

Relations between formal linguistic insecurity and the perception of linguistic insecurity. A quantitative study on linguistic insecurity in an educational environment at the Valencian Community (Spain)

Josep M. Baldaquí Escandell

Department of Catalan Studies / IIFV, University of Alicante, Spain

Ap. Correus 99, E-03080. Alacant (Spain)

(Received XX Month Year; final version received XX Month Year)

What is the relationship between the awareness of linguistic prestige and the security or insecurity in the use of minoritized languages? Is formal linguistic insecurity (as initially described by Labov) the same as the speakers' perception of linguistic insecurity? Which are the variables related to the various types of linguistic insecurity in bilingual environments? How does schooling affect this phenomenon? This paper attempts to provide an answer to these questions, based on the results of a research project on linguistic insecurity (LI) carried out in a social context, the Valencian Autonomous Community (Spain), where there are two official languages (Spanish and Catalan) competing against each other, but with a very different degree of social recognition. The choice of an academic environment for this study is justified by the importance that the LI phenomenon may have within the context of language learning in multilingual contexts.¹

Keywords: linguistic insecurity; Catalan language; linguistic prestige

Introduction

Since Labov (1966) formulated the initial definition of the notion of linguistic insecurity (LI), the studies on this construct have expanded and modified the notion, which is nowadays much more complex than Labov's original proposal. According to Labov (1966), LI is a measurement of the speaker's perception of the prestige of certain linguistic forms, compared to the ones the speaker remembers he or she normally uses. Labov's index of linguistic insecurity, or ILI (Labov 2006, 319, 421-22) is, therefore, a quantitative measure of the percentage of coincidences or discrepancies between forms that speakers recognize as more adequate, that is, more prestigious, and those which they recognize as part of their own idiosyncratic usage. It must be emphasized that this

description approaches LI as an intralinguistic phenomenon, since it compares linguistic variants belonging to different sociolects within the same language.

Within Labov's variationist model, LI appears as one of the causes of language change, because of the hypercorrection mechanisms it originates (Labov 2006, 318). The most insecure social groups regarding usage would be those with a greater sensitivity towards prestigious linguistic forms, who desire to rise within the social scale, especially the low-middle class and females.

Although the term "linguistic insecurity" may be felt as somewhat inadequate in order to refer to a process of evaluation of linguistic prestige, it would be justified by the consequences it has among speakers. Thus, hypercorrection, doubt, nervousness, self-correction, erroneous perception of one's own speech pattern, or an important fluctuation between different speech styles have been associated to the language usage of insecure individuals (Labov 2006, 322-23).

The original concept of LI suffered no important modification until the 1990's. It was then that French-speaking sociolinguistics made a sizeable contribution to the study of LI, replacing the original quantitative approach by a qualitative methodology. The new focus was on the description of speakers' representations of LI, based on the analysis of their own epilinguistic discourse.

One of the main promoters of this new approach was Michel Francard, who strived to lay the foundations for a new theoretical LI framework. According to Francard, the state of insecurity is seen in representations as those he describes for Belgium's French-speaking community: a) subjection to exogenous linguistic model, resulting in cultural and linguistic dependence on France; b) depreciation of one's own language practices and variety; c) ambivalence of linguistic representations, leading speakers to resort to compensation strategies, such as attributing qualities to their native

variety (affectiveness, complicity, warmth, coexistence...) which are denied to the dominant variety; d) experts' pessimism towards the future of French, a feeling of threat expressed especially concerning the role of French within the world language market, completely taken over by the English language (Francard 1993a, 63-68; Francard 1993b, 14-17).

It is under this new perspective that Francard (1997, 171-72) defines LI as

la manifestation d'une quête de légitimité linguistique, vécue par un groupe social dominé, qui a une perception aiguisée tout à la fois des formes linguistiques qui attestent sa minorisation et des formes linguistiques à acquérir pour progresser dans la hiérarchie sociale.

It was also Francard (1989; 1993a) who described the relationship between the degree of schooling and the degree of LI, emphasizing the role of schools as LI generators: "il n'est pas arbitraire d'attribuer à l'institution scolaire un rôle essentiel dans l'émergence d'attitudes d'insécurité linguistique" (Francard 1993a, 40). Indeed, in present times schools are the main institutions disseminating prestigious social norms regarding language usage. Therefore, the knowledge of the prestigious norm is directly related to the degree of schooling, and this knowledge allows speakers to be aware of the distance between their speech and the prestigious model. The paradoxical consequence is that speakers most familiar with the language norm are those who, at the same time, show a lower degree of confidence, that is, a greater insecurity regarding language usage: "les sujets les plus scolarisés ont les appréciations les plus négatives concernant les usages linguistiques" (Francard 1989, 151).

Years later, also in the 1990's, the studies on LI were expanded to multilingual environments, and the initial intralinguistic perspective became an interlinguistic one, including language contact situations (Bretegnier 1996; Calvet 1996; de Robillard 1996). As a result, there was a proposal to include within the notion of linguistic insecurity, issues such as the status of languages in contact, or the relations between

languages and individual and group identities, within the social and language dynamics of language contact situations (Calvet 1999; 2006).

Thus, Calvet described three different types of LI, which can appear separately or together: a) formal or Labovian insecurity, resulting from speakers' perception of the distance between their native language uses and those they consider most prestigious; b) statutory insecurity, the consequence of speakers' negative evaluation of the status of the language they use, compared to that of another language or variety; and c) identity insecurity, which takes place when speakers use a language or a linguistic variety different from that used by the community they identify themselves with and are members of (Calvet 2006, 133-45). As one can observe, the first type of LI (formal insecurity) is an intralinguistic phenomenon, between social varieties within the same language, whereas statutory or identity LI are, basically, interlinguistic phenomena taking place between clearly differentiated languages or linguistic varieties as perceived by speakers.

The context of this research

The research presented here has been carried out in the Autonomous Community of Valencia (Spain).² As is well known, this area has with two official languages, albeit with a different status: Spanish, the official language of the whole of Spain, is the dominant language, and is almost the only language present in most of the media, especially the most formal ones. The Catalan language, which in this area is given the official name of "Valencian", presents a great degree of minoritization, with little importance in public and prestigious usage, because it has only reached a relatively important presence (although a moderate one) in compulsory educational levels. The sociolinguistic situation greatly differs from that in Catalonia, better known in the international sphere.

