

**Principles of variationism for disambiguating
language contact phenomena:
The case of lone Spanish nouns
in Catalan discourse**

JOSÉ LUIS BLAS ARROYO AND
DEBORAH TRICKER
University Jaume I

ABSTRACT

Using the variationist comparative method, the status of ambiguous lone Spanish-origin nouns in Catalan discourse is determined by analyzing their distribution and conditioning and by comparing them to their counterparts in unmixed Spanish or in multiple-word code-switches. Some areas of the nominal grammar have been selected for contrastive purposes (determination, complementation, gender, number) because they represent sites of coincidence as well as conflict between the two languages in contact, Spanish and Catalan. The main conclusion of this research is that Spanish-origin nouns in an otherwise Catalan context present grammatical variability similar to that of Catalan nouns, and that they behave differently from Spanish nouns in a monolingual context. In short, the grammar of these nouns is Catalan, and their categorial status is that of loanwords and not that of code-switches.

LANGUAGE MIXING: LEXICAL BORROWING VERSUS
CODE-SWITCHING

Within the area of research into language contact, one of the thorniest problems facing scholars has been the question of the epistemological borders between different phenomena. From a theoretical point of view we have to decide when a foreign element in the discourse of a language loses the status of lexical borrowing and enters the conceptual area of language alternation or code-switching.

If we take as our starting point the criterion of linguistic integration, we may say that lexical borrowing implies the incorporation of words from a donor language in the discourse of another (borrowing) language. If we adapt this definition to the bilingual Catalan–Spanish corpus on which the present study is based, we would have to state that in (1) the italicized element is a conspicuous case of lexical borrowing from Spanish (*entonces*) in the spoken Catalan discourse of Valencia.¹

The authors would like to thank Shana Poplack and Emilio Ridruejo for their valuable comments on an earlier version of this article.

- (1) *Entonces*, com anava dient-te
 ‘So, as I was telling you’

Anyone familiar with the Valencian linguistic situation knows that the adverb *entonces* in its usage as connector and/or discourse marker is not a native Catalan element, but that it occurs frequently and certainly much more often than the normative form *doncs*, at least at a colloquial level, in the Valencian dialect. On the other hand, the word is adapted to Catalan phonetics via the substitution of the alveolar sibilant /s/ for the interdental consonant /θ/, which is unknown in Catalan.

Similarly, although here the movement is in the other direction, the italicized term in (2) is a case of lexical borrowing from Catalan in the spoken Spanish of Valencian-speaking communities. Its status as a borrowing—already consolidated in the speech community—resides, on the one hand, in its remarkable frequency in the discourse of bilingual and even monolingual Valencians and, on the other hand, in the process of linguistic integration which has taken place.

- (2) ¿Ya estás *dotoreando* otra vez? (Sp. *cotilleando*)
 ‘Are you gossiping again?’

Code-switching would be understood as the alternating use of two languages by the same speaker, without any full or partial adaptation of the languages’ constituents occurring as a result of this use. In Catalan–Spanish bilingual discourse the most common occurrences of code-switching have neither the social significance nor the recurrent nature observed in other speech communities (cf. Myers-Scotton 1993; Poplack 1990), but they may be noted relatively frequently, especially when various contextual factors contribute to their occurrence. Thus, (3) contains a piece of direct speech that leads to an undisputed code-switch. This is a clear alternation, in which the boundaries of each language are perfectly defined.

- (3) (Cat.) *La tia desgraciada va i me diu l’atre dia*: (Sp.) A ver si limpiamos el portal, eh!
 ‘The silly woman goes and says to me the other day: How about cleaning the entrance, eh!’

Despite the apparent clarity of these ideas, several writers have noted the difficulties involved in defining the exact limits of these concepts when analyzing naturally occurring data. Indeed, in practice, some researchers have queried the validity of these principles for disambiguating the linguistic consequences of contact, arguing, with greater or lesser emphasis, that the phenomena in question are in real life interpreted by bilingual individuals in the same way and thus are hardly distinguishable (Bentahila & Davis, 1991; Gardner-Chloros, 1991; Myers-Scotton, 1993; Treffers-Daller 1990).

One of the most intriguing areas of the controversy surrounding the status of language mixing concerns isolated nouns that originate in the lexifying language,

but that in discourse appear in a linguistic context that belongs completely to the receiving language. One of the reasons for our interest in these linguistic forms is the fact that, as various authors have shown (Berk-Seligson, 1986; Eze, 1998; Ghafar Samar, 1996; Poplack & Meechan, 1995; Poplack, Sankoff, & Miller, 1988; Turpin, 1995), the noun is the most frequent word class in any bilingual corpus.

However, formal differences between noun groups can be found. If we take as our starting point the criterion of integration sketched here and consider it a discriminating factor in lexical borrowing, we find that, in the analysis of bilingual discourse, the material of the other language is indeed often adapted to the grammatical patterns of the receiving language. In the corpus on which this study is based, we observe this adaptation of borrowings from Spanish in Catalan discourse in examples like (4) and (5).

(4) Entonses el *jues* [xwes] se ha conviscut de que jo tinc raó
 ‘So the judge decided that I’m right’

(5) I entonses ells me van denunciar per *injúries* [inyurjes]
 ‘And then they sued me for slander’

These examples represent different degrees of integration in the grammar of Catalan. In (4), the noun of Spanish origin, *juez*, is partially adapted to the phonetics of Catalan. While it still retains the voiceless velar fricative [x] of the source language, the interdental [θ], which is non-existent in Catalan phonetics, is replaced by a sibilant, [s]. In (5), the Spanish noun *injurias* has an even higher degree of integration in Catalan. The velar [x] is transformed into a palatal sound, [y], and, from the morphological point of view, the atonal vowel [a] is replaced by [e], which is characteristic of plural formation in Catalan (Badia, 1962).

However, many other nouns of Spanish origin in a Catalan linguistic context have a form which we might initially call “bare” (i.e., without the superficial marks of grammatical or phonetic integration characteristic of the items just discussed). This is the case with *aseo* and *duro*, as shown in (6) and (7).

(6) I a la que me vaig anar al *aseo* me veig plena de sang
 ‘And when I went to the toilet I saw I was covered in blood’

(7) I la meua familia li va dir que no li pasava ni un *duro*
 ‘And my family told him they wouldn’t give him a penny’

The status of this last group of nouns represents an important challenge for the theory of language mixing, since there are no superficial indications other than purely etymological ones that allow us to establish clearly the linguistic affiliation of these words. This is even more relevant in the case of contact between typologically similar languages such as Spanish and Catalan. The fact that nouns of Spanish origin can be inserted in an otherwise Catalan context at syntactic

points in the clause, congruent with the rules of both grammars, means that the syntactic criterion becomes less useful for distinguishing between the phenomena of code-switching and lexical borrowing (cf. Budzhak-Jones & Poplack, 1997; Meechan & Poplack, 1995; Poplack & Meechan, 1995; Poplack, Wheeler, & Westwood, 1987). Similarly, phonetic assimilation cannot be a criterion, since it has been widely shown to be variable among bilingual individuals of the community, both in cases of borrowing and code-switching. And it goes without saying that, from a strictly linguistic point of view, leaving the normative issue to one side, sentences like (6) and (7) are clearly grammatical. For this reason, it is difficult to assign this kind of noun simply to one language or the other.²

PRINCIPLES OF VARIATIONIST ANALYSIS:
THE COMPARATIVE METHOD

For some researchers these nouns are simply a manifestation of code-switching—changes that in some cases involve a violation of the grammar of the source language, the target language, or both (Backus, 1992; Bentahila & Davis, 1983; Berk-Seligson, 1986; Boeschoten, 1990; Boeschoten & Verhoeven, 1985; Bokamba, 1988; Muysken, 1987; Romaine, 1989). These conclusions usually lack empirical proof beyond that offered by a few isolated examples. However, a team of researchers led by Shana Poplack has proposed an analytical model that attempts to resolve in a rigorous manner the controversy surrounding the status of these isolated words.

Following the principles of the variationist approach—detailed research into the spontaneous speech of bilingual individuals, exhaustive analysis of the relevant data, and the clear definition of the aim of the study and, especially, of the variable context—these authors have developed a comparative method to test whether linguistic material should be treated as a manifestation of one or the other category of language mixing. Several studies based on this approach have been published (Adalar & Tagliamonte, 1998; Budzhak-Jones, 1998; Budzhak-Jones & Poplack, 1997; Meechan & Poplack, 1995; Poplack & Meechan, 1995, 1998). Most of these studies, though, have been concerned with comparing typologically different languages, such as French and the African languages Wolof and Fongbe (Meechan & Poplack, 1995; Poplack & Meechan, 1995), English and Ukrainian (Budzhak-Jones & Poplack, 1997), or English and Persian (Ghafar Samar & Meechan, 1998). As far as we are aware, there have been few, if any, applications of the comparative method to situations of contact between structurally similar languages.³

For this reason, one of the objectives of this article is to evaluate the method for studying contact between two typologically similar languages: Spanish and Catalan. Indeed, this analytical model may acquire a particular relevance in such contexts, since, precisely because of the structural and formal proximity between such pairs of languages, in many cases it would be impossible to decide which mechanisms of language mixing are involved. The comparative method uses the

information supplied by the patterns of grammatical variability to determine the linguistic affiliation of ambiguous nouns (in this case, nouns of Spanish origin in an otherwise Catalan context).

In applying this analytical model, it is accepted that the linguistic affiliation of ambiguous elements cannot be determined from isolated variables. Neither integration, whether it be phonological, grammatical, or semantic, nor recurrence can serve to determine the status of isolated lexical elements in bilingual discourse. Both offer notable difficulties. Integration has often proved to be a variable and gradual phenomenon, incapable of distinguishing clearly between various contact phenomena; frequency is problematic because of the lack of predictability of lexical elements in discourse.

Considering these difficulties and also recognizing that a case-by-case disambiguation of contact phenomena may be impossible, the comparative method argues that classes of items in bilingual discourse can be predominantly characterized as borrowings or code-switches (cf. Poplack & Meechan, 1998). This means that evidence for the status of these words can be obtained only through an empirical analysis of the linguistic structures in which they appear, in accordance with Labov's principle of accountability (cf. Labov, 1972; Poplack, 1993). The application of this principle in the present case (i.e., for comparing classes of contact phenomena in bilingual discourse) is summarized as follows by Meechan and Poplack (1995:173): "Their different rates of occurrence in each structure form a quantitative pattern which can be compared with that of their counterparts in a monolingual context."

HYPOTHESIS

In accordance with the line of argumentation set out here, the present study of Spanish–Catalan contact will analyze the distribution of a series of syntactic and morphological factors related to the nouns that are the object of our analysis (i.e., those of Spanish etymological origin in an otherwise Catalan context). These data will be compared with data of other noun groups, namely (1) nouns of Spanish origin in a Spanish monolingual context, (2) nouns of Spanish origin in a mixed context (unambiguous code-switches), and (3) Catalan nouns in a Catalan monolingual context.

In this theoretical framework, we propose the following hypothesis. If the patterns of grammatical variability of nouns of Spanish origin in a Catalan context coincide with Catalan nouns but not with Spanish nouns in a Spanish context, we shall have to conclude that, leaving aside their etymology, their grammar corresponds to Catalan. That is to say, in taxonomical terms, they are lexical borrowings and not code-switches. On the other hand, if the variability of these lexical elements coincides with that observed in Spanish nouns in a monolingual Spanish context or in unambiguous code-switches, we shall have evidence that the linguistic mixture in our corpus occurs basically in the form of language alternation.

TABLE 1. *Distribution of the nouns in the corpus by context*

Catalan	Ambiguous	Spanish	Code-switching
223	234	218	108

METHODOLOGY

In addition to the principles of variationist analysis, the study of contact phenomena has two essential methodological requirements: identifying a real speech community where these discourse modalities are frequently used and obtaining a sufficiently wide and representative corpus to serve as empirical support for the research.

