13).



In order to put the analysis of the relationship between this variable and the IUP into perspective, it should be said that the majority of young people in Catalonia state they have high standards of active oral proficiency. Graphic 7 shows that on a scale of 0 to 10, half of the cohort state their oral proficiency is 9 or 10, and three quarters, 7 or above. Nevertheless, it should also be noted that up to a quarter are situated in fairly wide range, varying between 3 and 7.

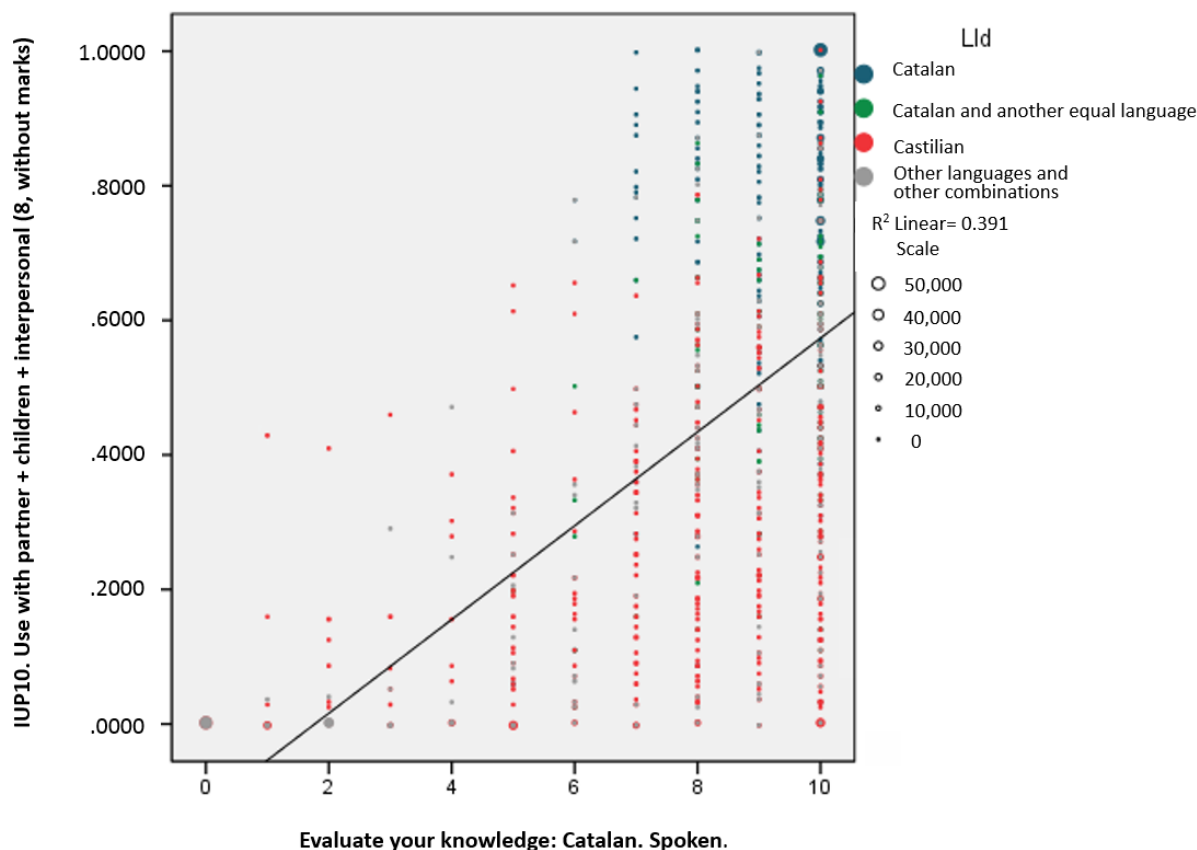
Graphic 7. Oral proficiency in Catalan (scalar survey) in a box plot, in which the median, the quartiles and the extreme cases are represented. Catalonia, young people, 2013 (EULP13).