The sociolinguistic situation in the Valencian Community presents a number of factors likely to provoke a high degree of LI in the minoritized language, especially in educational environments: (a) the high degree of minoritization of Catalan in the Valencian community; (b) the weak perception by Valencian speakers of their own ethnolinguistic identity;³ c) the existence, in part of the Valencian society, of a feeling of linguistic secessionism of the Valencian variety from Catalan; d) the existence of various official bodies competing for the normativization of Catalan in the Valencian Community (IEC and AVL)⁴ and, therefore, of various competing formal language models, albeit with small differences; e) the LI affecting many teachers, due to their limited linguistic training in the minoritized language, an LI which is easily transmitted to the students; f) the heterogeneous linguistic competence of the students in Catalan, due mostly to their first language, to the language prevailing in the educational programme they have been schooled in (Catalan or Spanish), and to the language prevailing in the environment where they live (Baldaquí 2009, 186-95).

Methodological issues

The results presented here are part of a wider research project on LI developed in the whole of the Valencian community, following a quantitative method. The quantitative method has been chosen for two reasons: because it allows us to compare formal LI (studied since its origins with quantitative methods) and IL perception (traditionally approached through qualitative methods), and because it allows us to study the phenomenon in a wider context, the Valencian Community.

The universe is formed by young people in Valencian-speaking areas⁵ studying the last two years of compulsory education (3rd and 4th year of Compulsory Secondary Education), aged mainly between fourteen and sixteen. In order to design the sample, a multi-stage design has been used, organized in clusters. These ages were chosen because, since the students are following the last years of compulsory education, the sample potentially includes all the young people in this cohort.

The sample has been clustered proportionally according to the variables considered most important for the study: area demography, sociolinguistic region,⁶ school ownership (public / privately owned, state-funded or not)⁷ and curriculum type, i.e. Progressive Incorporation Programme (*Programa d'Incorporació Progressiva, PIP*) or Valencian Teaching Programme (*Programa d'Ensenyament en Valencià, PEV*).⁸ The elements in the sample (class groups) were selected by means of simple random sampling, and the questionnaires were handed out to all the students in the class groups selected.

The number of students surveyed at the end of the field research, carried out between October 2007 and June 2008, was 2190. Therefore, it can be said that the sample is representative of the universe studied (students in third and fourth year of compulsory secondary education in Valencian-speaking areas in the Autonomous Community of Valencia) with a reliability of 95% and an error margin which, in the worst-case scenario ($p = q = 0.5$), is lower than 0.03.

Data collection instruments

The data collection instruments used in order to assess LI were the following:

- (1) *Sociolinguistic questionnaire*, with 53 questions concerning students' sociolinguistic profile (age, sex, social class, personal linguistic history, self-evaluation of their own linguistic competence, self-evaluation of language use of official languages in various speech environments, self-ascription to one or another ethnolinguistic group, etc.).
- (2) *Language attitude test*, consisting of 20 questions (10 concerning Spanish and 10 concerning Valencian), aimed at building two Likert scales on respondents' attitude regarding each of the official languages.⁹ The scales built have a reliability of $\alpha = 0.884$ for the scale regarding Catalan and of $\alpha = 0.717$ for the scale regarding Spanish.
- (3) *Formal linguistic insecurity questionnaire*, following the model originally developed by Labov, although considerably expanded concerning the number of linguistic variables studied. The questionnaire includes 45 items regarding phonetic variables (10 items), morphosyntactic variables (20 items) and lexical variables (15 items), and assesses insecurity in formal environments: classroom and interaction with teachers. In order to answer, respondents heard the 45 items

on two different occasions, each item consisting of two sentences with only one difference (e.g. a given phonetic variant, or two synonyms belonging to different formal levels). The first time respondents hear each sentence pair, they must indicate which one they consider more adequate in formal contexts (in the classroom or with teachers). The second time, they hear them they must say which one they would normally use. The disagreement between the forms considered most adequate and those normally used is the General Index of Formal Linguistic Insecurity (GIFLI).

- (4) *Linguistic insecurity perception questionnaire*. This is a completely original test, aimed at assessing the perception and beliefs regarding linguistic insecurity. It consists of 42 double questions (regarding the two official languages, Valencian and Spanish), on the various issues considered part of the LI construct. Respondents have to answer using a semantic continuum represented by a line divided into 10 equal segments (Figure 1).

Figure 1. LI perception questionnaire.

| | | | | | | | | | | |
|--|---|---|---|---|------------------|---|---|---|---|----------------|
| 10) When I speak to teachers in the classroom, I feel: | | | | | | | | | | |
| VALENCIAN: | | | | | | | | | | |
| Great insecurity | | | | | Average security | | | | | Great security |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| SPANISH: | | | | | | | | | | |
| Great insecurity | | | | | Average security | | | | | Great security |
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |

The double questions aim to obtain a figure for LI perception in each language and, at the same time, to calculate the differential value of LI perception. Indeed, we believe that in bilingual environments, such as this one, LI is an absolute phenomenon (LI in each of the languages), but also a relative one (LI compared between the two languages).

The 42 items in the LI perception questionnaire have been formulated considering the issues which, in our opinion, could influence LI perception in the context studied, and also the theoretical contributions by previous researchers, especially Calvet (1999; 2006). The sections in which the questionnaire is divided are the following:

- a.1) Self-evaluation of security / insecurity in one's personal language practices (12 items).

- a.2) Self-evaluation of desired security / insecurity in one's personal language practices (4 items).
- b) Self-evaluation of one's own linguistic quality (4 items).
- c.1) Evaluation of general linguistic quality of the area (4 items).
- c.2) Evaluation of quality of the reference formal language models and identification with some of these models (4 items).
- d) Evaluation of difficulty entailed by the formal variety and by the written variety of the language (2 items).
- e.1) Evaluation of security / insecurity in linguistic status (5 items).
- e.2) Evaluation of desirable security / insecurity in linguistic status (3 items).
- f) Evaluation of role of the language in the configuration of Valencian or Spanish identity as a group, both at present and in the future (4 items).¹⁰

Dependent variables

This study has two main dependent variables:

- Formal linguistic insecurity, measured through the General Index of Formal Linguistic Insecurity (GIFLI), which is the percentage of insecurity (lack of agreement between the most prestigious variety and the variety used everyday) corresponding to the set of 45 variables in the Formal LI questionnaire.
- Perception of linguistic insecurity, measured through the General Index of Differential Linguistic Insecurity Perception (GIDLIP). This index corresponds to the figures in a Likert scale¹¹ developed from the differential results in the answers to the LI perception questionnaire. This scale includes 36 items from the 42 originally featured in the questionnaire, with a very high reliability ($\alpha = 0.964$).