The corpus for the present study consists of the speech of 15 bilingual individuals, born in the bilingual regions of the Valencian community, who took part over several days in a local television program, presided over by a presenter who asked questions and gave permission to speak. The participants told various personal experiences related to a previously established subject. The programs were recorded on a VHS-Pal video recorder on the days of their broadcast.

The data extracted from this corpus are comparable for various reasons. All of the tokens shared a similar semi-informal register and an identical discourse genre—the narrative of personal experiences (Labov, 1997). All of the speakers had a common social background that might be classified as lower middle class, as evidenced by their use of nonstandard varieties of both Spanish and Catalan. Finally, the selected tokens all had Catalan as the main language (i.e., the language used for most of the discourse) with occasional changes to Spanish of a situational nature (i.e., alternations corresponding to changes in the elements that make up the communicative event: e.g., change of topic, modification of the interactional tone, etc.) (Gumperz, 1982).

The resulting corpus contains various examples of linguistic contact, including unambiguous manifestations of borrowing and code-switching as well as material, generally restricted to the word level, which does not appear to be superficially integrated into the receiving language. This bilingual corpus, made up of utterances where Spanish and Catalan coincide, was transcribed by the authors. The data were coded using the statistical program GOLDVARB 2.0 (Rand & Sankoff, 1990) for Macintosh computers. The quantitative differences were tested using a chi-squared statistical significance test. Following transcription, all nouns were extracted from the corpus and classified according to two overlapping criteria: (1) the language to which they belong etymologically (Spanish or Catalan), and (2) the discourse context (Spanish, Catalan, or mixed) in which they appear. Table 1 shows the data available for the whole study.

As the main objective of the study was to analyze the status of lone Spanish-origin nouns in an otherwise Catalan context, the first selection was of a subgroup of tokens with these characteristics, exemplified in (8). For ease of presentation we refer to these nouns as “ambiguous.”

- (8) El seu home es el seu *terreno*?
‘Is your husband your territory?’

However, as stated in note 1, this initial sample of bare nouns, which did not exhibit any superficial indications of linguistic integration, was later supplemented with some lone nouns which did exhibit some morphological adaptation.

Subsequently, with the aim of empirical comparison, three additional classes were selected: (1) Spanish nouns in a Spanish linguistic context, as in (9), (2) Spanish nouns in a mixed linguistic context (unambiguous code-switches; they appear in constituents greater than the word), as in (10), and (3) Catalan nouns in a Catalan monolingual context (11). For convenience, these three types of nouns are referred to as Spanish, code-switching, and Catalan, respectively.

- (9) Al final puede acabar en una *sanción* económica (Spanish)
‘It can end in an economic sanction’

- (10) Va pasar pel meu costat i no me va conèixer i diu *pero qué pelo te has cortao?*
(code-switching)
‘He walked past me and didn’t recognize me and said what have you done to your hair?’

- (11) Li va donar unes *herbes* (Catalan)
‘He gave her some herbs’

CODING AND ANALYSIS

As a basis of variationist analysis, a contrastive study of various areas of grammar that affect the noun in both languages is required, in order to show the similarities and differences both at the formal and at the syntagmatic level.⁴ We therefore propose an analysis in which, along with the information supplied by the grammars of each language in monolingual discourse, the most relevant contrastive aspects are highlighted; the aim is to determine to what extent there are significant differences between one group of nouns and another, according to their etymological origin and the linguistic context (monolingual or bilingual) in which they appear.

Two groups of structural factors were used for this task. The first group was concerned with the syntactic rules governing nouns within the noun phrase (NP). For this purpose, we turned to the criteria of determination and complementation as the two areas of nominal syntax with the greatest degree of structural diver-

gence and similarity between Spanish and Catalan. The second group involved morphological variables such as gender and number, which present points of similarity and structural conflict between the two languages and which assist in evaluating the use of grammatical variability as a disambiguating criterion.

In short, we propose an analysis in which, along with the information supplied by the grammars of each language⁵ in monolingual discourse, the most relevant contrastive aspects are highlighted. Our aim is to determine to what extent there is significant divergence between the different groups of nouns as a function of their etymology and the linguistic context in which they appear.

NP syntax in monolingual and bilingual discourses

Determiners traditionally refer to the group of linguistic elements that function as nominal adjuncts, although, with the exception of the article, they can also do so as head without the need for a transposer (Álvarez, 1986:129). For Hernanz and Brucart (1987:176), what unites all these elements is their actualizing ability, so that the term “determiner” should be used to refer to those units which can accompany a noun. Thus, the collocation of the group in an argumental position does not produce ungrammaticality.

From a syntagmatic point of view, determiners are characterized by the fact that, with a few exceptions, they appear in complementary distribution. Another relevant syntactic feature is their obligatory pre-head positioning within the NP. Finally, unlike qualifying adjectives, they appear in closed paradigms and do not admit the possibility of semantic grading (cf. Álvarez, 1986; Lorenzo, 1995).

Traditional grammar divides determiners according to various semantic features: on the one hand, deixis, which groups together articles, demonstratives, and possessives (which in turn differ according to additional features), and, on the other, quantity, whether inexact (indefinite) or of a definite number (numerals). In the present study, we classify all determiners of the second type together as quantifiers, regardless of whether they express an exact quantity.

The typological proximity of Spanish and Catalan means that the members of the determiner category have very similar semantic profiles. A detailed analysis of the two languages, however, reveals within each paradigm some forms and combinations that are shared by both languages, as well as ones that are exclusive to one or the other. For example, consider the identical manifestations of the singular article *el* ‘the’ in Spanish and Catalan, as in (12) and (13); we can observe others exclusive to each language in the formation of the plural (Sp. *las*; Cat. *les* ‘the’), as in (14) and (15):

- (12) (Sp.) Esperamos que con *el incremento* de los usuarios . . .
‘We hope that with the increase in users . . .’

- (13) (Cat.) Han [(g)overnat] allí *el sopar*?
‘Did they give you supper there?’

- (14) (Sp.) Tú das *las aperturas*
‘You give the openings’
- (15) (Cat.) Estaven totes *les indústries* fortes arropant a l'alcalde
‘All the strong industries were there backing the mayor’

Similarly, while some combinations of determiners and nouns coincide in Spanish and Catalan, as in (16) and (17), others are clearly exclusive to one language and are completely ungrammatical in the other. This is shown clearly by a comparison of (18) and (19), based on a sample of speech from our corpus.

- (16) (Cat.) I vosatros mireu *ixe programa* de ixe incult
‘And you watch that ignorant man’s programme’
- (17) (Sp.) Con *esta gente* se pueden hacer maravillas
‘With these people you can work wonders’
- (18) (Cat.) *El seu home és el seu terrenó?*
‘Is your husband your territory?’
- (19) (Sp.) *¿*El su hombre es el su terreno?* (cf. ¿*su marido es su terreno?*)
‘Is your husband your territory?’

To facilitate the comprehension of the coincidence and conflict sites between the two languages, let us now carry out a brief contrastive analysis of the different paradigms which make up this category.

As far as the definite article is concerned, it must be noted that the divergences between Spanish and Catalan are concentrated mainly in two areas: (1) plural formation, and (2) the characteristic contractions of Catalan, which are more numerous than in Spanish.

The forms of the article in Catalan are *el* and *la* for the masculine and feminine singular and *els* and *les* for the masculine and feminine plural. The Spanish singular forms are identical; the plurals are *los* and *las*. Catalan singular articles are reduced to *l'* in certain phonetic contexts: for example, before words beginning with a vowel or a (silent) *h*, as in *l'aire* ‘the air’ and *l'home* ‘the man’.⁶ The forms of the masculine article are also contracted when combined with the prepositions *a*, *de*, *per*, or *ca*, creating more extensive paradigms than those of modern Spanish (Cat. *al*, *als*, *del*, *dels*, *pel*, *pels*, *cal*, *can* vs. Sp. *al*, *del*). However, unlike Spanish, the union of preposition and article does not occur when the following word demands the contracted form of the article (Cat. *de l'home*; Sp. *del hombre* ‘the man’s’).

With regard to the indefinite article, there are fewer differences between Spanish and Catalan. These only occur in the plural: *uns* (masculine) and *unes* (feminine) in Catalan versus *unos* (masculine) and *unas* (feminine) in Spanish.

Next, we consider the formal divergences between the two languages in the expression of the zero article: in particular, when the noun is followed by a prep-

ositional phrase (PP). In such cases, the difference depends on whether the preposition preceding the noun is a homophone in the two languages (Sp. *cuadra de ladrones*; Cat. *quadra de lladres* ‘gang of thieves’), or whether it is a form exclusive to Catalan (*dona amb [en⁷] estil* ‘woman with style’).

In both Spanish and Catalan, the demonstratives are determiners characterized by (1) a closed paradigm, (2) variation in gender and number, in agreement with the accompanying noun,⁸ and (3) fixed position. That is, they generally appear before the noun, but if they are postposed, the article must occur in prehead position (Sp. *viene esos hombres, vienen los hombres esos*; Cat. *venen eixos homes, venen los homes eixos* ‘those men are coming’). Unlike other varieties of Catalan, which, like English, have only two dimensions of space–time (*aquest* ‘this’, *aquell* ‘that’), the Valencian dialect has a triadic system similar to that of Spanish, which gives rise to some common forms between the two languages (*este, estos, esta* ‘this’, *aquella* ‘that’). Other demonstratives are clearly differentiated, whether through the application of the morphophonological rules specific to each language in the formation of feminine plurals (Cat. *estes, aquelles*; Sp. *estas, aquellas* ‘this, those’) or through etymological differences (Cat. *eixe, eixa, aquell*; Sp. *ese, esa, aquel* ‘this, that, that over there’).

The possessives represent, without a doubt, the area of determination with the most radical divergences between the two languages; there are no variants common to both languages (see Table 2). These differences are not restricted to the formal aspect (Cat. *teu, seu*; Sp. *tu/tuyo, su/suyo* ‘your, it/his/her/their’), but affect the combinatory rules of the elements making up the NP. The main structural differences are concentrated in the subparadigm of the tonic forms. In Spanish these occur only when the possessive follows the noun (*el libro es mio* ‘the book is mine’),⁹ whereas in Catalan they can occupy either of the positions: prehead (*el meu gos* ‘my dog’) or posthead (*el gos meu*). Catalan offers the possibility of postponing the tonic forms in certain noun phrases without any determiner in prehead position (*a casa meua* ‘to my house’, *per culpa teua* ‘because of you’), which is completely ungrammatical in Spanish; Spanish would only admit the corresponding apocopated form (**a mi casa, por tu culpa*) (Tió, 1982:56).

Finally, we include the data on certain less frequent forms in a single group of “other” determiners. While less frequent, they exhibit patterns of structural variability of interest for the present study. Here we include quantifiers, generally divided into numerals, quantitatives, and indefinite items, as well as traditional exclamative–interrogative adjectives. Like the other paradigms, the components of this group exhibit three classes of variant: (1) variants common to both languages (*tres* ‘three’, *alguna* ‘some’, *qué* ‘what’), (2) variants exclusive to Catalan (*dues* ‘two’, *algú* ‘some’, *quina* ‘what’), and (3) variants exclusive to Spanish (*cuatro* ‘four’, *algún* ‘some’, *bastante* ‘enough’). Table 2 summarizes the factors indicated within the grammatical category of determination.