To understand the relationship between language of identification, oral proficiency in Catalan (ability to speak), and use of Catalan, Graphic 8 combines the three variables, with knowledge of Catalan on the horizontal axis, use of Catalan on the vertical axis, and language of identification expressed in the colour of the points. The number of points, which represents the number of respondents involved in each situation, shows that the poorer the knowledge of Catalan (left-hand side), the less frequent use of Catalan (bottom); these two positions are occupied by young people whose language of identification is Castilian (red) or other languages (grey). 5.1% of young people state they can't speak Catalan at all, and therefore don't use it at all (IUP of 0). The language of identification of exactly half of these young people is Castilian and the other half identify with other languages or combinations with other languages. By contrast, as knowledge of Catalan

increases so too does the range of language use: the section where knowledge of Catalan is good (right-hand side) includes people who use Catalan in almost all contexts (top right) with others who hardly use Catalan in any contexts (bottom right). If one looks more closely at the space where competence in Catalan is high (right-hand side) one can see that Catalan is predominantly used (top) by young people whose language of identification is either Catalan alone (blue) or Catalan and another language (green). By contrast, among those who say that although their Catalan is good, they use it in few domains (bottom), one finds young speakers who identify with Castilian (red) or with other languages (grey).

Graphic 8. Spoken language use index (IUP) (vertical axis) according to scalar knowledge of Catalan (horizontal axis) and language of identification (LId) of the respondent (colour of points). Catalonia, young people, 2013 (EULP13).



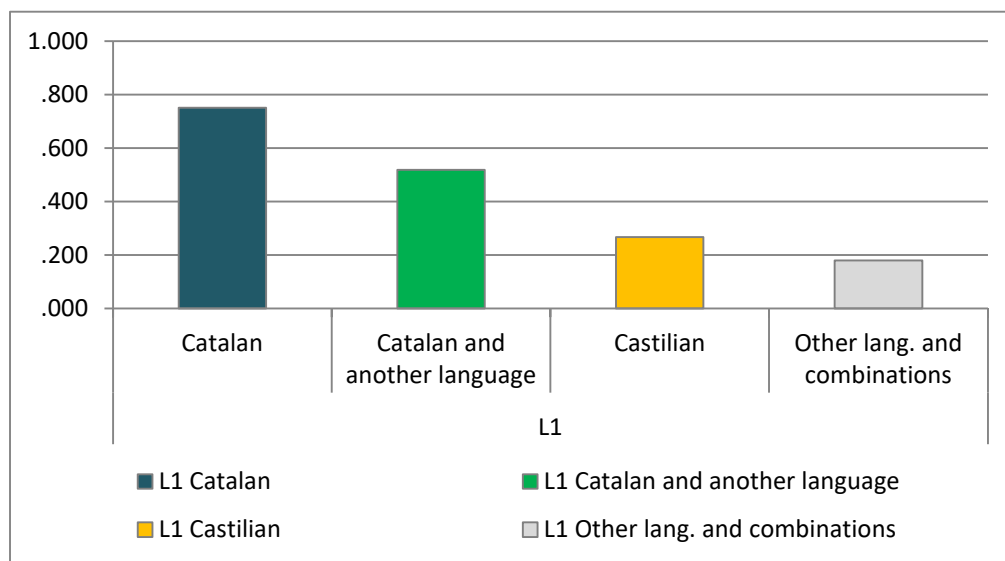
All the above indicates that the combination of both variables—the language of identification and active oral proficiency (ability to speak Catalan)—best explains the use of Catalan, which is low in those with a low level of competence, who also coincide with the group whose language of identification is not Catalan, but, by contrast, is more varied in the group that has a high level of competence, a segment in which the use of Catalan still tends to be low among those who don't include Catalan as one of, or as their exclusive language(s) of identification.

Thirdly, and lastly, the regression model suggests that it is important to incorporate the first language.⁹ In this case, the adjustment of the predicted values only improves 2.7 percentage points compared to the model that includes only language of identification and oral proficiency of Catalan.¹⁰ Nevertheless, the first language is highly important in this model. Graphic 9 demonstrates that it follows a similar dynamic to the language of identification: those who use Catalan most are above all native speakers of Catalan, and, to a less extent, those who learnt both Catalan and another language as first languages.