Structure of the LI perception scale

A principal component analysis (with orthogonal varimax rotation) has been performed on the 36 items finally included in the Likert scale, in order to know the structure of the LI perception scale. The preliminary statistical results confirm that the data are fully adequate in order to perform a principal component analysis.¹²

An analysis of the results has shown that the principal component analysis test fits a three-factor solution fairly well, since the factors altogether explain 59.55% of the variance in the LI perception construct.¹³ The interpretation of the factors from the items with the highest factor loads is the following:

The first factor includes most of the items in blocks *a.1* (self-evaluation of security / insecurity in one's personal language practices), *b* (self-evaluation of one's own linguistic quality) and *d* (evaluation of difficulty entailed by the formal variety and by the written variety of the language). We have interpreted it as "evaluation of security or insecurity in personal language practices, one's own linguistic quality and perception of the difficulty of the language". This factor appears to have some connection to formal security, because it includes the perception of one's own linguistic quality, but also items concerning the self-evaluation of one's language competence. As will be seen below, it is difficult to consider this factor as equivalent to the LI notion as developed by Labov (1966), which Calvet (1999) calls "formal linguistic insecurity".

The second factor includes most of the items in block *a.2* (self-evaluation of security / insecurity desired in one's personal language practices), *c.2* (evaluation of quality of the reference formal language models), *e.2* (evaluation of desirable security / insecurity in linguistic status) and *f* (evaluation of role of the language in configuration of Valencian or Spanish identity as a group, both at present and in the future), and also includes an item on the usefulness of the language. This may be interpreted as "evaluation of desirable security / insecurity, both in one's own language practices and in the status of the languages, and evaluation of role of the language in configuration of group identity". This is the factor most resembling identity security or insecurity (Calvet 1999), which in our scale is connected to the desire to improve one's LI.

Finally, factor 3 includes items from block *c.1* (evaluation of general linguistic quality of the area) and *e.1* (evaluation of security / insecurity in linguistic status). One possible interpretation is that this factor concerns the "evaluation of security / insecurity in the status of languages, including general linguistic quality of the area as an element

determining such status". The factor, therefore, would resemble status security or insecurity as proposed by Calvet (1999).

To sum up, it may be concluded that the LI perception scale developed here shows interesting coincidences with Calvet's theoretical LI construct, but also presents some divergences from Calvet's model.

Independent variables

A selection has been made of 21 independent variables obtained from the Sociolinguistic questionnaire and the Linguistic Attitudes Test, which may yield a thorough sociolinguistic description of the subjects:

- General sociolinguistic variables: 1) Sex; 2) Social class (calculated from parents' profession and educational level).
- Variables indicating sociolinguistic environment where respondents live: 3) language of environment (language perceived by respondents as predominant in the area where they live); 4) "Apitxat area" (respondent lives in a town where the apitxat variety is spoken, traditionally stigmatized as hardly correct).
- Variables concerning subjects' schooling: 5) Educational programme, with two possibilities: Valencian Teaching Programme (PEV), a bilingual programme with Catalan as the main teaching language, and Progressive Incorporation Programme (PIP), where teaching is mainly in Spanish, although Catalan is taught as an individual subject and may be the language in which some of the non-linguistic subjects are taught; 6) Ownership: school ownership: public or privately owned (state-funded or not).
- Variables related to subjects' linguistic history: 7) Place of birth (classified according to the language); 8) Years living in Valencian Community; 9) Mother tongue; 10) Languages spoken (number of languages spoken fluently, not including Catalan and Spanish).
- Variables related to subjects' linguistic competence in Catalan and Spanish: 11) Self-evaluation of oral comprehension in Catalan; 12) Self-evaluation of oral expression in Catalan; 13) Self-evaluation of reading comprehension in Catalan; 14) Self-evaluation of written expression in Catalan; 15) Self-evaluation of oral comprehension in Spanish; 16) Self-evaluation of oral expression in Spanish; 17) Self-evaluation of reading comprehension in Spanish; 18) Self-evaluation of written expression in Spanish.
- Variables regarding linguistic attitudes and identity: 19) Motivation towards learning Catalan (instrumental or integrative); 20) Differential attitude (index combining the two variables related to linguistic attitudes; calculated by subtracting, from the variable expressing the attitude towards Valencian, the value of the variable expressing the attitude towards Spanish); 21) Spanish identity subtractive index, SISI (index combining the two variables related to

Spanish and Valencian identity; calculated by subtracting the degree of Spanish identity from the degree of Valencian identity).

Discussion

Considering the purposes of this paper, an initial analysis is made of the relationship between formal LI and the perception of LI. This is followed by a study of the relations between the two types of LI and the independent variables. Finally, a detailed analysis is offered of the connections between language usage, the two types of LI and the other variables in the study.

Relations between formal LI and LI perception

The first issue to be analyzed is the relations between formal linguistic insecurity, measured through the General Index of Formal Linguistic Insecurity (GIFLI), and LI perception, measured through the General Index of Differential Linguistic Insecurity Perception (GIDLIP).

As seen in Table 1, the Pearson correlations between the two LI indexes are negative, low-intensity, and statistically significant ($r = -0.147$, $p = 0.000$). This suggests that the two types of insecurity actually correspond to different phenomena.

Table 1. Correlations.

Correlations

| | | GIFLI | GIDLIP | Fact-1 | Fact-2 | Fact-3 | Self-evaluation of Catalan usage degree | Self-evaluation of Spanish usage degree |
|---|---------------------|-----------|-----------|-----------|-----------|-----------|---|---|
| GIFLI (General index of formal linguistic insecurity) | Pearson Correlation | 1 | -,147(**) | -,117(**) | -,044 | -,093(**) | ,148(**) | -,135(**) |
| | Sig. (2-tailed) | | ,000 | ,000 | ,073 | ,000 | ,000 | ,000 |
| | N | 1893 | 1670 | 1670 | 1670 | 1670 | 1884 | 1884 |
| GIDLIP (General index of differential linguistic insecurity perception) | Pearson Correlation | -,147(**) | 1 | ,704(**) | ,569(**) | ,425(**) | -,328(**) | ,325(**) |
| | Sig. (2-tailed) | ,000 | | ,000 | ,000 | ,000 | ,000 | ,000 |
| | N | 1670 | 1887 | 1887 | 1887 | 1887 | 1878 | 1880 |
| Fact-1 | Pearson Correlation | -,117(**) | ,704(**) | 1 | ,000 | ,000 | -,261(**) | ,259(**) |
| | Sig. (2-tailed) | ,000 | ,000 | | 1,000 | 1,000 | ,000 | ,000 |
| | N | 1670 | 1887 | 1887 | 1887 | 1887 | 1878 | 1880 |
| Fact-2 | Pearson Correlation | -,044 | ,569(**) | ,000 | 1 | ,000 | -,096(**) | ,075(**) |
| | Sig. (2-tailed) | ,073 | ,000 | 1,000 | | 1,000 | ,000 | ,001 |
| | N | 1670 | 1887 | 1887 | 1887 | 1887 | 1878 | 1880 |
| Fact-3 | Pearson Correlation | -,093(**) | ,425(**) | ,000 | ,000 | 1 | -,205(**) | ,227(**) |
| | Sig. (2-tailed) | ,000 | ,000 | 1,000 | 1,000 | | ,000 | ,000 |
| | N | 1670 | 1887 | 1887 | 1887 | 1887 | 1878 | 1880 |
| Self-evaluation of Catalan usage degree | Pearson Correlation | ,148(**) | -,328(**) | -,261(**) | -,096(**) | -,205(**) | 1 | -,679(**) |
| | Sig. (2-tailed) | ,000 | ,000 | ,000 | ,000 | ,000 | | ,000 |
| | N | 1884 | 1878 | 1878 | 1878 | 1878 | 2179 | 2178 |
| Self-evaluation of Spanish usage degree | Pearson Correlation | -,135(**) | ,325(**) | ,259(**) | ,075(**) | ,227(**) | -,679(**) | 1 |
| | Sig. (2-tailed) | ,000 | ,000 | ,000 | ,001 | ,000 | ,000 | |
| | N | 1884 | 1880 | 1880 | 1880 | 1880 | 2178 | 2181 |