As far as complementation is concerned, it should be noted that in both Spanish and Catalan this function is carried out by different lexical and syntagmatic

TABLE 2. *Paradigms of determination in the NPs of the corpus*

Determiner	Examples
Definite article	Spanish (<i>los, las</i>) Catalan (<i>les, els</i>) Common (<i>el, la</i>)
Indefinite article	Spanish (<i>unos, unas</i>) Catalan (<i>uns, unes</i>) Common (<i>un, una</i>)
Preposition + zero article	Catalan (<i>sense, per</i>) Common (<i>a, de</i>)
Demonstrative	Spanish (<i>esa, aquellas</i>) Catalan (<i>eixe, aquell</i>) Common (<i>esta, este</i>)
Possessive	Spanish (<i>mi, tu</i>) Catalan (<i>el meu, el teu</i>)
Others	Spanish (<i>algunos, otros</i>) Catalan (<i>alguns, altres</i>) Common (<i>qué, tres</i>)
Contraction	Catalan (<i>pel, als</i>) Common (<i>al, del</i>)

categories, which, with the exception of some adjectives, always follow the head. Some of the possibilities corresponding to monolingual Spanish or Catalan discourse are presented in (20) through (26).

(20) No modification

(Sp.) Nos tocó la *lotería*

'We won the lottery'

(Cat.) No li agrada res la *gent*

'He doesn't like people at all'

(21) Preposed adjective

(Sp.) Yo creo que es la zona que más mejorará en los *próximos años*

'I think it's the area that's going to improve most in the next few years'

(Cat.) Còmpre unes *bones polaines* una temporada

'Buy yourself some good legwarmers for a season'

(22) Postposed adjective

(Sp.) Todos los *genes políticos* los tiene Carlos

'Carlos has got all the political genes'

(Cat.) Porte casi ja vint dies en el *peu fastidiat*

'My foot's been playing me up for nearly twenty days now'

- (23) PP
 (Sp.) Lo que más me impresionó fue la *limpieza del pueblo*
 ‘What most impressed me was the cleanliness of the town’
 (Cat.) Tres-cents milions val la *farmàcia de Benicasim*
 ‘The pharmacy in Benicasim is worth 300 million pesetas’
- (24) Subordiante clause
 (Sp.) Voy a decirle un *problema que tengo*
 ‘I’m going to tell you about a problem I’ve got’
 (Cat.) Jo tinc una *filla que està estudiant farmàcia*
 ‘I’ve got a daughter who’s studying pharmacy’
- (25) Appositive NP
 (Sp.) Llegar tú con el *apellido Fabra* también
 ‘You turning up with the name Fabra as well’
 (Cat.) I demà en el *periòdic Mediterraneo* tots els jugadors del Playas
 ‘And tomorrow in the *Mediterraneo* newspaper all the Playas players’
- (26) Multiple complementation¹⁰
 (Sp.) Seguimos con Miguel, con *temas interesantes del presente*
 ‘We continue with Miguel, with interesting present-day issues’
 (Cat.) Això ha donat molta *ajuda moral i social per al meu poble*
 ‘That’s helped my town a lot morally and socially’

As these examples show, what we are dealing with here are areas of strong structural equivalence between the two languages. There are very similar syntagmatic combinations around the head of the NP, with the single possible exception of the preposing of the adjective (21). This preposing according to some scholars (e.g., Badia, 1994) is less characteristic of Catalan than of Spanish, even though we may not consider it incorrect in all cases because of its expressive importance.

Morphological factors

In addition to the syntactic factors of determination and complementation, we have taken into account the role played by certain grammatical morphemes in the formal modifications undergone by nouns to facilitate the functions of relation and articulation within the clause. Gender and number act within these Romance languages as elements that allow both the categorization of nouns as nouns and the agreement of various units within the NP and even, occasionally, outside it. In both languages, gender and number are bound morphemes.

The system of the gender morpheme in Spanish and Catalan is formed by the masculine/feminine opposition. From another point of view, gender in both languages falls among those morphological attributes introduced by flecional affixes, as opposed to those in which the grammatical information is incorporated exclusively by the lexis (Picallo, 1991:212; Ritter, 1991:50). However, it is clear that neither Spanish nor Catalan meets all the criteria of this morphological char-

acterization, since in both languages words habitually have invariable gender (Sp. *mesa*, **meso*; Cat. *taula*, **taulo* ‘table’) (cf. Lorenzo, 1995:40). In both languages, some nouns, mainly those referring to human beings and animals, distinguish masculine and feminine genders through heteronymy. Finally, there are other common procedures, such as the determination of gender through the article (Sp. *el cónyuge*, *la cónyuge*; Cat. *el cònjuge*, *la cònjuge* ‘spouse’), that in some cases give rise to changes in meaning (Sp. *el cólera* ‘cholera’, *la cólera* ‘rage’; Cat. *el còlera*, *la còlera*).

Nevertheless, as Lloret and Viaplana (1997) observed, where there is gender variation, there is a formal difference between the two languages. In Spanish gender is articulated around two possible oppositions: either the opposition of the binary feature (\pm fem) or the exclusive opposition of masculine and feminine features. In Catalan gender is articulated around an exclusive opposition: the feminine is marked formally by the central vowel /ə/, whereas the masculine is not marked by anything.

Tió (1982:55) noted that the principal points of structural conflict between the two languages with regard to the gender morpheme occur in the Catalan oral and written variations of the feminine with respect to the corresponding masculine radicals in those words that admit gender variation. By contrast, in Spanish, the alternation occurs simply between the morphs *-o/-a* (compare Cat. *nou/nova* vs. Sp. *nuevo/nueva* ‘new’; Cat. *amic/amiga* vs. Sp. *amigo/amiga* ‘friend’; Cat. *menjat/mentjada* vs. Sp. *comido/comida* ‘eaten’; and Cat. *llop/lloba* vs. Sp. *lobo/loba* ‘wolf’). This does not include the other marginal procedures that affect limited nominal paradigms, such as the use of special suffixes, which are different for each language (Sp. *-esa*, *-isa*, *-triz*; Cat. *-essa*, *-iu*, *-ina*).

Other formal gender-related differences between the two languages lie in the type of ending. Thus, a masculine Catalan word ending in *-e*, a consonant, or a semivowel corresponds to a feminine Spanish word ending in *-a*. Similarly, a feminine Catalan word ending in *-a* corresponds to a masculine Spanish word ending in *-o*, *-e*, or a consonant. For example, words that are masculine in Catalan but feminine in Spanish include Cat. *el cep*, Sp. *la cepa* ‘stock’, Cat. *el cim*, Sp. *la cima* ‘summit’, Cat. *el deute*, Sp. *la deuda* ‘debt’; words that are feminine in Catalan but masculine in Spanish include Cat. *l’arada*, Sp. *el arado* ‘plough’, Cat. *la xocolata*, Sp. *el chocolate* ‘chocolate’, Cat. *les despulles*, Sp. *los despojos* ‘spoils’ (Badia, 1962:135). Thus, we coded the nouns in our corpus according to the gender morpheme, as shown in the following examples:

Masculine nouns: Cat. *transporte*; Sp. *transporte* ‘transport’

Feminine nouns with no variation: Cat. *taula*; Sp. *mesa* ‘table’

Feminine nouns with variation: Cat. *amic/amiga*; Sp. *amigo/amiga* ‘friend’

Finally, number is the morpheme which qualifies the increase in quantity over that provided by the noun lexeme. As in the majority of Romance languages,

Spanish and Catalan only have a distinction between singular and plural. However, the rules of plural formation vary from one language to the other. For example, the plural is marked in Spanish through the allomorphs *-s* and *-es*, depending on the previous phonetic context (*casa-s* ‘house-*s*’ vs. *camión-es* ‘truck-*s*’); this is a general rule, to which there are a few well-known exceptions affecting certain tonic vowels (e.g., *esquí-s* ‘skis’). In Catalan, the plural of nouns is in general formed by adding *-s* to the singular, regardless of whether the word ends in a vowel or a consonant; in the latter case, this gives rise to combinations of consonants unknown in Spanish (*camions* ‘trucks’, *senyors* ‘gentlemen’). The addition of plural *-s* occasionally gives rise to phonetic or spelling changes in some nouns. In this context, the closing of the vowel /a/ to /ə/ in the formation of plurals (*taula, taules* ‘table-*s*’) should be noted for its importance in the Valencian dialect, where it affects not only spelling, but also pronunciation. Apart from the general rule, we should point out that in Catalan the *-s* is often accompanied by other phonemes in various contexts, leading to consonant changes in the root. Thus, most nouns ending in a tonic vowel form the plural with *-ns* (*pa, pans* ‘bread, loaves’), with an underlying nasal consonant which becomes explicit in derived forms (*pa, panificador* ‘bread, baker’). However, there are numerous exceptions which follow a general rule: that is, the use of *-s* alone (*robí-s* ‘ruby, rubies’). Nouns ending in certain sibilants, such as *ç, s, x*, or in a combination of consonants containing one of these sibilants, such as *tx, sc, st, xt*, add *-os* (*braç-os, cas-os* ‘case-*s*’). There are also exceptions here, which are usually resolved by turning to the general rule. For example, Payrató (1988:96) noted a tendency in colloquial Catalan to regularize the flectional paradigm through various morphs that are nonstandard for the phonic contexts in which they appear, as in the case of *-os* in *llàpissos* ‘pencils’ and *dillunsos* ‘Mondays’ or *-es* in *masses* ‘masses’ and *forces* ‘forces’.¹¹ Thus, all the nouns of the corpus in the present study were coded in accordance with the following factors:

Singular: Cat. *metge* ‘doctor’; Sp. *pueblo* ‘village’

Forms exclusive to the Catalan plural: *senyors* ‘gentlemen’

Forms exclusive to the Spanish plural: *camiones* ‘trucks’

Common forms of the plural: Cat. *diumenges* ‘Sundays’; Sp. *hermanos* ‘brothers’

In short, as we have indicated on several occasions, decisive evidence on the linguistic status of lone Spanish-origin nouns in an otherwise Catalan context can be determined only by the underlying grammar in its realization in the hierarchy imposed by the grammatical factors conditioning the variability.

TABLE 3. *Distribution of articles and contractions in the NPs of the corpus*

	Catalan	Ambiguous	Spanish	Code-switching
	% (N)	% (N)	% (N)	% (N)
Definite article				
Spanish	1.2 (1)	0 (0)	20 (21)	15.4 (6)
Catalan	24.4 (20)	16.2 (12)	0 (0)	0 (0)
Common	74.4 (61)	83.8 (62)	80 (84)	84.6 (33)
Indefinite article				
Spanish	0 (0)	0 (0)	3.2 (1)	0 (0)
Catalan	25.8 (10)	2 (1)	0 (0)	0 (0)
Common	74.4 (29)	98 (49)	96.8 (30)	100 (14)
Zero article				
Catalan preposition	4.3 (2)	13.3 (8)	0 (0)	0 (0)
Common preposition	95.7 (42)	86.7 (52)	100 (28)	100 (26)
Contraction				
Catalan	8.3 (1)	5.5 (1)	0 (0)	0 (0)
Common	91.7 (11)	94.5 (17)	100 (11)	100 (1)
Total	177	202	175	80

Note: We have followed the convention of highlighting in boldface and/or italics any similarities in the pattern of quantitative distribution. Thus, figures for categories and nouns in boldface have similar distributional patterns, which differ from those in boldface italics.

RESULTS

SYNTACTIC FACTORS

Determination

Tables 3, 4, and 5 show the results of quantitative frequency analysis for the types of determination in the corpus.¹²

Before commenting on the parallelisms and the distributional differences observed in the tables, the data were tested for statistical significance using chi-squared tests to measure the validity of the comparisons of the different population samples. From the data produced by this test the following assumptions could be made.

1. The frequency differences between Catalan nouns and Spanish nouns as well as between Catalan nouns and code-switching nouns are all highly significant ($p < .001$).
2. The frequency differences between ambiguous nouns and Spanish nouns and between ambiguous nouns and code-switching nouns are also all highly significant ($p < .001$).