⁹ It should be noted that the variable for current use of language with parents achieved a significantly better model, with an increase of 1.5% over first language. Despite this, the first language model was maintained as the difference is not very great, and, above all, as this second variable is a more accurate reflection of primary socialisation. On the relationship between these two variables, see Tenorio (2013).

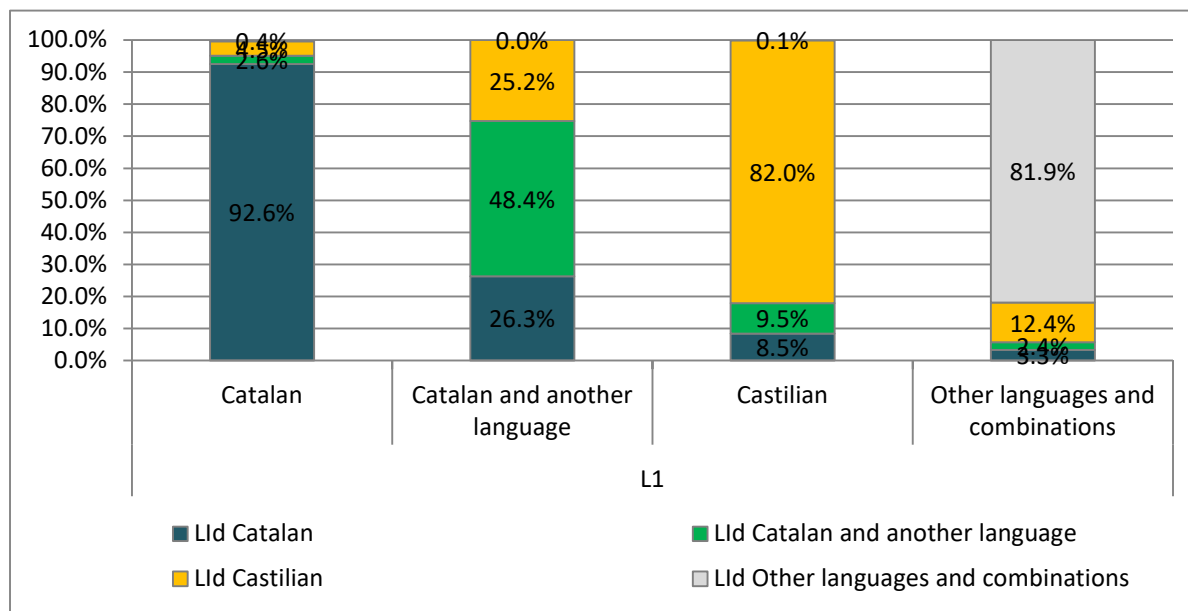
¹⁰ Catalan as first language achieves an adjustment in the prediction of values of 49.7%.

Graphic 9. Spoken language use index (IUP) according to first language (L1). Catalonia, young people, 2013 (EULP13).



In fact, there is a close association between first language and language of identification, as is shown in Graphic 10,¹¹ as the majority of young people declare that their first language is their language of identification. Nevertheless, there are certain important incongruities. Firstly, half of native bilinguals identify principally with only one of the declared first languages, and these are spread equally between those who go on to identify exclusively with Catalan (26.3%) and Castilian (25.2%). Secondly, a fifth of native speakers of Castilian identify with Catalan, either exclusively (8.5%) or together with Castilian (9.5%). Lastly, there are a large number of native alloglots who declare Castilian as language of identification (12.4%).

Graph 10. Language of identification (LId) according to first language (L1). Catalonia, young people. 2013 (EULP13).



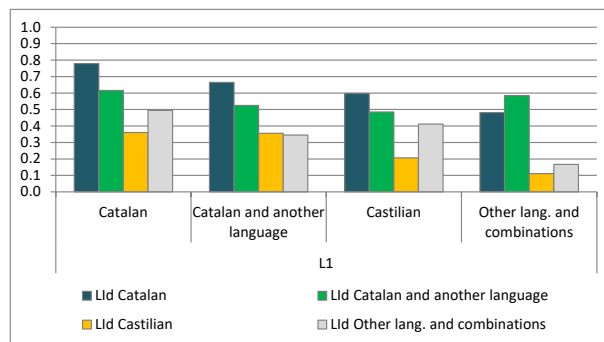
We have seen above that the spoken language use index is more closely associated with the language of identification than with the first language. The greater weight of the language of identification can be verified in Graphics 11 and 12, which display the spoken language use index both according to language of identification and the first language of the respondents. Graph 11 shows frequency of use of Catalan according to first language, and then subdivided into groups of first language depending on the declared language

¹¹ The Cramér's V between the two variables is 0.715, which highlights the close association between the two variables.