** Correlation is significant at the 0.01 level (2-tailed).

There are also low, negative correlations between formal LI and each of the factors described earlier from the principal component analysis of the General Index of Differential Linguistic Insecurity Perception (GIDLIP). More specifically, there is a negative, low-intensity correlation between the General Index of Formal Insecurity Perception (GIFLI) and Factor 1 ($r = -0.117$, $p = 0.000$), a factor including the self-evaluation of LI, of the linguistic quality of speakers and of the difficulty of the formal varieties of the language. The lack of intensity of this correlation confirms our previous statement that Factor 1 in our study does not correspond to the formal or Labovian LI notion. The complete absence of correlation between the three factors in the GIDLIP is a logical one, since they are the result of an orthogonal varimax rotation.

Finally, the correlations are analyzed between the two LI types and language usage. The correlation between formal LI (GIFLI) and Catalan usage is significant, though a positive, low-intensity one, whereas the correlation of the same index with Spanish usage is also a low, significant, positive one. Therefore, it may be concluded that the usage of the minoritized language is not negatively affected by formal LI because, in any case, there is a positive correlation between very low-intensity variables which cannot be interpreted as a cause-effect relationship.

The correlation between LI perception (GIDLIP) and language usage is more important, statistically significant, and also has an opposite sign in each of the languages. In this case it may be concluded that the use of the minoritized language is negatively related to LI perception in this language ($r = -0.328$, $p = 0.000$), although the covariance, a moderate one, cannot be interpreted either as a cause-effect relationship.

Analysis of formal LI

Given the high number of independent variables in this study, it was decided to explore the data through the segmentation analysis technique. This methodology will make it possible to find the independent variables most related to the dependent variables, and also to describe the relations between the dependent variables (the two types of LI) and the remaining variables in the study (predictor variables). In fact, an awareness of the factors with the greatest influence upon LI is a necessary condition for a full explanation of the phenomenon. Besides, from an applied point of view, identifying the groups of subjects most likely to experience formal LI (what in medical language would be called the risk groups) is a first step in the struggle against LI, especially when it appears in school environments as an unexpected result of language learning.

The methodology used for the segmentation analysis is the Exhaustive CHAID (Chi Automatic Interaction Detection) algorithm, which, unlike other segmentation methods, makes it possible to divide subjects into more than two groups simultaneously.¹⁴

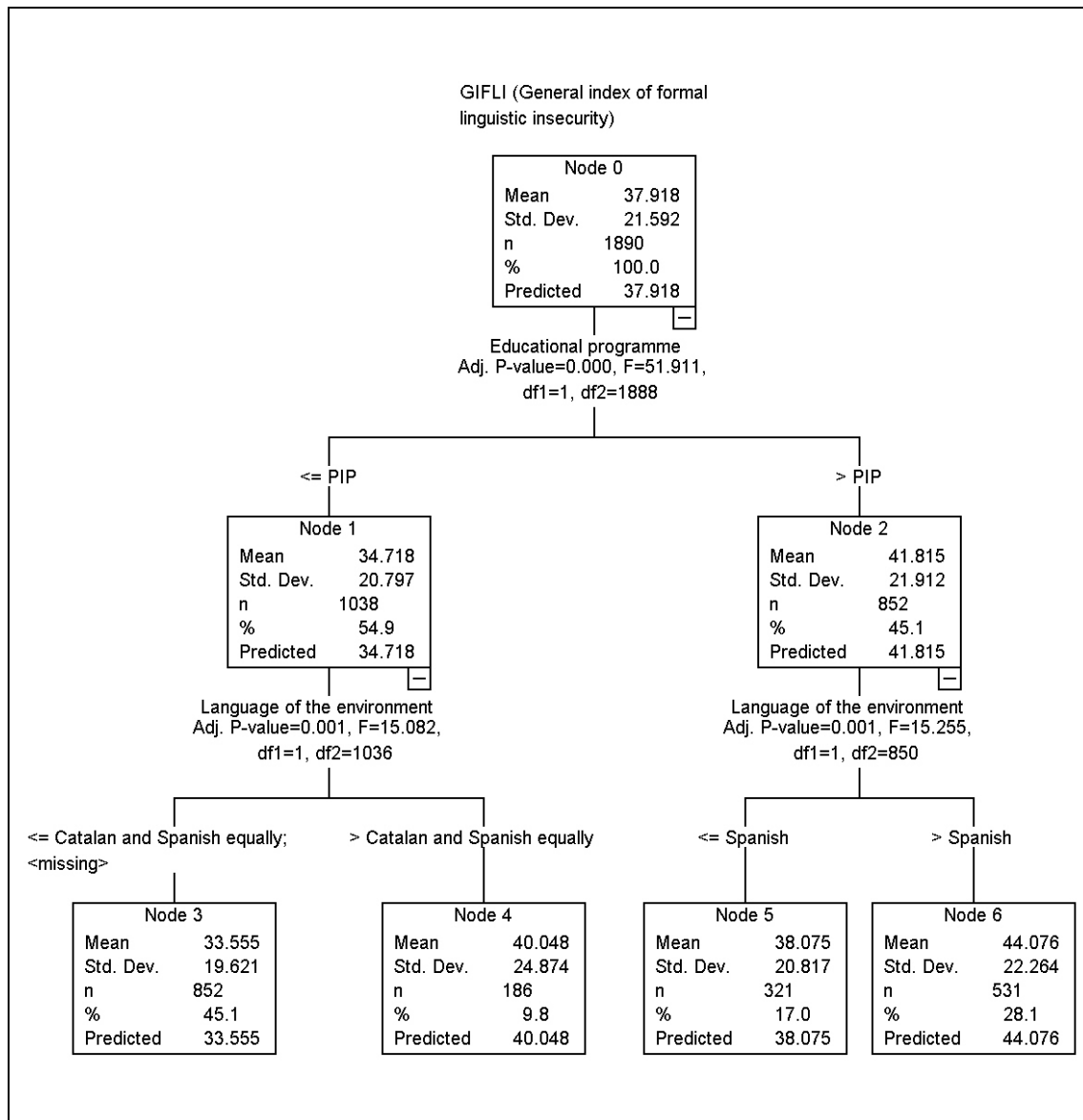
For the analysis of formal LI in the use of Catalan, in this study the General Index of Formal Linguistic Insecurity (GIFLI) will be the dependent variable, whereas the independent variables described earlier will be used as predictor variables. The variables appearing as the best predictors of formal LI are the following (Table 2):