TABLE 4. *Distribution of definite/indefinite articles among the nouns in the corpus*

	Catalan	Ambiguous	Spanish	Code-switching
	% (N)	% (N)	% (N)	% (N)
Definite	67.7 (82)	59.7 (74)	77.3 (105)	73.6 (39)
Indefinite	32.3 (39)	40.3 (50)	22.7 (31)	26.4 (14)
Total	121	124	136	53

TABLE 5. *Distribution of other kinds of determination in the NPs of the corpus*

	Catalan	Ambiguous	Spanish	Code-switching
	% (N)	% (N)	% (N)	% (N)
Demonstrative				
Spanish	10 (1)	0 (0)	64.7 (11)	33.3 (2)
Catalan	60 (6)	66.6 (4)	0 (0)	0 (0)
Common	30 (3)	33.4 (2)	35.5 (6)	66.4 (4)
Possessive				
Spanish	0 (0)	7.7 (1)	100 (14)	100 (15)
Catalan	100 (14)	92.3 (12)	0 (0)	0 (0)
Others (interrogatives, numerals, etc.)				
Spanish	5.5 (1)	16.6 (2)	58.5 (7)	57.1 (4)
Catalan	86.4 (19)	58.3 (7)	0 (0)	0 (0)
Common	9.1 (2)	25.1 (3)	41.5 (5)	42.9 (3)
Total	46	31	43	28

3. However, the frequency differences between Catalan nouns and ambiguous nouns are only significant on two occasions (indefinite article, other determiners), with a probability for these cases close to 5%.
4. Finally, the differences between Spanish nouns and code-switching nouns are not significant, except in the case of the demonstratives, which have an index of probability close to the threshold of acceptable significance (i.e., 5%).

These data on the statistical significance of the differences observed at first glance lead to an interesting point. The frequencies of nouns in a Catalan context, whatever their etymological origin, have similar patterns of grammatical variability, which are different from those noted for Spanish nouns in the other two contexts (monolingual Spanish, code-switching).

Definite article. It must be stressed that Spanish nouns in a Spanish context occur with this form of determination more frequently than do nouns in a Catalan

context, regardless of their origin (see Table 4). Catalan nouns and ambiguous nouns also share a closer distribution as compared to the other two noun types of the corpus.

It is true that the highest percentages in all the groups are for the forms of the article which are common to both languages: Catalan 74.4%; ambiguous 83.8%; Spanish 80%; code-switching 84.6%. This is not surprising if we consider that these common forms, which represent points where the structures of the two languages coincide, are concentrated in the singular of the two paradigms. Indeed, the singular is by far the most frequent number among all of the nouns of the corpus. These forms are illustrated in (27), (28), (29), and (30)¹³ for all classes of nouns.

- (27) *Porte casi ja vint dies en el peu fastidiat* (Catalan)
‘My foot’s been playing me up for twenty days’
- (28) *I ho tire a la papelera* (ambiguous)
‘And I threw it in the bin’
- (29) *Tuvimos la desgracia de perderlo* (Spanish)
‘We were unlucky enough to lose it’
- (30) **Prohibido tocar el género**, pero no m’ho palpeu (code-switching)
‘Please do not touch the goods, don’t touch them’

The differences are concentrated, however, in the exclusive units of each language. As shown in Table 3, there is a strong discrepancy between the forms of article used by nouns in a Catalan context, independent of their linguistic origin, and those used by Spanish nouns or code-switching nouns. Note that exclusive forms of the Catalan definite article do not appear in any of these examples, whereas, as expected, exclusive forms of the Spanish article do appear, as in (31) and (32):

- (31) *Todos los genes políticos los tiene Carlos* (Spanish)
‘Carlos has got all the political genes’
- (32) *que demà és Sant Valentí, el día de los enamorados* (code-switching)
‘tomorrow’s St Valentine’s Day, the day for lovers’

This is exactly the opposite of what occurs with nouns in a Catalan context, where only Catalan forms are used and exclusively Spanish articles do not appear.¹⁴ This occurs with Catalan nouns and with ambiguous nouns.

- (33) *Qué listos van ser els socialistes!* (Catalan)
‘Weren’t the socialists clever!’
- (34) *Porta els equipos* (ambiguous)
‘He’s carrying the equipment’

Indefinite article. The general distribution pattern is like that discussed for definite articles, with a few interesting differences. It should be observed that, for all the groups, the indefinite article is less frequent than the definite article. Table 4 shows the respective frequency of both forms of determination through the article for each of the classes of noun considered in our analysis.

However, if we turn to the data for Spanish nouns, it becomes clear that there is a marked quantitative difference between some of the groups. The figures for Spanish nouns and code-switching nouns are very similar (77.3% and 22.7% vs. 73.6% and 26.4%, respectively), whereas the frequency distribution of ambiguous nouns is closer to that of Catalan nouns; the differences between the two kinds of article, while still significant, are nevertheless smaller. In any case, it should be noted that ambiguous nouns are not patterning as an intermediate or mixed group between the two (Spanish and Catalan) grammars.

These similarities occur in other data for determination with the indefinite article. Apart from the considerable presence of forms of the article common to both languages (*un, una*), again explicable because they are concentrated in the singular paradigms of Spanish and Catalan, it should be noted that only the nouns in a Catalan context use forms exclusive to this language. Nevertheless, it is still significant that, whereas ambiguous nouns use forms of the Catalan indefinite article when they appear in the plural, as in (35), Spanish nouns (36) or code-switching nouns (37) never do so, preferring common forms or those exclusive to Spanish.

(35) I al costat havien *uns joyeros* txicotets (ambiguous)
'And next to it there were some little jewellers' shops'

(36) Vamos con *unos consejos* de la abuela (Spanish)
'Let's continue with some of granny's tips'

(37) I la xiqueta lo tenia i va i me diu, **dice: Rosa que aquel sitio hay unas señoras**
(code-switching)
'And the girl had it and she said to me, she said, Rosa, there are some ladies there'

Zero article. We distinguished only two possibilities, limiting our analysis to those cases in which the noun follows a preposition whose form is either common to both languages or exclusive to Catalan. The results, shown in Table 3, demonstrate once again that the only nouns with zero article with an exclusively Catalan preposition appear in a Catalan linguistic context, independent of their etymological origin. In (38) and (39), Catalan prepositions can be observed preceding both kinds of nouns.

(38) perquè està en verd en oscur *sense llum* (Catalan)
'because it's green, dark, without any light'

(39) Me van escriure i van tocar *per telèfono*. (ambiguous)
'They wrote to me and they called me on the phone'

In contrast, Spanish nouns and code-switching nouns are never preceded by these prepositions, but can be by others that are formally identical to the Catalan forms, as illustrated in (40) and (41).

(40) las riñas *de hermanos* de siempre (Spanish)
‘the typical rows between brothers and sisters’

(41) i me posava *de vuelta y media perejil* (code-switching)
‘And he gave me a piece of his mind’

To sum up, so far two clearly defined paradigms can be deduced: on the one hand, nouns in a Catalan context, regardless of their origin, and, on the other hand, Spanish nouns and code-switching nouns. Although in all the cases analyzed the most frequent forms are always those common to both languages, this is a result of the numerous points of structural similarity between the two languages for this paradigm of determination. The fact that most of the nouns in the corpus appear in the singular explains, for example, why the four groups coincide in the use of the common forms of the definite and indefinite article, since it is in the singular that the paradigms are formally identical in both languages (*el, la* and *un, una*).

When the nouns appear in the plural, the use of exclusively Catalan forms (*els, les, uns, unes*) or Spanish ones (*los, las, unos, unas*) leads to significant differences. Thus, we have seen how nouns in a Catalan context only use Catalan articles, regardless of whether they are Catalan or ambiguous.

Finally, Spanish nouns or code-switching nouns never use forms of the article that are exclusive to Catalan. Rather, they use forms common to both languages, in the case of singular articles, or exclusively Spanish ones, in the case of plurals.

The differences that have allowed us to articulate the two groups of nouns also extend to those cases in which the determination of the noun is realized through the zero article, when it is preceded by a preposition. Here again the discrepancies are clear. Leaving aside the prepositions whose forms are identical in both languages (*a, de*), differences arise in the rest of the prepositional system. In particular, the fact that Spanish nouns and code-switching nouns are never preceded by Catalan prepositions, but that nouns in a Catalan context are, is very revealing.

Contractions. We now turn to those cases in which the definite article is combined with a preposition in a single lexical unit. This is a grammatical characteristic of both Spanish and Catalan, although in Catalan the number of contractions is higher than that in present-day Spanish, in which only two forms, *al* and *del*, remain. These contractions are formally identical to the corresponding Catalan forms. However, in Catalan the contraction paradigm is wider, including the plural of the masculine article (*als, dels* ‘to the, of the’) and involving more prepositions than in Spanish (*pel, pels* ‘for the’) (cf. Badia, 1994:202–203).

The results of the quantitative analysis do not allow us to draw any definitive conclusions because of the scarcity of data for some areas. For example, the

overall figures for exclusively Catalan contractions in the corpus are low, in contrast to the much higher frequency of forms common to both languages. This is understandable, though, if we take into account the fact that these prepositions (*a* and *de*) and singular forms are much more frequent in the corpus than other prepositions or the plural.

In any case, the scarcity of data should not prevent the recognition of a significant fact related to the earlier discussion. The only nouns that are occasionally preceded by contractions exclusive to Catalan are those that appear in the discourse context of that language. This occurs regardless of etymology, affecting Catalan nouns and ambiguous nouns, as illustrated in (42) and (43), respectively.

(42) *Als xiquets* vende-les loteries (Catalan)
‘The children, sell them lottery tickets’

(43) Depén *dels apellidos* que els posen (ambiguous)
‘It depends on the surnames they give them’

Demonstratives. The distribution patterns shown in Table 5 demonstrate once again the existence of two models of structural variability among the nouns of the corpus. Unlike the results discussed in the analysis of the article, here the most striking fact is that, apart from code-switching nouns, the most frequent demonstrative forms are those belonging to each language. This occurs because the formal differences between the languages in this case do not affect the singular/plural dichotomy, as with the article, but the degrees into which the system of demonstratives is divided. Thus, for example, the radicals for the second and third level differ in Spanish and Catalan (Cat. *eixe*, Sp. *ese* ‘this’; Cat. *aquell* [əkél], Sp. *aquel* [akel] ‘that’). These differences in the feminine plural extend to all the degrees because of the usual Catalan morphophonological transformations (Cat. *estes*, Sp. *estas* ‘these’; Cat. *eixes*, Sp. *esas* ‘those’; Cat. *aquelles*, Sp. *aquellas* ‘those over there’).

This explains the fact that, when the demonstratives belong to one of these subparadigms, the speaker must necessarily use forms exclusive to each language, which allows us to distinguish between the grammatical variability of the types of nouns in which the corpus is divided.

In this context, it is revealing that, once more, it is only nouns in a Catalan context that co-occur with Catalan forms of the demonstrative, whether they are etymologically Catalan elements, as in (44), or ambiguous nouns, as in (45).

(44) I vosatros mireu *ixe programa* de *ixe* inclut (Catalan)
‘And you watch that ignorant man’s program’

(45) Perquè si *ixe grupo* que a mi no me coneixeu (ambiguous)
‘Because if that group that doesn’t know me’

Neither of these noun types uses exclusively Spanish demonstrative forms,¹⁵ in contrast to Spanish nouns (46) or code-switching nouns (47), which for their part never use Catalan demonstratives.

(46) *Con esta gente* se pueden hacer maravillas (Spanish)
 ‘With these people you can work wonders’

(47) La meua dona no volia, diu: **Antonio, que tenemos ese dinero para la casa**
 (code-switching)
 ‘My wife didn’t want to, she says, Antonio, we’ve got that money for our house’

Possessives. The two grammars of determination can be seen even more clearly in the case of possessives. It is here that we encounter the point of greatest structural conflict between the grammars of Catalan and Spanish, since their formal diversity is joined by an important syntactic difference. In Catalan the article and possessive adjective appear together before the noun, whereas in contemporary peninsular Spanish this combination is completely ungrammatical (Cat. *el seu home*; Sp. **el su marido, su marido* ‘her husband’) (Badia, 1962, 1994; Solà, 1973). Consequently, there are no forms common to both languages in this paradigm, so that the speaker is forced to choose one of the two grammars.