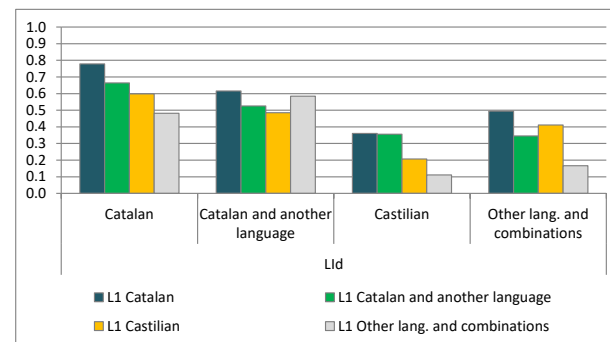
of identification. Graphic 12 shows the same data on the frequency of Catalan use, but then according to language of identification, subdivided according to first language.

These two arrangements show why Catalan use is more closely associated with the language of identification than the first language. For example, if one looks at those who state that their first language is Catalan and another language—the second block in Graphic 11—we can see that despite them all having a similar linguistic origin, their use of Catalan varies considerably when it comes to the range of languages they identify with. This diversity is greater than that seen in the second block in Graphic 12, the IUP of those who identify with Catalan and another language depending on their first language. There is greater homogeneity in use of Catalan in this case, despite the diversity of linguistic origin. In fact, the greater consistency in use of Catalan based on language of identification rather than first language is repeated in the rest of the groups. Thus one can confirm, even if only intuitively, that language of identification is more closely connected to language use in young people.

Graphic 11. Spoken language use index (IUP) according to first language (bars grouped by language of identification). Catalonia, young people, 2013 (EULP13).



Graphic 12. Spoken language use index (IUP) according to language of identification (bars grouped by first language). Catalonia, young people, 2013 (EULP13).



4.3 How are these variables associated with language use?

Finally, we present the results of a path analysis that incorporates the four variables dealt with up to now. Table 2 illustrated the association among the four variables. Apart from revealing the use of Catalan associated with language of identification, oral proficiency and first language, it also shows how these three *independent* variables are associated with each other. Path analysis enables one to verify empirically the dependency relationships among these variables hypothesised in the theoretical proposal. Graphic 1 showed the hypothesis regarding these relationships, which suggested that the three variables were associated directly with use of Catalan and at the same time, first language preceded both the language of identification and language competence, and that there may also be a certain association between increased proficiency and identification with the language.

As we noted in the previous section, the multiple linear regression model that incorporates the three independent variables—first language, knowledge of Catalan, and language of identification—explains 69.1% of variation in the use of Catalan by young people in Catalonia. If one analyses the stages¹² involved, this model shows that in the first instance, the language of identification alone would explain 59.5% of the variation in the use of Catalan. When proficiency is added, the improvement is 8.3%, increasing to 67.8%. When the last variable, the first language, is added in, the explanatory power of the model is increased by just 1.2%.

Pathway analysis shows firstly that, when the group effect of the three variables directly associated with use of Catalan is controlled, the one with the closest correlation is language of identification, with a standardised

¹² The variables are entered in a manner which shows how the model is improved by incorporating the variables stepwise.

coefficient of -0.459.¹³ This coefficient is higher than for knowledge of Catalan, which is 0.319. And finally, first language is relegated to a secondary position with a coefficient of -0.194. This can be observed in the beta values for the sub-model 1 in Table 3. Therefore, we can see that in the direct relationship, identification with the language and even oral proficiency in Catalan are more important in explaining the level of Catalan use than linguistic origin (first language).