Table 2. Best formal LI predictors.

| | | | | |
|----|-----------------------------|-------------|------------|-----------------|
| a. | Educational programme | F = 51.9959 | DF: 1.1891 | p = 0.000000000 |
| b. | Language of the environment | F = 23.6553 | DF: 2.1890 | p = 0.000000001 |
| c. | Mother tongue | F = 16.9450 | DF: 3.1889 | p = 0.000000002 |
| d. | Differential attitude | F = 19.8288 | DF: 1.1891 | p = 0.000053784 |
| e. | Motivation | F = 12.2353 | DF: 1.1891 | p = 0.002878674 |

The results of this first segmentation level (Figure 2) show that young people enrolled in the Valencian Teaching Programme (PEV), taught mainly in Catalan, experience a higher degree of formal LI when using Catalan ($\bar{X} = 41.815$) than students enrolled in the programme taught mainly in Spanish (PIP), whose average formal LI is lower ($\bar{X} = 34.718$). The result is consistent with the studies indicating that school is an institution creating LI. Perhaps the most interesting detail is that this phenomenon, which had already been described in qualitative studies on LI perception (Francard 1989, 1993a), is also found here in a quantitative study on formal or Labovian LI.

Figure 2. GIFLI segmentation tree.



The results of the second segmentation level indicate that, for both groups of young people, the perception of which language has the greater social presence in the environment in which they live, determines the degree of formal LI in the use of Catalan. Thus, considering the two educational programmes, the students with a higher formal LI are those perceiving a greater presence of Catalan in their environment. A description of the terminal nodes shows that the students with the highest formal LI are those enrolled in the programme taught in Catalan (PEV), who consider that in their district or town Catalan prevails, or Catalan and Spanish are equally used (node 6, \bar{X} =

44.076). Next, one can find students taught mainly in Spanish (PIP) perceiving that Catalan prevails in their environment (node 4, $\bar{x} = 40.048$), followed by students in the Valencian Teaching Programme who perceive that Spanish prevails in their environment (node 5, $\bar{x} = 38.075$). Finally, the students with a lower formal LI in Catalan are those schooled mainly in Spanish who perceive that Spanish is the language with the greater presence in their environment ($\bar{x} = 33.555$).

Our interpretation of these results is that formal LI in the use of Catalan appears especially through the confluence of two factors:

- The degree of knowledge of the formal variety of the minoritized language, Catalan, acquired mainly through schooling, and more specifically, by enrolling in the programme taught mainly in Catalan (PEV), as could be seen in the first segmentation level.
- The degree of social use of Catalan in one's immediate environment, which determines the degree of contact with the traditional/colloquial language model.

This is why youths with a greater formal LI are those with a high knowledge obtained at school of the formal language model, living in environments allowing great contact with the traditional/colloquial language model. Therefore, the opportunity to compare the two language models would lead to formal linguistic insecurity, which is after all a sign that the awareness among these young people of the linguistic prestige of the formal variety of Catalan, is greater than that among other youths.

It is interesting to mention the lack of predictor variables for some of the theme groups in which we classified them initially: those related to linguistic competence of young people and the sociolinguistic variables of a more general nature (sex and social class). It must be emphasized that the fact that they do not appear in the segmentation tree does not mean that they are not related to formal LI, but simply that such relation is not as strong as that of the variables analyzed here.

Analysis of LI perception

LI perception has been measured through the General Index of Linguistic Insecurity Perception (GIDLIP), the result of a Likert scale expressing the differential or relative LI perception between the two official languages. The scale may range from 10 (a maximum LI perception for Catalan and a minimum one for Spanish) to -10 (a maximum LI perception for Spanish and a minimum one for Catalan). The value 0 would indicate a balanced LI perception in the two languages.

The average value of this variable is $\bar{X} = 1.5642$, which indicates that, for the sample as a whole, LI perception is greater in the use of Catalan than in Spanish. This is an unexpected result, not regarding the direction of LI, logically higher in the minoritized language, but the intensity of the LI, which appears to be very low considering the degree of minoritization of Catalan in the Valencian Community.

A segmentation analysis has been used also for the analysis of formal LI perception. In this case, the dependent variable is the General Index of Differential Linguistic Insecurity Perception (GIDLIP), whereas the independent variables are the other variables in the study.

Table 3 shows the variables which may prove the best options for the first segmentation:

