International Journal of Applied Linguistics • Vol. 16 • No. 2 • 2006

Connecting linguistic description and language teaching: native and leaner use of existential *there*[1]

Ignacio Palacios-Martínez *Universidade de Santiago de Compostela*
Ana Martínez-Insua *Universidade de Vigo, Spain*

This article emerges from the need to connect linguistic theory and language teaching to find concrete solutions to problems Spanish students confront when learning English. This study looks at existential *there* constructions taken from the following native and non-native written English corpora: the International Corpus of Learner English and the Santiago University Learner of English Corpus for the non-native set, and the Louvain Corpus of Native English Essays, Biber et al. (1999) and a subcorpus of the BNC for the native English group. This contrastive study reveals important differences in the use of *there* constructions as regards their frequency, structural complexity, polarity and pragmatic value. Important implications for the presentation and the pedagogical treatment of the *there* constructions can be derived from the results.

Keywords: existential constructions, learner corpus, language teaching, contrastive linguistics, corpus linguistics

El presente artículo surge de la necesidad de conectar la teoría lingüística y la práctica pedagógica, tratando de encontrar soluciones concretas a problemas con los que se enfrentan alumnos españoles de inglés como lengua extranjera. Este trabajo estudia las construcciones existenciales con *there* (CTs) a partir de los siguientes corpus de textos escritos de hablantes nativos y no nativos: International Corpus of Learner English y Santiago University Learner of English Corpus para los no nativos, y Louvain Corpus of Native English Essays, Biber et al. (1999) y un subcorpus del BNC para los nativos. Este estudio contrastivo constata diferencias importantes en el uso de las CTs relativas a su frecuencia, complejidad estructural, polaridad y valor pragmático. De todos estos resultados se derivan importantes implicaciones para la presentación y tratamiento pedagógico de las CTs.

Palabras clave: construcciones existenciales, corpus de estudiantes, enseñanza de lenguas, lingüística contrastiva, lingüística de corpus

Introduction

Those of us who devote our careers to both language teaching and research in linguistics face the problem of how to draw direct connections between

IJAL_114 | Pages: 19

these two main facets of our profession; in other words, how our investigations can provide useful guidelines which may illuminate our daily teaching. We believe that by carefully examining the data provided by our students of English and comparing it with the language typical of native speakers, we can draw our learners' attention more easily to the difficulties they may have when learning English. This approach is particularly applicable to those grammatical structures of the English language that differ considerably from their counterparts in the students' L1, in our case, Spanish. Existential *there* constructions (e.g. *There were a lot of guests at the royal wedding*) are a case in point.

From our teaching experience we know that Spanish students of English have certain difficulties in the use of existential sentences. Elementary learners are prone to confuse *it* with *there*, like other learners of different language backgrounds do, such as Norwegian students. There is even a tendency to use the possessive *their* instead of *there*, as these two words are pronounced the same and elementary-level learners are not able to distinguish them from the context. Intermediate and advanced students for their part do not have serious problems with *there* constructions, although they do not exploit the pragmatic possibilities to their full extent.

We hope that this article will throw light on the teaching of *there* constructions. Traditionally these structures have not been given a special place in grammar teaching in general and, in our view, they require specific pedagogical treatment. Moreover, to our knowledge, no large-scale investigations of *there* constructions in any language have been conducted up to now with particular reference to language teaching. Lastly, it may be the case that some of our conclusions can also be extrapolated to other thematically marked structures, such as clefts and extraposed clauses.

Review of the literature

Existential *there* constructions have been the object of a wide number of studies conducted from both a synchronic and a diachronic perspective (Milsark 1979; Hannay 1985; Breivik 1990 [1983]; Aniya 1992; Johansson 1997; Pérez-Guerra 1999). Some linguists have concentrated, for example, on the characteristic syntactic features of these structures and have considered existential *there* as a dummy element (Radford 1997; Huddleston and Pullum 2002), as the subject of these sentences (Downing and Locke 1995; Biber et al. 1999), or even as a simple pro-word (Pérez-Guerra 1999; Huddleston and Pullum, 2002). Other scholars, such as Erdman (1976), Breivik (1990 [1983]), and Huddleston (1988), have tried to explain the semantics of existential *there* by contrasting it with the locative adverb *there*. Most of these semantic accounts seem to agree that existential *there* originated from the locative form. In addition to this, *there* constructions have also been examined from a pragmatic perspective. *There* constructions are commonly viewed in the

literature as the product of derivations or movements of elements from their canonical positions (Huddleston 1971: 322, 1988: 184; Quirk et al. 1985). Taking this line, the existential and non-existential counterparts of these structures have been regarded as two thematic variants expressing the same propositional content (Huddleston 1988: 173; Gómez-González 2000: 245ff), with users choosing between them depending on the part of the message they want to emphasise or on what they consider as known to the addressee. However, this is not the only pragmatic function of *there* constructions, as these constructions are very commonly used to introduce new information into the discourse (Quirk et al. 1985; Givón 1993; Downing and Locke 1995), to focus on the existence or occurrence of something, and even to bring something to mind rather than stating explicitly that something really exists (Biber et al. 1999: 951ff).