In the second submodel, the two variables preceding language of identification, first language and oral proficiency, are analysed. In this case, one can see that with regard to subjectivisation (linguistic identification), the effect of primary socialization (first language) is much more important than language knowledge, since the language of identification dependency coefficient is clearly higher for first language (0.753) than for language competence (-0.155).

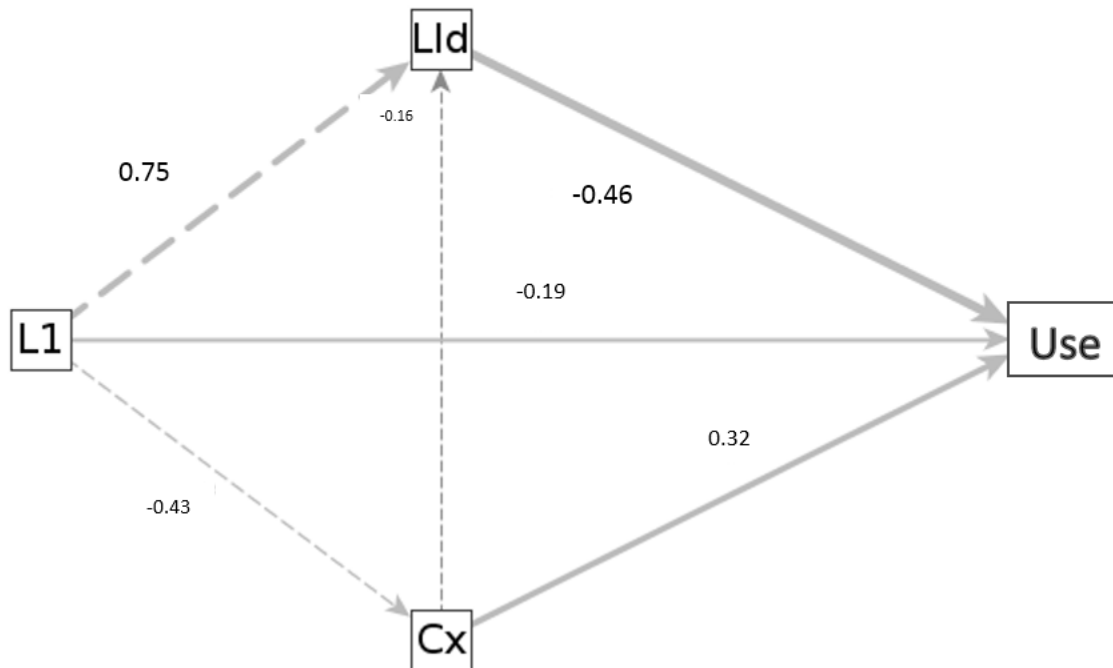
Finally, the third submodel analyses the dependence of language competence on first language and it can be seen that the association is fairly low ($r^2 = 0.188$), with a coefficient of 0.434. All these coefficients are modelled in Graphic 13.

Table 3. Multiple regressions that conform each of the three implicit submodels involved in the path analysis

Dependent variable	Indep. variable	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B		Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Tolerance	VIF
Submodel 1-VD: Use of Catalan (IUP) ($r^2=.691$)	(Constant)	.664	.021		31.191	.000	.622	.706		
	LId	-.163	.008	-.459	-20.608	.000	-.179	-.148	.308	3.247
	Knowledge.	.036	.002	.319	22.572	.000	.033	.039	.761	1.314
Submodel 2-VD: LId ($r^2=.692$)	(Constant)	.837	.057		14.753	.000	.726	.948		
	L1	.758	.014	.753	54.925	0.000	.731	.785	.809	1.236
	Knowledge	-.049	.004	-.155	-11.300	.000	-.057	-.040	.809	1.236
Submodel 3-VD: Knowledge. ($r^2=.188$)	(Constant)	11.06	.159		69.463	0.000	10.75	11.38		
	L1	-1.394	.064	-.434	-21.755	.000	-1.520	-1.269	1.000	1.000

¹³ When interpreting the path analysis results one must take into account that the (positive or negative) sign is not very important, and is due to the way the categories of each variable are ordered. Catalan use and knowledge are ordered in such a way that higher variable values are associated with greater use or higher proficiency of this language. By contrast, first language and language of identification are ordered in such a way that code 1 is the category *Catalan*, code 2 is *Catalan and other languages*, and code 3 is the category *other languages and other combinations*. Therefore, the ordering is the opposite of the other two values: for higher values, less use of Catalan. It is for this reason that the negative sign in the association between language of identification and language use, for example, refers to the fact that when the code of one increases, the other decreases, and therefore, for less identification with Catalan (increase of codes), use of the language goes down (decrease of codes).