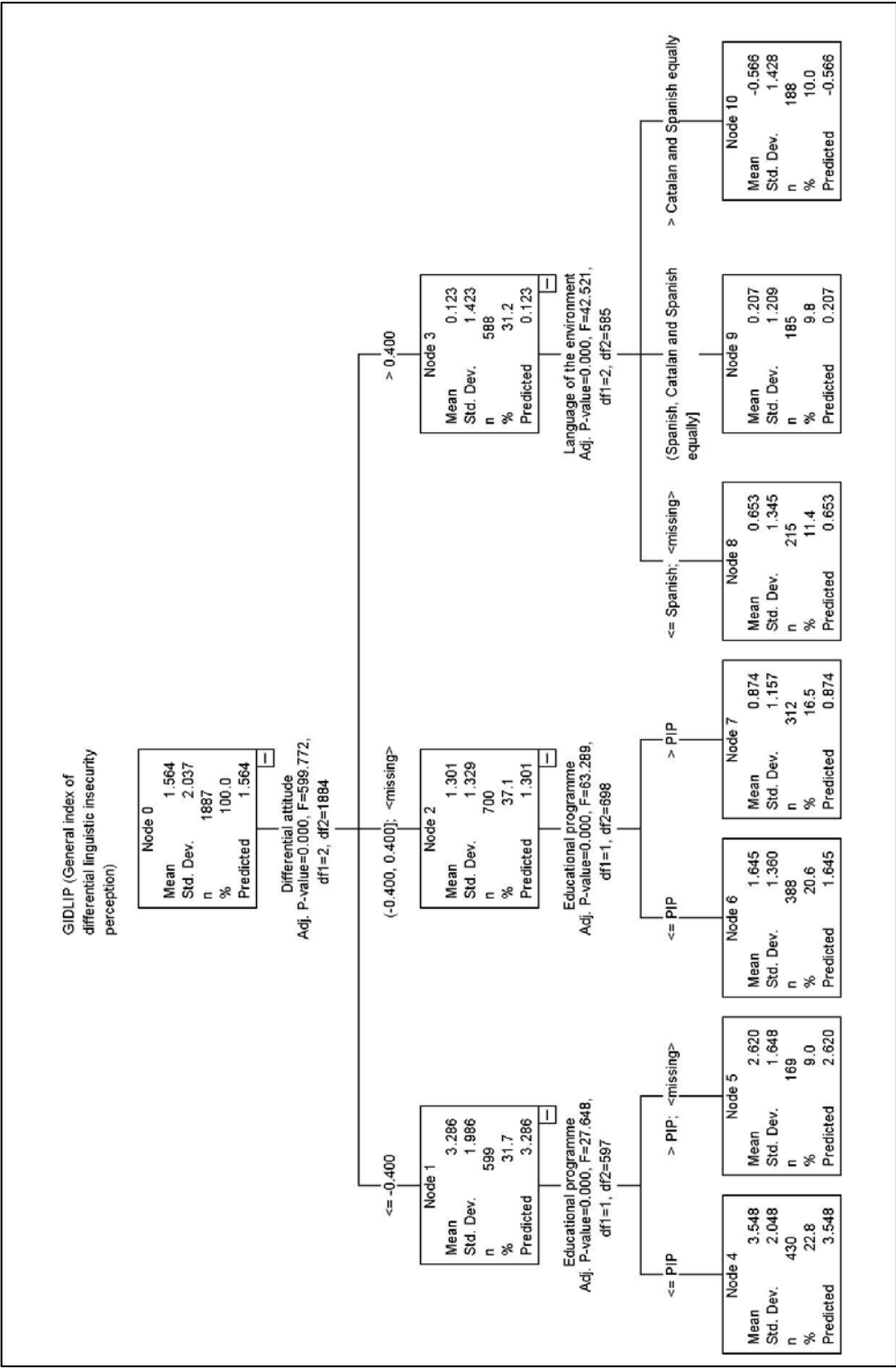
Table 3. Best predictors of perception of LI insecurity.

| | | | | |
|----|-----------------------------|--------------|------------|-----------------|
| a. | Differential attitude | F = 399.8933 | DF: 3.1883 | p = 0.000000000 |
| b. | Educational programme | F = 236.5666 | DF: 1.1885 | p = 0.000000000 |
| c. | Catalan: oral comprehension | F = 78.6257 | DF: 1.1885 | p = 0.000000000 |
| d. | Language of the environment | F = 72.9938 | DF: 2.1884 | p = 0.000000000 |
| e. | Spanish: oral expression | F = 64.7101 | DF: 1.1885 | p = 0.000000000 |

A first element to be pointed out is that LI perception is mainly influenced by an attitudinal variable. Positive values in the differential attitude variable indicate a

linguistic attitude in favour of Catalan, negative values a more favourable attitude towards Spanish, and 0 would represent a balanced attitude. Therefore, in the first segmentation level in the tree (Figure 3), it may be seen that young people with a smaller LI perception in Catalan are those with the most favourable attitude towards Catalan (node 3, $\bar{X} = 0.123$), whereas those with attitudes more favourable to Spanish express a higher perception of Catalan LI (node 1, $\bar{X} = 3.286$). The most balanced linguistic attitude values present GIDLIP values closer to the average in the sample (node 2, $\bar{X} = 1.301$).

Figure 3. GIDLIP segmentation tree.



In a second segmentation level two explanatory variables can be found, which indicates the interaction between the variables, since the most influential variable explaining LI perception is not the same one for each of the groups classified according

to their linguistic attitude. These are variables with different characteristics: one concerns the language of the environment where respondents live, whereas the other is related to the educational programme they have been exposed to.

In the case of young people positively inclined towards the minoritized language (> 0.400), the language they perceive as most widely used in their environment is the variable which best helps to explain LI in this segmentation level. Thus, an increased social presence of Catalan in their environment results in a lower LI perception in this language. On the contrary, in the case of young people with balanced linguistic attitudes ($-0.400, 0.400$) or attitudes more inclined towards the dominant language (≤ -0.400), the linguistic characterization of the educational programme where they have been enrolled is the most influential explanatory variable for the segmentation of this second level. Thus, for these students, an increased presence of Catalan in the school curriculum leads to a lower LI perception in this language.

An analysis of the terminal nodes shows a continuum, from a greater LI perception in Catalan among young people with a more favourable attitude to Spanish who have been also taught in Spanish (node 4, $\bar{X} = 3.548$), to a slightly higher LI perception in Spanish among students taught in Catalan who live in an environment where Catalan prevails (node 10, $\bar{X} = -0.566$).