The results in this context are thus definitive: whereas Spanish nouns or code-switches always co-occur with Spanish possessives, as in (48) and (49), respectively, and never with the Catalan equivalent, the opposite is true for nouns in a Catalan linguistic context, regardless of their origin. Thus, both Catalan nouns and ambiguous nouns exhibit the forms and combinations characteristic of the Catalan possessive, as illustrated in (50) and (51), respectively.

(48) Gracias Alberto por haber cumplido *tu palabra* (Spanish)
 ‘Thank you, Alberto, for keeping your word’

(49) **que ni podía entrar ni era su mujer** i va i me diu (code-switching)
 ‘that I couldn’t come in and I wasn’t his wife, and he goes and says’

(50) Això es treballar pel *seu poble* (Catalan)
 ‘Now that’s working for your town’

(51) i en el *seu equipo* de govern (ambiguous)
 ‘and in his management team’

Only on one occasion (7.7%) does an ambiguous noun co-occur with an exclusively Spanish possessive form. This figure contrasts clearly with the remaining 92.3% of determination using Catalan possessive forms.

Other determiners. The remaining determiners (numerals, indefinites, quantifiers, and interrogatives), which we have grouped together in view of their lower frequency in discourse, also exhibit considerable grammatical differences between the classes of nouns. Once again we encounter Spanish nouns, with a balanced distribution between forms exclusive to this language (58.5%), as in (52), and forms common to Spanish and Catalan (41.5%), as in (53).

(52) Ha tenido ya acceso a *dos obras* (Spanish)
‘He’s already had access to two jobs’

(53) *Tres veces* a la semana (Spanish)
‘Three times a week’

These nouns are never preceded by one of the classes of other determiners with a purely Catalan form.

Code-switching nouns behave in exactly the same way, in almost the same proportions, using either exclusively Spanish forms (57.1%), as in (54), or forms common to Spanish and Catalan, as in (55).

(54) **te tenía que haber pegao dos *hostias* más**, en perdó de la paraula, com sona (code-switching)
‘I should have given you a couple more slaps, pardon my language, as it sounds’

(55) no pot ser, ***qué cuesta de enero*** (code-switching)
‘it’s not right, that straight after Christmas’

The grammatical behavior of both classes contrasts with that of Catalan nouns, with a remarkable use of forms belonging to Catalan (86.4%), as in (56), and only very occasional use of forms common to both languages (9.1%), as in (57).

(56) que el *atre dia* vaig portar el meu marit al metge (Catalan)
‘that the other day I took my husband to the doctor’s’

(57) Pero el Cara al Sol *alguna vegada* lo hauràs cantat? (Catalan)
‘But you must have sung Cara al Sol at some time?’

Between the two extremes are ambiguous nouns, but their distribution patterns are closer to Catalan nouns than to Spanish nouns or code-switching nouns. Although in this case the test for statistical significance reveals a greater difference between ambiguous and Catalan nouns, we should not ignore the fact that, in no less than 58.3% of cases, ambiguous nouns are preceded by a Catalan determiner from one of the paradigms included in this section, as in (58)—something that never occurs with Spanish nouns or code-switching nouns.

(58) El primer que se revente ***alguns globos*** ixé guanyarà (ambiguous)
‘The first person who bursts some balloons wins’

By comparison, ambiguous nouns using formally Spanish determiners, as in (59), make up a mere 16.6% of the total.

(59) sense parar in ***ningún puesto*** (ambiguous)
‘without stopping at any stall’

TABLE 6. *Frequency distribution of noun complements in the corpus*

	Catalan	Ambiguous	Spanish	Code-switching
	% (N)	% (N)	% (N)	% (N)
No complement	67.7 (151)	69.5 (612)	50.4 (110)	67.6 (73)
Preposed adjective	1.3 (3)	2.5 (6)	2.3 (5)	2.7 (3)
Postposed adjective	8.1 (18)	5.9 (14)	11.5 (25)	10.2 (11)
PP	13.9 (31)	13.7 (32)	18.8 (41)	11.1 (12)
Subordinate clause	6.3 (14)	6.8 (16)	11 (24)	6.4 (7)
Appositive	1.8 (4)	1.2 (3)	2.3 (5)	4.6 (2)
Multiple complements	0.9 (2)	0.4 (1)	3.6 (8)	0 (0)
Total	223	234	218	108

To sum up our results so far, we can say that in the corpus there is a homogeneous and regular distribution of the patterns of grammatical variability related to determination among the different groups of nouns considered in the study. This view is reinforced, as we have seen, by chi-squared tests, which confirm the parallelisms and differences indicated in almost every case.

It is striking that most of the determiners chosen belong to the group of forms common to both of the languages in contact. This occurs above all in the case of the articles, which is quite comprehensible in view of the fact that most of the nouns in the corpus appear in the singular, and that it is precisely here that the forms of the article in Spanish and Catalan are identical. However, when the noun is plural, the choices affect the forms exclusive to each language. In this case, there is a marked quantitative distance between the two groups of noun types: on the one hand, those which occur in a Catalan context, whatever their etymology, and, on the other hand, Spanish nouns and code-switching nouns.

With other types of determination (e.g., demonstratives, possessives, numerals, indefinites), however, the choices tend towards the exclusive forms, given their greater frequency in discourse compared with those common to both languages. In some cases, such as possessive determination, the selection of exclusive forms is categorical, since the points of conflict between Spanish and Catalan are most pronounced in this context. In these kinds of determination, the most important fact is that ambiguous nouns clearly pattern with Catalan nouns in terms of linguistic strategies and frequency of distribution. In short, ambiguous nouns adopt a clearly Catalan model of determination and distance themselves from Spanish, in direct contrast with Spanish nouns and code-switching nouns.

Complementation

Table 6 illustrates the frequency distribution of the different types of complementation examined in the present study: no complementation, preposed adject-

tive, postposed adjective, PP, subordinate clause, appositive, and multiple complementation.

Chi-squared statistics once again establish significant differences between Catalan and Spanish nouns ($p < .001$) and between ambiguous nouns and Spanish nouns ($p < .002$). There are no other statistically significant results.¹⁶

It is noteworthy that the most common syntactic framework for NPs in this discourse is absence of complementation, which comprises over half of the overall data for all groups. There is, nevertheless, a striking difference between the percentages for Spanish nouns and the other noun types. While the percentage of non-complementation for Spanish nouns barely exceeds 50% (50.4%), the other noun types have figures approaching 70%.

If we consider more closely the similarities and differences observed between the distribution patterns of the different groups, it is again Spanish nouns that stand out from the others. In addition to the differences in connection with non-complementation, they also behave differently in connection with PPs, subordinate clauses, appositives, and multiple complementation. Only in connection with adjectives, both preposed and postposed, is some quantitative convergence with code-switching nouns apparent. The data in Table 6 would also seem to indicate that, as with determination, Catalan nouns and ambiguous nouns show similar distributions, with the exception of postposed adjectivization, where the frequency of Catalan nouns slightly exceeds that of ambiguous ones.

More detailed analysis reveals that the similarity in distribution patterns extends to other groups. This means that the percentages for Catalan and ambiguous nouns are indeed very similar (the test for statistical significance gives the corresponding samples as significant: $p < .001$), but in some cases this similarity is shared by other nouns. This occurs mainly with code-switching nouns in connection with non-complementation, preposed adjectives, PPs, subordination, or multiple complementation. Only with appositives do the percentages between the groups separate slightly.

One possible interpretation of these data is that with complementation we are dealing with a continuum whose extremes are represented by the nouns of each language in their corresponding monolingual context, with the remaining two groups between the two extremes. Thus, the mixed nature of ambiguous nouns and code-switching nouns would justify this intermediate position, halfway between the grammars of Spanish and of Catalan.

However, there are important arguments against the hypothesis of a third grammar, which would characterize these elements in a bilingual context to a greater or lesser degree in the case of ambiguous nouns and code-switching nouns, respectively. Such a continuum would imply a steady gradation, which is not apparent in any of our data. Catalan, ambiguous, and code-switching nouns converge on similar figures, without any hierarchy among them. Of even greater importance is the fact that the chi-squared tests for the differences observed between ambiguous nouns and code-switching nouns, on the one hand, and Catalan nouns

TABLE 7. *Distribution of the number morpheme among the nouns of the corpus*

	Catalan	Ambiguous	Spanish	Code-switching
	% (N)	% (N)	% (N)	% (N)
Singular	74.4 (166)	84.6 (198)	78.9 (172)	86.1 (93)
Spanish plural	0.9 (2)	2.1 (5)	10.9 (22)	0 (0)
Catalan plural	20.7 (46)	0.4 (1)	0 (0)	0 (0)
Common plural	4 (9)	12.9 (30)	11.2 (24)	23.9 (15)
Total	223	234	218	108

and code-switching nouns, on the other, are not significant, so that they may be due to chance and related to fluctuations in the sample.

It should be borne in mind that, from a syntactic point of view, the structural similarities carry more weight in connection with complementation than do the differences. In contrast to determination, where there are certain points of structural conflict that in some cases affect a considerable part of certain paradigms (article, demonstratives, etc.) and in others affect the whole paradigm (possesives), the grammar of complementation is basically common to both languages. In other words, this factor is a point of convergence rather than conflict. It is this fact that is responsible for the lack of significant differences between the types of nouns rather than some kind of mixed grammar. In our view the absence of significant differences between the types of nouns in connection with complementation is mainly due to the fact that this is a context in which the similarities between the two languages carry more weight than do the differences.

MORPHOLOGICAL FACTORS

Number

Table 7 presents the frequency distribution of the number morpheme for the nouns in the corpus. It will be recalled that the forms of the singular are distinguished from the plural, and that the plural is classified according to a distinction among Spanish mechanisms of plural formation from the singular, structural procedures exclusive to Catalan, and finally forms common to both languages. The test for statistical significance gives significant differences for all the pairs of nouns studied. As expected, the principal differences in frequency among the groups appear among the forms of the plural. By contrast, the singular is the most common variant in discourse, with very high percentages for all the cases.

To form the plural, Catalan nouns mainly adopt the characteristic procedures of this language (80%), as in (60); only on a few occasions (15.5%) do they use ones common to both Spanish and Catalan, as in (61).¹⁷

(60) Perquè porte casi ja vint *dies* en el peu fastidiat (Catalan)
‘Because my leg’s been playing me up for nearly twenty days’

(61) Anem a donar els *premis* (Catalan)
‘Let’s give out the prizes’

There is also a slight presence of Spanish morphs in the formation of the plural in Catalan (4.5%). Thus, in (62), the speaker uses the form *-es*, characteristic of Spanish, instead of *-s*, which is obligatory in Catalan.

(62) I unes *oraciones*¹⁸ que jo quan me xitava (ambiguous)
‘And some prayers when I went to bed’

Spanish nouns never use morphs exclusive to Catalan. The formation of the plural is evenly distributed among exclusively Spanish variants (47.8%) and forms common to both languages (52.8%). Thus, in (63), we note the use of *-s*, obligatory in Spanish in nouns ending in unstressed *a*, which contrasts with the transformation into *-es* that occurs in Catalan in similar contexts.

(63) Lo que más me ha impresionado es que does *familias* (Spanish)
‘What most struck me is that two families’

In (64), the same morph *-s* is obligatory in both languages when the singular ends in unstressed *o*.

(64) Cuidar a los *ciudadanos*, que os quieren (Spanish)
‘Look after the citizens, they love you’

In the plural formation of ambiguous nouns, there appears to be a distribution intermediate between those considered up to now. Unlike Spanish nouns or code-switching nouns, which never adopt exclusively Catalan mechanisms of plural formation, ambiguous nouns do adopt them, although to a small extent (3.6%). In (65), the vowel *a* is transformed into *e* in the plural of a feminine noun.