Graphic 13. Path analysis of Catalan use (Use) according to first language (L1), language of identification (LId), and oral proficiency in Catalan (Cx). Empirical evidence.



5 Conclusions

The EULP13 data indicates that the spoken language use index (IUP) of Catalan millennials is fairly similar to the rest of cohorts born from the 1930s onwards, and is around 0.4 on a scale of 0 to 1, in which 0 represents no use of the language in any of the contexts selected and 1 represents the exclusive use of Catalan in all the contexts. There is only greater use in the youngest generation surveyed, those born in the 1990s, a trend linked to the limited presence of persons born abroad in this cohort compared to that of young adults and adults. Against this background, regression models were carried out to establish which sociodemographic and sociolinguistic variables associated with Catalan use by young people. The analysis indicated that sociolinguistic variables were those most closely correlated with language use and that, by contrast, sociodemographic variables had less explanatory power. Though sociodemographic variables may be very important from a sociological point of view, this led us to focus on the sociolinguistic variables in this study, for which path analysis was employed.

Without losing sight of the possible limitations of this initial multivariate approach to language use, whose results must be interpreted with a degree of caution, the analysis we have presented shows that the language of identification is the variable most closely associated with use of Catalan by young people in Catalonia. It is important to point out that this is a variable related to attitudes and sociolinguistic representations, and, while being closely related to first language, may also be a matter of choice, unlike geographic or linguistic origin, which are structural and outside the agency of the individual and the processes of internalization and subjectivisation described by Berger & Luckmann. The development of a certain affective connection with Catalan seems to be an important prerequisite for the consolidation of extensive use of Catalan by Catalan millennials, though one cannot go so far as to claim causality—that is to say, to claim that language identification is the cause of Catalan use—but rather that the two phenomena co-occur and reinforce each other: the more frequent the use, the greater the identification, and vice-versa. This reality is followed by the influence of oral proficiency in Catalan and speakers' first language. The results show that among young people in Catalonia identification, attitudes, values, and even willpower, are slightly more important than other structural variables, such as first language, in explaining language use.¹⁴

¹⁴ This theoretical proposal presents a number of opportunities for future research, including, firstly, that of carrying out the same

As highlighted above, the association between first language and language of identification is strong, despite important differences between the two variables. Therefore we stress that the language of identification cannot be *isolated* from the first language, nor, therefore, from linguistic origin. Nevertheless, considering them separately helps to provide a clearer explanation for young people's language use in Catalonia and to establish that linguistic origin has less weight as an explanatory factor in Catalan millennials language use. The same is true of language knowledge, which is now widespread in Generation Y as a result of the Catalan conjunction model of education. The tendencies observed in this demolinguistic analysis add to the evidence provided by studies that have suggested a certain de-polarisation and de-ethnicisation in language use by young Catalans and to a corresponding change in public discourse on Catalan language and identity, where there has been a rise in arguments that stress the elective nature of Catalan identity as a *project identity*, that, in principle, any person can adopt, regardless of their origin (Woolard, 2016).

The continuity of a language education model that guarantees a command of Catalan and Castilian is a key ingredient in this commitment to blurring the bound between language use and language origin and to the reduction of sociolinguistic inequalities based on origin. Nevertheless, the results of our research also indicate that, while it is true that extensive use of Catalan is only possible when there is a high level of language proficiency, it is not merely a matter of increasing language knowledge among young people, due to the strong link between language use and the language of identification.