LI as a predictor variable of the use of the minoritized language

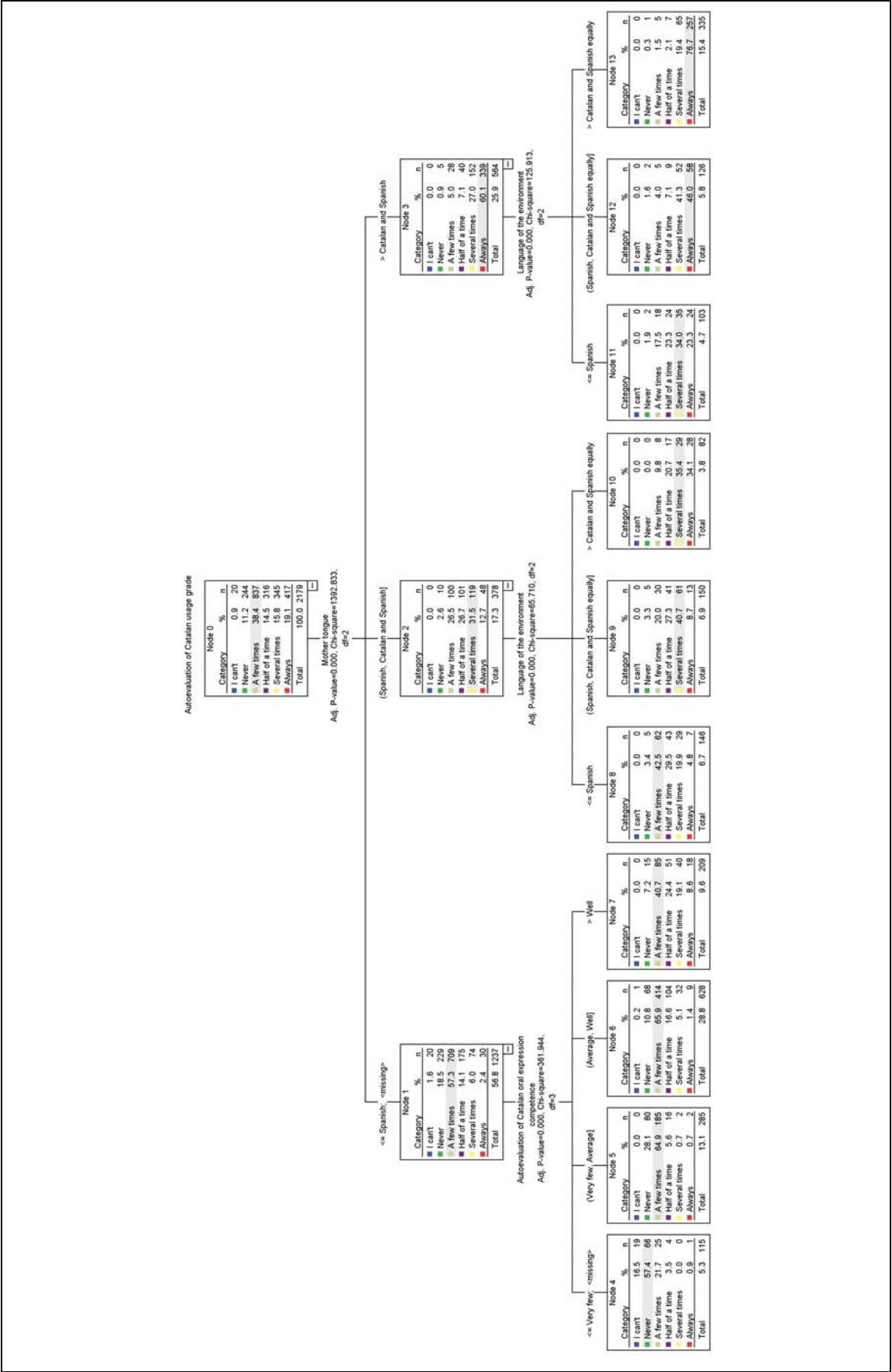
The last analysis carried out is aimed at deciding whether the variables indicating LI in Catalan are related to language use. Therefore, a final segmentation analysis is proposed, in order to describe the most influential variables explaining the use of the minoritized language. The analysis includes the two variables indicating LI (GIFLI and GIDLIP) within the group of the predictor variables.

As can be observed from Figure 4, LI does not appear under any shape (either formal LI –GIFLI–, or LI perception –GIDLIP–) as a predictor variable of language usage. In fact, the mother tongue is the variable which best helps to predict the use of Catalan: subjects who have learned Catalan as their mother tongue are those using it more frequently, followed by bilingual students who have learned Catalan in the family environment. This is a logical result, especially considering that, Catalan being a minoritized language in the Valencian Community, it enjoys little public formal use, and is therefore greatly restricted to family and colloquial environments.

In the second segmentation stage, it may be seen that, for those students whose mother tongue is Catalan, or who are family bilinguals, the variable best determining the use is the higher or lower presence of the language in their environment. As these young people usually speak Catalan, the greater the presence of the language in their environment, the greater the use. Conversely, when a young person has not learned Catalan in the family environment, the use of the language (which, in any case, is lower than in the other groups) is mostly related to the oral competence they have reached, usually after learning it at school.

An examination of the modal values in the terminal nodes shows that youths that state they always use Catalan are those who learned it as their mother tongue, and live in an environment where they perceive Catalan is used at least half of the time or more. This category also includes bilingual students who have learned Catalan and another language (other than Spanish) in their family environments.

Figure 4. Use of the Catalan language tree.



The respondents who use Catalan very frequently are those who have learned it as their mother tongue and live in contexts where it is little used, or Catalan-Spanish bilinguals who learned Catalan in their families, and live in environments where Catalan is used, at least, on equal footing with the dominant language.

The youths who use Catalan only sometimes are either young Catalan-Spanish bilinguals who learned Catalan within their families and live in environments where Spanish prevails, or students whose mother tongue is Spanish, or are family bilinguals in Spanish and another language and have acquired an average or good competence in Catalan. Finally, respondents whose mother tongue is Spanish or Spanish-other language family bilinguals, who have only managed to reach a low or non-existent oral linguistic competence in Catalan, are those who, quite expectedly, never use it.

Finally, and to sum up, the variables best explaining language usage in the minoritized language are mother tongue, language of the environment and oral competence, whereas the two variables expressing the various LI types are not among the most influential when it comes to explaining language usage.

Brief conclusions

This study has made it possible to describe two different aspects of LI, formal LI and LI perception, and the connections between the two types of LI in the context studied. It has also allowed us to make a contribution on the nature of the LI construct and on the relations between LI and other sociolinguistic variables, especially regarding language usage.

This study shows that, in the context studied, the two LI types are very little related, arise from very different factors, affect the various groups of youths in different degrees, and their consequences regarding language usage are also different.

It has been shown that formal LI, as initially proposed by Labov, is basically a measure of linguistic prestige. In bilingual environments, or at least in the environment studied here, the appearance of formal LI in the use of the minoritized language is mainly related to the degree of awareness of the formal, normativized uses of the language. This awareness is acquired by comparing the formal language model learned at school with traditional colloquial varieties of the minoritized language, learned in the family and in one's nearer language environment. The possibility of gaining access to these two language models will determine the degree of awareness of linguistic prestige and, therefore, the degree of formal LI. This is a type of LI affecting mostly young people who have learned the minoritized language in their families or in their nearer environment. The presence of this type of LI does not seem to be related to language usage, and should not be necessarily considered a negative event, since it seems desirable that there should be some degree of awareness of linguistic prestige, which can be, after all, a catalyst for language learning (Coste 2001).

However, LI perception in the minoritized language is negatively related to language use, although to a moderate extent. IL perception, as measured in our questionnaire, may be considered as a variable self-assessing the speakers' willingness to use a language, and although it is a complex construct, statistical analysis shows that it is mainly related to the speakers' attitude towards the minoritized language, and to the availability of sufficient linguistic competence to use such language. Such IL affects especially youths who have only learnt the language at school, and have not reached sufficient competence to feel comfortable when using it. Thus, this LI type fits much better within the notion of "linguistic insecurity" than the formal or Labovian LI, which rather corresponds to the idea of linguistic prestige.