(65) Las *llamades* eran totes per al defensor del poble (ambiguous)
‘The calls were all for the town’s ombudsman’

It should be noted, however, that this figure does not do justice to the importance of the application to these nouns of characteristically Catalan plural formation rules. This is the case because, in our initial selection of the items making up this noun type, all those with superficial signs of morphological integration in the

TABLE 8. *Distribution of plural markers among ambiguous nouns*

	% (N)
Spanish	9.4 (5)
Catalan	34 (18)
Common	56.6 (30)
Total	100 (53)

target language were omitted. We were led to make this decision by the fact that in the first stages of this study we had focused on the grammatical behavior of bare elements (i.e., ones without signs of phonological and/or morphological adaptation). Thus, for example, a noun like *bufandes* in (66) is an ambiguous noun in which the plural formation rule is characteristic of Catalan for nouns with unstressed final *a*.¹⁹

- (66) Li vaig fer unes *bufandes* per al hivern (ambiguous)
 'I made him some scarves for the winter'

In order to avoid skewing of the final interpretation of the data, we decided to carry out a new frequency count, this time taking into consideration an additional sample of ambiguous plural nouns whose plural reflects some degree of integration with the morphophonological rules of Catalan. The result of this new count is shown in Table 8. As the table reveals, no fewer than 34% of ambiguous nouns form their plurals following morphophonological rules exclusive to Catalan, whereas the use of mechanisms characteristic of Spanish represents only 9.4% of the total, a percentage far lower than that obtained for corresponding Spanish nouns (83.6%).

How can the high frequency (56.6%) of common solutions for the formation of the plural in these forms be explained?²⁰ Note that there is a certain parallelism between these nouns and code-switching nouns, which never exhibit forms exclusive to either Catalan or Spanish, so that all of the plurals corresponding to this group mark the plural via identical procedures in both languages. Might we then interpret this as a solution characterizing nouns in a bilingual context, whether they are isolated (i.e., ambiguous in our terms) or appear in code-switching contexts, which would distinguish these nouns from nouns in a monolingual context, whether Catalan or Spanish? If we answer this question in the affirmative, we are faced with a continuum of contact phenomena for plural forms similar to that formulated as a hypothesis earlier, which can be graphically represented as follows:

(Spanish grammar) Spanish nouns—code-switching nouns—ambiguous nouns—
 Catalan nouns (Catalan grammar)

Here, the extremes are occupied by nouns of Spanish and Catalan origin, which mark the plural mainly via the procedures of the relevant language. Between the two extremes, code-switching nouns and ambiguous nouns would represent intermediate solutions, although they would be closer to Spanish and Catalan nouns, respectively. The common denominator of these intermediate classes would be the use of mechanisms of plural formation common to both languages, with greater frequency the further they are from Catalan, as in the case of code-switching nouns. The high figures for exclusively Catalan solutions for ambiguous nouns would also seem to justify positioning them closer to the Catalan extreme.

Although superficially attractive, this interpretation does not appear to be valid for several reasons. It is hardly logical to assume that a speaker's choice of a noun from these supposed intermediate categories is due to the fact that he or she knows that the plural is formed using the same mechanisms in both languages. But in any case, Table 7 shows that common solutions are characteristic not only of code-switching nouns, but also of Spanish nouns. Note that the latter has a percentage of these forms (52.2%) that is even higher than that (47.8%) for exclusive Spanish markers.²¹

The reason for this convergence in the adoption of forms of the plural common to both languages lies in the fact that the majority of the items in the corpus, regardless of their context, have a phonological makeup which imposes the use of formal markers that are identical in both languages. The morph *-s* is the key here, as it is a very frequent plural marker in Spanish and the only one in Catalan. It is thus hardly surprising that there should be such high figures for this paradigm among the nouns considered here.

To sum up, we can infer that, while the differences between the types of Spanish noun, regardless of context, are attenuated at the points of structural convergence between the two languages, others which represent conflict sites give us a clearer glimpse of the linguistic affiliation of the lexical items. As far as plural formation is concerned, when the phonetic configurations of ambiguous nouns favor structural solutions common to both languages, their distributional pattern follows that of other classes. However, when the phonetic context favors the adoption of exclusive solutions, the linguistic affiliation of these nouns is closer to Catalan, since, as with the Catalan nouns, they tend to adopt formal procedures exclusive to Catalan.

Gender

Table 9 gives the data for the frequency distribution of the gender morpheme among the nouns in the corpus. As discussed earlier, gender is a grammatical morpheme which in both languages only affects the masculine/feminine opposition. It should also be remembered that, apart from a few special cases, the gender of a noun in one language is the same as in the other. Some of the figures in Table 9 should thus be discussed in the light of these facts.

To analyze the categorial affiliation of ambiguous nouns in relation to the grammars of Catalan and Spanish, we turn to one of the few points of struc-

TABLE 9. *Distribution of the gender morpheme of the nouns in the corpus*

	Catalan	Ambiguous	Spanish	Code-switching
	% (N)	% (N)	% (N)	% (N)
Masculine	60.5 (135)	75.7 (177)	48.1 (105)	56.5 (61)
Feminine	39.5 (88)	24.3 (57)	51.9 (113)	43.5 (47)
Total	223	234	218	108

tural conflict between the two languages exhibited by the gender morpheme: the procedures for feminine gender inflection, yielding the marked forms in the masculine/feminine opposition. Formal variation is not acceptable in either Spanish or Catalan for most feminine nouns, so that gender is given via syntactic procedures such as agreement with determiners and/or adjectives. Formally, however, a large percentage of these nouns also has, in both languages, the final *-a* marker.

Those nouns which do admit gender variation, whatever the semantic category giving rise to the opposition (differences of sex, size, etc.), use various formal mechanisms to carry it out. In addition to a few minority specialized suffixes in both languages (Cat. *-essa*, Sp. *-esa*), most of the elements in this paradigm resolve the question of marking through the addition of phonetic markers to the masculine form, and it is here that the differences between the two languages are to be found.

Unlike Spanish, Catalan in certain contexts adds phonetic modifications to the simple addition of *-a*, the procedure common to both languages. Thus, for example, when the relevant masculine form of a Catalan word ends in a stop or unvoiced sibilant, that consonant becomes voiced in the feminine, as illustrated in (67).

- (67) Me dia que jo era una *lloba* [λόβə] (masc. *llop* [λόπ]) (Catalan)
 'He told me I was a she-wolf.' (Sp. masc. *loba* [lóbo], fem. *loba* [lóba])

Thus, bearing in mind the aim of the present study to determine the grammar of ambiguous nouns, it is of interest to analyze in detail the formal procedures accompanying feminine forms. Figure 1 shows the percentages for the three possibilities considered: (1) feminine forms without gender variation ending in *-a*; (2) feminine forms without gender variation not ending in *-a*; and (3) feminine forms which admit gender variation.

The results show that significant differences between the grammars of Spanish and Catalan cannot be found on the basis of ambiguous nouns in the corpus analyzed because all the categories for marking gender are common to both languages. Thus, 89.4% of the feminine nouns do not admit gender variation, as

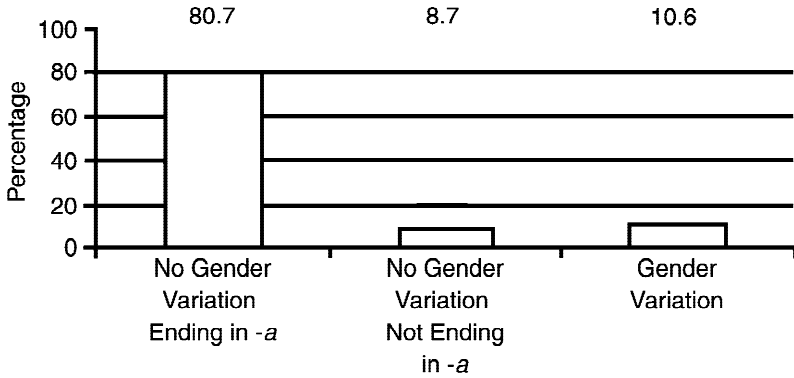


FIGURE 1. Distribution of the markers of feminine gender among the nouns of the corpus in all contexts.

opposed to only 10.6% that do. Of the former, the vast majority are nouns ending in *-a*, a marker for both Spanish and Catalan nouns, with 8.7% ending in phonemes other than *-a*. With a single exception (Sp. *calle* 'street'), these are nouns ending in a dental consonant which designate abstract entities (*libertad, entidad, enfermedad, particularidad* 'rights, entity, illness, peculiarity'), which in Spanish are feminine, exactly like the corresponding nouns in Catalan.

Finally, the nouns admitting gender variation in all cases correspond to forms in which the masculine/feminine opposition reflects sexual differentiation (*chica, compañera, peluquera, novia* 'girl, female workmate, female hairdresser, girlfriend'). Similarly, for all the examples, final *-a* appears as marker of the feminine, together with other features of agreement. More importantly, this marking in no case affects phonetic contexts in which Catalan not only adds the *-a*, but also alters the phonetic configuration of the previous consonant, since here the morph *-o* of the masculine is replaced by the *-a* of the feminine.

To sum up, it can be inferred that, precisely because of the high degree of structural similarity between the two languages involved in contact, no significant differences can be discerned between the types of nouns considered in the present study. Only at those points where the grammars of the two languages show some structural differences in the system for marking the feminine might there appear additional data of interest. However, such data do not appear in our corpus, since the procedures found in all cases are identical in both languages. In any case, this is no more than a consequence of the vast frequency of these mechanisms in both Spanish and Catalan discourse.

CONCLUSION

Both the variationist approach and the comparative method which we have used in the present investigation lead us to some conclusions about the status of am-

biguous nouns. The originality of these conclusions lies mainly in the fact that they have been obtained through applying a rigorous method to a bilingual Catalan–Spanish spoken corpus. The two main conclusions are as follows. First, ambiguous nouns present a pattern of grammatical variability different from that of nouns of Spanish origin, whether these appear in a monolingual Spanish or in a code-switching context. Second, the grammar of ambiguous nouns, in contrast to Spanish nouns and code-switching nouns, resembles that of Catalan nouns.

These two conclusions are based on a comparison of the patterns of grammatical distribution in those areas of the grammar where the differences between the two languages, Spanish and Catalan, are structurally clearest. In contrast, the differences diminish, or simply disappear, in the analysis of other factors where the similarities have greater weight than the differences.

In the development of the comparative method applied in the present study to the Catalan–Spanish context, the epistemological opposition between structural conflict sites and coincidence sites has a decisive importance, as our data have indeed confirmed. Although we have commented on these concepts throughout this study, we now focus in more detail on their main theoretical implications.

As Poplack and Meechan (1998) pointed out, if two languages exhibit parallel structures at certain coincidence sites, it is impossible to decide which of the two grammatical systems bilingual utterances derive from. We have observed that this occurs in our analysis of nouns viewed through the lenses of complementation and gender. Among the syntactic categories, complementation is one in which Spanish and Catalan are closest structurally: the languages share the same categories (non-complementation, multiple complementation, preposed and postposed adjectives, PPs) at similar levels of frequency, and the differences in many cases are not statistically significant. Similar conclusions can be drawn for the morphological factor of gender. In both Catalan and Spanish, gender opposition is limited to the masculine/feminine terms, and any discrepancies between Spanish nouns and their corresponding Catalan forms are simple exceptions. Only in certain aspects of the gender paradigm, such as the mechanisms for marking the feminine gender, do the grammars of the two languages diverge. However, all the data of our corpus fall into the area of common structures.

How, then, can we elucidate the grammatical status of lone nouns—even bare nouns—from one language in the discourse of another? To quote Poplack and Meechan (1998:132): “To unambiguously ascertain the language membership of the bilingual construction, we need to examine at which the structures of the language pair do not match, i.e. sites where the grammars of the two languages in contact *conflict*.” That is, when two languages differ with regard to certain structural features, whether morphological or syntactic or both, it is possible to assess the linguistic status of lone lexical elements in the discourse of the other language by comparing the distribution patterns of this grammatical variability. This can be carried out through contrasting the distribution patterns of nouns in a monolingual context with the distribution patterns of nouns of the same language origin in an unambiguous code-switching context.