Integrative motivations and the idea that Catalan is the language needed for full integration in the country seem still to be key reasons for adopting extensive use of Catalan. Current language policy and the current state of the language market in Catalonia, in which Catalan does not have the integrative/discriminatory capacity of established languages, to use Lamuela's terminology (1994, 2004), seems to prevent the consolidation of an important segment of "unidentified users", that is to say, people who use Catalan extensively because of instrumental/pragmatic motivations but do not identify actively with the language. The choice faced by planners therefore, is between adopting measures to make Catalan a necessary tool for accessing more important material and symbolic resources and for full participation in Catalan society—if the aim is to consolidate the group of unidentified users—or, on the other hand, to increase the importance of the use of Catalan in discourses on the definition of Catalan integration and identity—if the aim is to consolidate the group of identified users. Intervention in both areas, however, is clearly also an option.

Reference list

- Arnau, Joaquim, & Vila, F. Xavier. (2013). Language-in-education policies in the Catalan language area. In Joaquim Arnau (Ed.), *Reviving Catalan at school. Challenges and instructional approaches* (pp. 1-28). Bristol: Multilingual Matters.
- Berger, Peter L., & Luckmann, Thomas. (1966). *The social construction of reality. A treatise in the sociology of knowledge*. London: Penguin Books.
- Boix-Fuster, Emili. (1993). *Triar no és trair. Identitat i llengua en els joves de Barcelona*. Barcelona: Edicions 62.
- Boyle, Richard P. (1970). Path analysis and ordinal data. *American Journal of Sociology*, 75(4), 461-480.
- Direcció General de Política Lingüística (Coord.), Bretxa, Vanessa, Farré, Mireia, Martínez Melo, Montserrat, Masats, Marta, Navarro, Montse, Oller, Eulàlia, Rosselló, Carles de, Rovira, Cristina, Sorolla, Natxo, Torres-Pla, Joaquim, Ubalde, Josep, & Vila, F. Xavier. (2018). [Anàlisi de l'Enquesta d'usos lingüístics de la població a Catalunya 2013. Volum 1: Coneixements, usos, transmissió i actituds lingüístics](#). Barcelona: Generalitat de Catalunya, Departament de Cultura.

analysis on the whole Catalan population, in order to test the relationship between the variables considered and language use in different age groups, and to test the extent to which the tendencies described are exclusive to the young people or are common to other age groups. Secondly, and more ambitiously, it could be extended to different language use surveys in Catalan language territories and their regions, making it possible to study the bound between language use and the variables studied with the different language policies and social dynamics affecting Catalan in the different contexts.