The results obtained, both in the segmentation studies and in the principal component analysis of the LI perception scale developed here, contribute towards a better understanding of the LI construct and, especially, of the important conceptual differences between formal LI and LI perception. Finally, it is also our hope that this paper has made some interesting methodological contributions towards the quantitative study of LI.

Notes

1. Project “Linguistic insecurity in the use of the minoritized language in the Valencian educational system: analysis and improvement proposals” (HUM2006-10229), approved by the Spanish Ministry of Education and Science within the National Plan for Scientific Research, Development and Technological Innovation 2004-2007, jointly sponsored through ERDF funding. Thanks are due to Dr. Miguel Ángel Campos Pardillos, responsible for the English version of this paper.
2. Although Spain is not a federal state, autonomous communities in Spain enjoy great autonomy, especially regarding education, culture, health or public administration.
3. Only 26% of Valencians consider that speaking their native language is a distinctive trait of the Valencian identity (Siguan 1999).
4. Traditionally, the Philological Section of the Institute of Catalan Studies (*Institut d’Estudis Catalans*, IEC), located in Barcelona, has been the institution having competence on the normativization of Catalan in the whole linguistic territory. In 1998 the Valencian Academy of the Language (*Acadèmia Valenciana de la Llengua*, AVL) was created as an official institution for the Valencian area. Although this academy implicitly recognizes the unity of the Catalan language, it also considers itself as a normative body at the same level as the IEC, which has provoked tension between the two institutions.
5. The Valencian Language Teaching and Use Act 4/1983 (*Llei d’Ús i Ensenyament del Valencià*) divides the Valencian Community into areas of Valencian linguistic predominance and areas of Spanish linguistic predominance, depending on whether the historically dominant language has been Catalan or Spanish.
6. The area studied has been divided into five sociolinguistic regions, depending on whether there is a higher or lower social presence of Catalan.
7. Some private schools, though privately owned, are state-funded.
8. The main difference between the two educational programmes is the main language used for learning: in the case of the Progressive Incorporation Programme, the main language is Spanish, although some non-linguistic modules must be taught in Catalan. In the case of the Valencian Teaching Programme, the main learning language is Catalan. This latter programme may be considered as a fully bilingual educational programme, whereas the Progressive Incorporation Programme is bilingual to a lower extent, and shows great variation from one school to another.
9. This test has been adapted from that originally designed by Sharp (Great Britain Schools Council and Sharp 1973), and has been frequently used in research within the Catalan language context. The statistical software used in the construction of Likert scales, the correlation studies and the principal component analysis was SPSS v. 15.
10. The full version of the *Linguistic insecurity perception questionnaire* may be consulted at <http://hdl.handle.net/10045/16889>
11. In order to develop the scale, we have considered the statistical analysis of the integration of each item in the scale and the value of Cronbach’s alpha, if any of the items has been excluded.
12. The value of the matrix determinant is very low (2.61 E-013), and an analysis of the matrix shows that all the correlations are statistically significant ($p < 0.001$). Bartlett’s statistical test for sphericity shows a very high value (54270.33) and is statistically significant ($p < 0.001$); the

- anti-image correlation matrix shows that all the values are low, close to 0, and Kaiser-Meyer-Olkin's measure of sampling adequacy is highly adequate (0.969). Finally, all the coefficients in the Measures of Sampling Adequacy Sample have values over 0.9 (the lowest one is 0.919).
13. The first factor explains 29.73%, the second one 18.08%, and the third one 11.73%.
 14. We set the segmentation in two levels and the significance filters, in order to allow group division or category grouping, at $p = 0.01$ and $p = 0.05$ respectively. The statistical software used in the segmentation analysis was AnswerTree v. 3.1.

References

- Baldaquí, J. M. 2009. La inseguretat lingüística i l'aprenentatge de les llengües minoritzades. Reflexions des del País Valencià. *Revista Catalana de Pedagogia* 6: 177-197.
- Bretegnier, A. 1996. L'insécurité linguistique: objet insécurisé? Essai de synthèse et perspectives. In *Le Français dans l'espace francophone. Description linguistique et sociolinguistique de la francophonie*, ed. D. de Robillard and M. Beniamino, 903-19. Paris: Honoré Champion éditeur.
- Calvet, L.-J. 1996. Les 'Edwiniens' et leur langue: sentiments et attitudes linguistiques dans une communauté créolophone. *Revue québécoise de linguistique théorique et appliquée* 13: 9-50.
- Calvet, L.-J. 1999. *Pour une écologie des langues du monde*. Paris: Plon.
- Calvet, L.-J. 2006. *Towards an Ecology of World Languages*. Cambridge: Polity.
- Coste, D. 2001. Compétence bi/plurilingue et (in)sécurité linguistique: L'École Valdôtaine. <http://www.scuole.vda.it/Ecole/Atti/05.htm> (accessed July 8, 2010).
- De Robillard, D. 1996. Le concept d'insécurité linguistique: à la recherche d'un mode d'emploi. In *Français régionaux et insécurité linguistique*, ed. C. Bavoux, 55-76. Paris: L'Harmattan.
- Francard, M. 1989. Insécurité linguistique en situation de diglossie. Le cas de l'Ardenne belge. *Revue québécoise de linguistique théorique et appliquée* 8, no. 2: 133-63.
- Francard, M. 1993a. Trop proches pour ne pas être différents. Profils de l'insécurité linguistique dans la Communauté française de Belgique. *Cahiers de l'Institut de linguistique de Louvain* 19, no. 3-4: 61-70.
- Francard, M. 1993b. *L'insécurité linguistique en Communauté française de Belgique*. Brussels: Service de la Langue Française.
- Francard, M. 1997. Insécurité linguistique. In *Sociolinguistique. Concepts de base*, ed. M. Moreau, 170-76. Sprimont: Pierre Mardaga éditeur.
- Great Britain Schools Council, and D. Sharp. 1973. *Attitudes to Welsh and English in the Schools of Wales*. London and Cardiff: Macmillan and University of Wales Press.
- Labov, W. 1966. *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.
- Labov, W. 2006. *The Social Stratification of English in New York City*. 2nd ed. Cambridge UK and New York: Cambridge University Press. (Orig. pub. Washington D.C.: Center for Applied Linguistics, 1966.)
- Siguan, M. 1999. *Conocimiento y uso de las lenguas: investigación sobre el conocimiento y uso de las lenguas cooficiales en las comunidades autónomas bilingües*. Madrid: Centro de Investigaciones Sociológicas.