It is therefore not sufficient to demonstrate that lone elements share the same grammatical restrictions as the corresponding lexical categories in the target language. To affirm that the former share the same grammatical system as the latter, they must do so in similar proportions.

Several researchers have shown the way in which this analytical model can be used to study isolated lexical elements in a target language. In a pioneering study, Poplack and Meechan (1995) demonstrated that, in Wolof–French contact, bare nouns of French origin exhibited the same grammatical restrictions as nouns belonging to Wolof, and that they differed notably from French grammar both in monolingual contexts and in multiple code-switching. Budzhak-Jones and Poplack (1997) discovered a similar distributional pattern in English–Ukrainian contact. Budzhak-Jones (1998) noted that non-inflected English nouns in Ukrainian discourse were subject to restrictions in gender inflection similar to those found for corresponding Ukrainian nouns in nonstandard varieties.

Other studies have examined various contact situations. In most of these studies, the structural conflict sites in the morphology of the languages involved have enabled researchers to reach a conclusion similar to ours: that lone nouns from the donor language in the discourse of the receiving language exhibit a high level of grammatical integration, even in the absence of markers of superficial integration, and so they are considered borrowings and not code-switches. This is the conclusion reached, for example, by Adalar and Tagliamonte (1998) on plural formation of nouns in an English–Turkish context. Ghafar Samar and Meechan (1998) found similar results for marking of reference in English–Persian, as did Eze (1998) for the inflection of English verbs in Igbo discourse.

These morphological criteria are strengthened by others of a syntactic nature. This is the case of earlier studies with factors such as the markers of case and agreement in English–Ukrainian (Budzhak-Jones, 1998) or word order within the noun phrase, as in Adalar and Tagliamonte (1998) for English–Turkish or Ghafar Samar and Meechan (1998) for English–Persian.

The main conclusions of these researchers can be summarized as follows. First, contrary to what some researchers claim, the phenomena of code-switching and lexical borrowing are essentially different. Second, lone lexical material most frequently occurs in the form of borrowings, behaving grammatically in an identical way to unambiguous borrowings, which are sufficiently settled and linguistically and socially integrated in the receiving language.

It is within this hermeneutic context that the present study is located. However, unlike the majority of the studies mentioned, the situation of contact in this case involves two languages with numerous points of structural convergence. As far as we know, the comparative method has only been applied in one context similar to ours: namely, in Turpin's (1995, 1998) studies of French–English contact in bilingual speech communities in Canada. Turpin (1998) showed how morphological factors such as plural marking or syntactic factors like determination or word order in the NP permit the claim that most lone English nouns in a French context reflect French rather than English grammar.

However, despite some structural similarities between English and French, no study has tackled the problem in “sister” languages like Spanish and Catalan. Derived from the same Romance branch and historically developing in intense contact, Spanish and Catalan share numerous structural similarities. It is a commonplace, for example, that the main differences between the two are concentrated at the phonetic level, whereas in syntax and vocabulary the similarities are more numerous than the differences. As far as syntax is concerned, Catalan and Spanish exhibit some idiosyncratic structures that distance them from each other and bring them closer to other Romance languages.²²

In this context, the interest of the present study lies in comparing the grammar of these structural coincidence sites with that of the sites of conflict with the aim of elucidating the status of ambiguous nouns. We have thus considered factors like complementation and gender, where the two languages coincide, along with others, such as determination and plural marking, where they do not. As far as complementation is concerned, it should be remembered that, together with some common areas with coinciding forms, there are others exhibiting formal and/or syntagmatic differences. Conversely, in the category of number and plural formation, Catalan and Spanish also display points of conflict.

The data supplied by the factors involving structural conflict sites between the two languages coincide, as has been demonstrated. The distribution patterns of the different noun types considered here coincide in those subparadigms that are formally and functionally identical in the two languages. However, they differ significantly when the grammars diverge. For example, in the case of determination, the most frequent occurrences in discourse, in categories like the article, correspond to those forms of the singular that are identical in Catalan and Spanish. This would be explained by the fact that the majority of the nouns in the corpus appear in the singular. However, with the plural forms of the article, the differences between the noun types become more clearly defined, and the nouns fall into two groups: those in a Catalan context, regardless of their etymological origin, and those of Spanish origin, whether in a Spanish monolingual or in a code-switching context. While the members of the second group always follow the restrictions characteristic of the determination of the Spanish NP, the nouns in the first group obey the grammatical rules of Catalan, regardless of their etymological origin.

The results are even clearer in those areas of determination where the differences between the grammars of both languages are greater, as in the case of the possessives. Here, the speaker’s choice always affects forms exclusive to each language. Whereas Spanish nouns or code-switching nouns never resort to Catalan possessives, ambiguous nouns do so systematically.

These hypotheses were also confirmed in the analysis of plural marking. Thus, the distribution patterns of the nouns are closer for those points of the number morpheme where differences are not noted, which would explain the preference of all groups for the singular forms. However, within the plural forms, in which there are structural conflict sites between Spanish and Catalan, we should stress two facts that condense what we have seen so far. On the one hand, we have found

that, whereas Spanish nouns and code-switching nouns never use Catalan rules for marking the plural, ambiguous nouns do and in percentages which should not be underrated. Nevertheless, in this category we find that all the groups of Spanish-origin nouns coincide quantitatively in their use of common forms. Contrary to the thesis of a third grammar to explain the status of nouns in a code-switching context, we have found that the proximity of the distribution patterns of these nouns is a result of the fact that most of them have a phonetic configuration that necessarily involves marking the plural through procedures shared by Spanish and Catalan.

In view of these data, we may conclude that the characterization of lone nouns in Catalan as borrowings is basically correct, and we may reject the alternative view of them as code-switching. In other words, Spanish-origin nouns in a Catalan linguistic context are adapted to the receiving language (Catalan), even in the absence of any surface indications that this is so.

Budzhak-Jones and Poplack (1997:252), who reached similar conclusions in their study of bilingual English–Ukrainian discourse, argued that native speakers interpret nouns from another source, even when they are not (yet) consolidated and widely used by the society, as fully integrated in the receiving language. In this respect, it is interesting to observe that the now classic criterion of integration is undergoing a notable epistemological review in relation to the idea popularized almost three decades ago by Mackey (1970). As we saw earlier, the criterion of frequency in discourse as well as other linguistic criteria have demonstrated their limited use in the numerous corpora in which their application has been attempted. By contrast, the variationist comparative method considers that integration can only be measured in a rigorous empirical manner, based on models of linguistic variability in a bilingual corpus such as the one analyzed in this article.

NOTES

1. For scientific reasons, we have chosen to use the term “Catalan” to refer to the regional variety analyzed here, popularly known as Valencian. In this regard, it should be noted that during recent years secessionist attempts supported by the Valencian right wing have vociferously expressed their disagreement with the judgments of the scientific community.

2. Although the main interest of this study lies in the analysis of bare nouns, which do not exhibit superficial signs of linguistic integration because of the considerable structural proximity of the two languages in contact, in some areas of our research we use a small additional sample of lone nouns which already show certain signs of morphological adaptation. This is the case, for example, with the analysis of the plural morpheme of Spanish nouns in a Catalan context, since it would otherwise have been impossible to complete the comparative study.

3. A possible exception is the work of Turpin (1995, 1998), which analyzed this matter in French–English bilingual discourse in an Acadian corpus. Nevertheless, the structural proximity between those languages is less than that between two Romance languages, such as Spanish and Catalan.

4. Unlike other typologically different pairs of languages, such as those analyzed in previous studies, in the case of Catalan–Spanish contact the only differences are formal or combinatorial, since the same functional categories operate within the NP in both languages.

5. For Spanish, apart from classic grammars like those of Bello (1847) and the Spanish Academy (R.A.E., 1931, 1973), which serve as a basis for others, we considered the following general works:

Gili Gaya (1961), Alarcos (1972, 1995), Seco (1972), Alcina and Bleuca (1975), Marcos Marín (1980), Pilleux and Urrutia (1982), Hernández (1984), Fernández Ramírez (1986), Hernanz and Brucart (1987), López (1998), Marcos Marín, Satorre, and Viejo (1998), and Bosque and Demonte (1999).

For Catalan, there are fewer grammatical works, although in recent times a few grammars have appeared, along with monographs which, in greater or lesser detail and within different theoretical frameworks, have treated the subject from a synchronic viewpoint. Among them, the following can be highlighted: Salvà (1830), Sanchis-Guarner (1950), Fabra (1956), Badia (1962, 1994), Solà (1973, 1977, 1983, 1987, 1994), Barberà (1983), Marvà (1984), Bonet and Solà (1986), Valor (1987), Payrató (1988), Institut d'Estudis Valencians (1990), Lacreu (1992), Ramos (1993), and Rodríguez-Vida and Rúaix (1997).

In addition to the above general bibliography, there are many studies focusing on particular aspects of the grammar of both languages, particularly for Spanish. Some of these studies will be referred to as appropriate, although our interest here is centered on the more purely comparative aspects rather than on the often controversial interpretations of the different areas of morphosyntax made by scholars of the two languages. From a contrastive point of view, we make mention of the following works for the syntax of NPs for both Catalan and Spanish grammars: Jordana (1933), Palet (1987), Piot (1993), Vallduví (1994), Gorski et al. (1996), and Lloret and Viaplana (1997).

6. Catalan feminines beginning with an atonal *i* or *u* (*la idea* 'the idea', *la unitat* 'the unit') are an exception. Spanish has the peculiarity of adopting the masculine form of the article before words beginning with atonal *a* (*el agua* 'the water'), unlike Catalan (*l'aigua*).

7. This is the Valencian variant of the preposition (Valor, 1987).

8. Nevertheless, the demonstratives show a neuter form when they are the head of the NP (Sp. *mira esto* 'look this'; Cat. *aquello/dona'm esto* 'give me this').

9. In Old Spanish, the combination of article + possessive was frequent, although unlike contemporary Catalan the corresponding possessive form was always the atonic and not the tonic form (Lapesa, 1971). In certain areas of northern Spain, there are still dialectal traces of this combination (see Álvarez, 1986).

10. By multiple complementation we mean those cases in which the head of the NP is complemented by two or more of the syntagmatic categories discussed.

11. Payrató also noted the influence of Spanish in the regularization of the creation of singulars (*pantaló*), despite the fact that the pressure to regularize is the same for both languages, so that in his opinion, "Attribution of such cases to a single exclusive cause is risky and often unjustified" (1988:96). López del Castillo (1976) increased the social range of these forms to include the educated nonstandard level of contemporary Catalan, particularly in words ending in *ig*, *sc*, and *st* (see also Fabra, 1956).

12. Figures cited henceforth refer to the percentages of each linguistic variable. To discover the relation of these figures to absolute frequency, the relevant table should be consulted.

13. The grammatical categories on which the analysis focuses are shown in italics, whereas bold-face is used to highlight material of Spanish origin in Catalan discourse. In each case, every noun is underscored.

14. There is only one isolated case which could be attributed to linguistic error on the speaker's part, as it occurs curiously in the group of Catalan nouns and not among the ambiguous nouns.

15. The fact that the only case appears among Catalan nouns and not among ambiguous nouns suggests that this may be an isolated case, produced by an error in linguistic performance.

16. Some of the figures for code-switching nouns are low, which may affect the validity of the significance test in the comparisons involving this noun group.

17. These figures are not based on Table 7, since they are calculated exclusively from the plural forms. Thus, the 46 occurrences of Catalan plural forms represent 80% of the Catalan nouns, and the 9 appearances of nouns with common plurals represent 15.5%.

18. The word is pronounced with the interdental /θ/ of Spanish.

19. The corresponding Spanish form would be *bufandas*, without modification of the final vowel.

20. It is this figure, which contrasts with the much more modest figure for Catalan nouns, that makes the global differences observed between the two classes of nouns statistically significant.