- Domínguez, Màrius, & Simó, Montse. (2003). *Tècniques d'investigació social quantitatives*. Barcelona: Edicions Universitat de Barcelona.
- Flors-Mas, Avel·lí. (2016). Joves i llengües a Catalunya. Un estat de la qüestió. *Linguapax Review, 2015. The role of youth in language revitalisation / El paper dels joves en la revitalització lingüística*, 11-29.
- Flors-Mas, Avel·lí. (2017). *Usos lingüístics i identitats socials entre adolescents catalans i valencians* (Doctoral thesis, University of Barcelona).
- Flors-Mas, Avel·lí, & Vila, F. Xavier. (2014). [Justificant les preferències. Com argumenten les opcions lingüístiques els adolescents catalans](#). *Treballs de Sociolingüística Catalana, 24*, 173-199.
- Garrett, Paul B., & Baquedano-López, Patricia. (2002). [Language socialization: reproduction and continuity, transformation and change](#). *Annual Review of Anthropology, 31*, 339-361.
- Generalitat de Catalunya. (2015). [Enquesta d'usos lingüístics de la població 2013](#). Generalitat de Catalunya.
- González, Isaac, Pujolar, Joan, Font, Anna, & Martínez, Roger. (2014). [Llengua i joves. Usos i percepcions lingüístics de la joventut catalana](#). Barcelona: Generalitat de Catalunya, Departament de Benestar Social i Família, Direcció General de Joventut.
- Ho, Robert. (2006). *Handbook of univariate and multivariate Data analysis and interpretation with SPSS*. Boca Raton (USA): Chapman & Hall/CRC Press.
- Howe, Neil, & Strauss, William. (2009). *Millennials rising: the next great generation*. New York: Vintage Books.
- Lamuela, Xavier. (1994). *Estandardització i establiment de les llengües*. Barcelona: Edicions 62.
- Lamuela, Xavier. (2004). [Instal·lació o establiment? Encara sobre els objectius de la promoció lingüística](#). *Caplletra. Revista Internacional de Filologia, 37*, 215-242.
- Milian, Antoni. (1984). [De la separació a la conjunció lingüística a l'ensenyament: El títol II de la Llei 7/1983, de 18 d'abril](#). *Revista de Llengua i Dret, 3*, 33-41.
- Newman, Michael, & Trenchs-Parera, Mireia. (2015). Language policies, ideologies and attitudes in Catalonia. Part 1: reversing language shift in the twentieth century. *Language and Linguistics Compass, 9*(7), 285-294.
- Newman, Michael, Trenchs-Parera, Mireia, & Ng, Shukhan. (2008). Normalizing bilingualism: The effects of the Catalan linguistic normalization policy one generation after. *Journal of Sociolinguistics, 12*(3), 306-333.
- Pujolar, Joan. (2008). [Els joves, les llengües i les identitats](#). *Noves SL. Revista de Sociolingüística, Hivern 2008*, 1-11.
- Pujolar, Joan, & González, Isaac. (2013). Linguistic «mudes» and the de-ethnicization of language choice in Catalonia. *International Journal of Bilingual Education and Bilingualism, 16*(2), 138-152.
- Rodríguez, Jose Antonio, & C. Bosch, Jose Luis. (unpublished). *Manual de anàlisis causal*.
- Rubinfeld, Sara, Clément, Richard, Lussier, Denise, Lebrun, Monique, & Auger, Réjean. (2006). Second language learning and cultural representations: beyond competence and identity. *Language Learning, 56*(4), 609-631.
- Schieffelin, Bambi B., & Ochs, Elinor. (1986). [Language socialization](#). *Annual Review of Anthropology, 15*, 163-191.

- Sorolla, Natxo. (2012). [Estabilitat i canvi de la llengua inicial en els usos lingüístics interpersonals. Una anàlisi comparativa del País Valencià, la Franja, les Illes Balears i Catalunya](#). In F. Xavier Vila (Ed.), *Posar-hi la base: Usos i aprenentatges lingüístics en el domini català* (pp. 91-108). Barcelona: IEC.
- Sorolla, Natxo, & Vila, F. Xavier. (2018). Els grups segons els usos lingüístics i l'evolució de l'ús del català entre 2003 i 2013. In Direcció General de Política Lingüística (Coord.), [Anàlisi de L'Enquesta d'usos lingüístics de la població a Catalunya 2013. Volum 1: Coneixements, usos, transmissió i actituds lingüístics](#). Barcelona: Generalitat de Catalunya, Departament de Cultura.
- Tenorio, Xavier. (2013). [Llengua inicial i llengua amb progenitors. Una mateixa realitat o dues variables diferents?](#) *Treballs de Sociolingüística Catalana*, 23, 407-426.
- Trenchs-Parera, Mireia, & Newman, Michael. (2015). Language policies, ideologies, and attitudes. Part 2: International immigration, globalization and the future of Catalan. *Language and Linguistics Compass*, 9(12), 491-501.
- Vila, F. Xavier. (2000). El moment d'adquisició del català i les seves repercussions sociolingüístiques. In Joaquim Noguero (Ed.), *L'ús del català entre els joves. Cicle del 21 al 30 de setembre de 1999* (pp. 53-61). Sabadell: Caixa de Sabadell.
- Vila, F. Xavier. (2005). [Language, education and ideology in an integrationist society](#). In Els Witte et al. (Ed.), *Language, attitudes & education in multilingual cities* (pp. 53-86). Brussels: Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten.
- Woolard, Kathryn A. (1989). *Double talk: bilingualism and the politics of ethnicity in Catalonia*. Stanford: Stanford University Press.
- Woolard, Kathryn A. (2016). *Singular and plural. Ideologies of linguistic authority in 21st Century Catalonia*. Oxford: Oxford University Press.
- Wooldridge, Jeffrey M. (2015). *Introductory Econometrics: a modern approach*. Melbourne: South-Western Cengage Learning.