21. This figure is obtained by calculating the percentages affecting exclusively the plural forms: 22 forms exclusive to Spanish (47.8%) vs. 24 common forms (52.2%).

22. Traditionally, the links between Catalan and French have been stressed, giving rise to a wide-ranging controversy in which some claim that the proximity between the two is greater than that between Catalan and Spanish.

REFERENCES

- Adalar, N., & Tagliamonte, S. (1998). Borrowed nouns, bilingual people: The case of the Londrali in the Northern Cyprus. *International Journal of Bilingualism* 2:139–159.
- Alarcos, E. (1972). *Estudios de gramática funcional del español*. Madrid: Gredos.
- (1995). *Gramática de la lengua española*. Madrid: Espasa Calpe.
- Alcina, J., & Blecua, J. M. (1975). *Gramática española*. Barcelona: Ariel.
- Álvarez, M. A. (1986). *El artículo como entidad funcional en el español de hoy*. Madrid: Gredos.
- Backus, Ad. (1992). *Patterns of language mixing: A study in Turkish–Dutch bilingualism*. Wiesbaden: Harrassowitz.
- Badia, A. (1962). *Gramática Catalana*, Madrid: Gredos.
- (1994). *Gramàtica de la llengua catalana descriptiva, normativa, diatòpica, diastràtica*. Barcelona: Enciclopèdia Catalana.
- Barberà, J. (1983). *Curs de gramàtica normativa per a ús dels valencians. Grau superior*. València: ICE.
- Bello, A. (1847). *Gramática de la lengua castellana destinada al uso de los americanos*. Santiago de Chile: EDAF. (New edition, with notes by R. J. Cuervo, Bogotá, 1874).
- Bentahila, A., & Davies, E. (1983). The syntax of Arabic–French codeswitching. *Lingua* 59:301–330.
- (1991). Constraints on code-switching: A look beyond grammar. *Papers for the Symposium on Code-Switching in Bilingual Studies. Theory, significance and perspectives*. Barcelona: European Science Foundation. 369–403.
- Berk-Seligson, S. (1986). Linguistic constraints on intrasentential code-switching: A study of Spanish/Hebrew bilingualism. *Language in Society* 15:313–348.
- Boeschoten, H. (1990). Asymmetrical code-switching in immigrant communities. *Papers for the Workshop on Constraints, Conditions and Models*. London: European Science Foundation. 85–100.
- Boeschoten, H., & Verhoeven, L. (1985). Integration niederländischer lexikalischer Elemente ins Türkische. Sprachmischung bei Immigranten der ersten und zweiten Generation. *Linguistische Berichte* 98:347–364.
- Bokamba, E. (1988). Code-mixing, language variation, and linguistic theory: Evidence from Bantu languages. *Lingua* 76:21–62.
- Bonet, S., & Solà, J. (1986). *Gramàtica generativa de la llengua catalana*. Barcelona: Edicions de la Abadía de Monserrat.
- Bosque, I., & Demonte, V. (eds.). (1999). *Gramática descriptiva de la lengua española*. Madrid: Espasa-Calpe.
- Budzhak-Jones, S. (1998). Against word-internal code-switching: Evidence from Ukrainian English bilingualism. *International Journal of Bilingualism* 2:161–182.
- Budzhak-Jones, S., & Poplack, S. (1997). Two generations, two strategies: The fate of bare English-origin nouns in Ukrainian. *Journal of Sociolinguistics* 1/2:225–258.
- Eze, E. (1997). *Aspects of language contact. A variationist approach to codeswitching and borrowing in Igbo–English bilingual discourse*. Doctoral dissertation, University of Ottawa.
- Fabra, P. (1956). *Gramàtica catalana*. Barcelona: Teide.
- Fernández Ramírez, S. (1986). *Gramática española*. Madrid: Arco/Libros.
- Gardner-Chloros, P. (1991). *Language selection and switching in Strasbourg*. Oxford: Clarendon.
- Ghafar Samar, R. (1996). *It's really null. Persian–English language contact and the null theory of code-switching*. Paper presented at NWAVE-XXV, University of Nevada.
- Ghafar Samar, R., & Meechan, M. (1998). The null theory of codeswitching versus the nonce borrowing hypothesis: Testing the fit in Persian–English bilingual discourse. *International Journal of Bilingualism* 2:203–219.
- Gili Gaya, S. (1961). *Curso superior de sintaxis española*. Barcelona: Bibliograf.
- Gorski, R., et al. (1996). Rodzaj żywotny/osobowy: dzieli czy łączy języki? Analiza kontrastywna wybranych języków słowiańskich i romańskich na przykładzie rzeczownika. *Biuletyn Polskiego Towarzystwa Językoznawczego/Bulletin de la Société Polonaise de Linguistique* 52:127–134.
- Gumperz, J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.
- Hernández, C. (1984). *Gramática funcional del español*. Madrid: Gredos.
- Hernanz, M., & Brucart, J. (1987). *La sintaxis. Principios teóricos. La oración simple*. Barcelona: Crítica.
- Institut d'Estudis Valencians. (1990). *Gramàtica de la llengua valenciana*. València: IEV.
- Jordana, C. (1933). *El català i el castellà comparats*. Barcelona: Barcino.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

- _____. (1997). *The language of life and death*. Paper presented at NWAWE-XXVI, University of Quebec.
- Lacreu, J. (1992). *Manual d'ús de l'estandard oral*. València: Servei de Publicacions, University of Valencia.
- Lapesa, R. (1971). Sobre el artículo ante posesivo en el castellano antiguo. In *Sprache und Geschichte. Festschrift Harri Meier*. Madrid: Gredos. 277–296.
- Loret, A. R., & Viaplana, J. (1997). On the morphological category of gender in Catalan and in Spanish. In W. U. Dressler, M. Prinzhorn, & J. R. Rennison (eds.), *Advances in morphology*. Berlin: Mouton de Gruyter. 171–188.
- López, A. (1998). *Gramática del español III*. Madrid: Arco Libros.
- López del Castillo, L. (1976). Les interferències català-castellpa. In *Bilingüismo y educación en cataluña*. Barcelona: Teide. 87–91.
- Lorenzo, G. (1995). *Geometría de las estructuras nominales. Sintaxis y semántica del SDET*. Oviedo: Universidad de Oviedo.
- Mackey, W. F. (1970). Interference, integration and the synchronic fallacy. In J. E. Alatis (ed.), *Bilingualism and language contact: Anthropological, linguistic, psychological, and sociological aspects*. Washington, DC: Georgetown University Press. 195–227.
- Marcos Marín, F. (1980). *Curso de gramática española*. Madrid: Cincel.
- Marcos Marín, F., Satorre, F., & Viejo, M. L. (1998). *Gramática española*. Madrid: Síntesis.
- Marvà, J. (1984). *Curs pràctic de gramàtica catalana, grau superior*. Barcelona: Barcino.
- Meechan, M., & Poplack, S. (1995). Orphan categories in bilingual discourse: Adjectivization strategies in Wolof–French and Fongbe–French. *Language Variation and Change* 7:169–194.
- Muysken, P. (1987). Neutrality in code mixing. In F. J. von Ingen, B. P. F. Al, S. R. Slings, J. H. Hulstijn, & C. Blotkamp (eds.), *Eigen en vreemd. identiteit en ontlening in taal, literatuur en beeldende kunst. Handelingen van het 39ste Nederlands Filologencongres*. Amsterdam: Vrije Universiteit Uitgeverij. 359–373.
- Myers-Scotton, C. (1993). *Duelling languages*. Oxford: Clarendon Press.
- Palet, M. T. (1987). Un problema de lingüística contrastiva: Las preposiciones en español y catalán. *Revista de la Sociedad Española de Lingüística* 17:69–84.
- Payrató, L. L. (1988). *El català col.loquial*. Valencia: University of Valencia.
- Picallo, C. (1991). Possessive pronouns in Catalan and the avoid pronoun principle. *Catalan Working Papers in Linguistics*. 211–234.
- Pilleux, M., & Urrutia, H. (1982). *Gramática transformacional del español*. Madrid: Cincel.
- Piot, M. (1993). Recherche et méthodologie en syntaxe comparée des langues romanes. *Studia Romanica Posnaniensia* 17:175–187.
- Poplack, S. (1990). Variation theory and language contact, concept, methods and data. In *Papers for the Workshop on Concepts, Methodology and Data*. Strasbourg: European Science Foundation. 33–66.
- _____. (1993). Variation theory and language contact: Concepts, methods and data. In D. Preston (ed.), *American dialect research*. Amsterdam: Benjamins. 251–286.
- Poplack, S., & Meechan, M. (1995). Patterns of language mixture: Nominal structure in Wolof–French and Fongbe–French bilingual discourse. In L. Milroy & P. Muysken (eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. Cambridge: Cambridge University Press. 199–232.
- _____. (1998). How languages fit together in code-mixing. *International Journal of Bilingualism* 2:127–138.
- Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26:47–104.
- Poplack, S., Wheeler, S., & Westwood, A. (1987). Distinguishing language contact phenomena: Evidence from Finnish–English bilingualism. In P. Lilius & M. Saari (eds.), *The Nordic languages and modern linguistics*. Helsinki: University of Helsinki Press. 33–56.
- R. A. E. (1931). *Gramática de la lengua castellana*. Madrid: Espasa-Calpe.
- _____. (1973). *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe.
- Ramos, J. R. (1993). *Introducció a la sintaxi*. València: Tandem.
- Rand, D., & Sankoff, D. (1990). *GOLDVARB: A variable rule application for the Macintosh, Version 2*. Montreal: Centre de recherches mathématique, Université de Montréal.
- Ritter, E. (1991). Two functional categories in noun phrases: Evidence from Modern Hebrew. In S. Rothstein (ed.), *Perspectives on phrase structure: Heads and licensing*. San Diego: Academic. 37–62.

- Rodríguez-Vida, S., & Ruaiç, J. (1997). *El català de la A a la Z*. Moia: Ruaiç.
- Romaine, S. (1989). *Bilingualism*. Oxford: Blackwell.
- Salvà, V. (1830). *Gramàtica de la lengua castellana según ahora se habla*. Valencia: Mallén.
- Sanchis-Guarner, M. (1950). *Gramàtica valenciana*. Valencia: Torres.
- Seco, M. (1972). *Gramàtica esencial del español*. Madrid: Destino.
- Solà, J. (1973). *Estudis de sintaxi catalana* (vols. 1 and 2). Barcelona: Edicions 62.
- (1977). *A l'entorn de la llengua*. Barcelona: Edicions 62.
- (1983). Qüestions sintàctiques pendents en la normativa actual. *Actes del 6 Col.loqui Internacional de Llengua i Literatura catalanes*. Barcelona: Publicacions de l'Abadia de Montserrat. 153–194.
- (1987). *Qüestions controvertides de sintaxi catalana*. Barcelona: Edicions 62.
- (1994). *Sintaxi normativa estat de lo qüestió*. Barcelona: Empuries.
- Tió, J. (1982). *L'ensenyament del català als no—catalanoparlants*. Vic: EUMO.
- Treffers-Daller, J. (1990). Towards a uniform approach to code-switching and borrowing. In *Papers for the Workshop on Concepts, Methodology and Data*. London: European Science Foundation. 259–277.
- (1994). *Mixing two languages*. Cambridge: Cambridge University Press.
- Turpin, D. (1995). 'Le français c'est le last frontier'. *Le syntagme nominal dans le discours bilingue français/anglais*. Master's thesis, University of Ottawa.
- (1998). 'Le français c'est le last frontier': The status of English-origin nouns in Arcadian French. *International Journal of Bilingualism* 2:221–233.
- Vallduví, E. (1994). Detachment in Catalan and information packaging. *Journal of Pragmatics* 22: 573–601.
- Valor, E. (1987). *Curs Mitjà de gramàtica catalana referida especialment al País Valencià*. València: Papers Basics.
- Vigara, A. (1992). *Morfosintaxis del español coloquial*, Madrid: Gredos.