The large number of studies conducted on numerous syntactic, semantic and pragmatic features of these constructions in L1 English are in clear contrast with the sparse research carried out to date on this issue from the point of view of both first and second language acquisition. This contrast is even more noticeable if we consider that the literature on language acquisition contains abundant examination of the developmental patterns of several other syntactic structures, such as interrogatives, negatives and relative clauses.[2]

Among the few studies on the acquisition[3] of *there* constructions is Johnson (2001), which concentrates on the connections between English deictic expressions with *there* (e.g. *There sits Peter*) and existential *there* constructions (e.g. *There's a dog in the yard*) in child first language acquisition. After analysing three different corpora from the CHILDES archive (MacWhinney 1995), Johnson concludes that "children base the central existential construction on the central deictic" (2001: 131). In fact, children seem to go through three main stages in the acquisition of *there* sentences: in the first stage, *there* is used as a deictic locative both in initial and final position (e.g. *There lion*; *Tower right there*), whereas in the second, deictics overlap with indefinite NPs and final phrases (e.g. *There's one for you*; *There's a radio over there*). In the third and final stage, *there* constructions are clearly distinguished from deictic ones (e.g. *Are there more than there*? *There was a big kangaroo*). Johnson then draws a parallel between what he calls *constructional grounding* (the use of occurrences of one construction as the basis for the acquisition of another construction) and *historical reanalysis*, which consists of assigning a new meaning to a particular structure as presented in a particular context.

Johansson and Lysvåg (1987: 321ff) and Hasselgård, Johansson and Lysvåg (1998: 326) in their contrastive Norwegian-English grammars for students of English explain how Norwegian learners tend to confuse the dummy subjects *it* and *there*, which both correspond to *det* in Norwegian. The following example reported by Johansson and Lysvåg (1987: 321) clearly shows *it* used instead of *there*:

1) In our complicated society **it** is a great need for skilled and educated workers.

Apart from these contributions, to our knowledge, nothing else on the acquisition of *there* constructions has been carried out. This explains our interest in exploring the learning of *there* constructions by Spanish students of English as a foreign language.

The study

Purpose

The focus of this article is the close analysis of *there* constructions from the perspective of second language learning. We are also concerned with comparing the use of these structures by native and non-native speakers of English – Spanish learners in this case. We hypothesise that differences will be identified between these two groups as regards frequency of use, the accompanying verb, the complexity of *there* constructions, polarity, concord and pragmatic value. We have selected these specific features of *there* constructions because they are the most widely discussed in the literature. It is interesting to know how non-natives of English make use of *there* constructions because, on the one hand, identifying differences can lead to remedial teaching; on the other hand, non-native data can also provide relevant information about the processing of language by native speakers and the evolution and development of this English construction. Native data cannot supply this information without longitudinal studies of the acquisition of these grammatical structures by native speakers at different points or periods in time, which is beyond the scope of this study.

Materials and method

We used data from five different corpora as the basis of our analysis. For the non-native data, the Spanish component of ICLE (International Corpus of Learner English) and part of SULEC (Santiago University Learner of English Corpus) were used. For the native English data, we took material from three sources: LOCNESS (Louvain Corpus of Native English Essays), Biber et al. (1999), and Martínez-Insua (2004). Although not all the corpora selected are directly comparable in terms if their internal organisation and general structure, we feel that conclusions can be drawn from their comparison.

The Spanish component of ICLE (Granger 1998a) contains about 200,000 words produced by university-level advanced students of English, and it consists of extracts from argumentative and literary essays. For the purpose of this study, only the argumentative compositions were considered, a total

of approximately 125,000 words. SULEC is a learner corpus currently being compiled by Palacios-Martínez and his team at the University of Santiago[4] of written and spoken language produced by intermediate and advanced Spanish secondary and university learners of English. For the purpose of this study, a sample of 100,000 written words was selected. As in ICLE, the essays these were extracted from discussed current controversial issues, such as the role of the monarchy in modern countries, same-sex marriages, the pros and cons of smoking policies, the usefulness of university degrees, and the necessity and value of armies.

The LOCNESS corpus contains native speaker (British and American) argumentative essays and exceeds 320,000 words. It is thus comparable with the two non-native corpora. Biber et al. (1999) used the LSWE (Longman Spoken and Written English) corpus to produce their well-known corpus-based grammar. This corpus contains 40 million words of modern British and American English. It comprises samples of many different registers and styles, and it is not limited to any particular genre. However, we have focused only on data extracted from the written part of the corpus. Martínez-Insua (2004), following Butler (1985) in her statistical approach, used a stratified sample of approximately 500,000 words both of spoken and written English, consisting of text-type passages of the same length and drawn from the BNC (British National Corpus). For the purpose of this study, only the written sample of 505,534 words was considered. Martínez-Insua (2004) provides figures and tables for syntactic, semantic and pragmatic features of *there* constructions.

As the size of the samples varied in each of these five databases, normalised frequencies were calculated when required. The use of these five primary sources allows us to discuss a large number of comparable native and non-native *there* constructions, since the samples share a number of common features: medium of expression (written language), variety of English (modern English and primarily British English), genre (argumentative in the case of LOCNESS, SULEC and ICLE), and education level in the native-speaker corpora.

The material extracted from ICLE, SULEC and LOCNESS was processed with the aid of the Concapp4 concordancer. However, it was also necessary to filter the data since this tool produced right and left concordances with all occurrences of *there* in the corpora, whether locative or existential in meaning. All the sentences containing a locative *there* were excluded from our study since we are only interested in existential *there* constructions here, but we looked at some unclear cases in the non-native data, such as the following:

2) Many of these women gave up **there** children for adoption. (LOC-302)[5]
3) so that I think that **there** more be places for. (SUL-356)

In example (2) it is not clear whether *there* has a locative meaning ('gave up at that place') or whether the writer really meant its homonym *their*. In (3)

Table 1. Frequency of *there* constructions in the different corpora

Corpus	total words	*there* constructions	ratio per 10,000 words
ICLE	125,550	475	37.80
SULEC	100,000	395	39.50
subtotal, non-native	*225,550*	*870*	*38.65*
LOCNESS	324,134	998	30.70
LSWE (Biber et al. 1999)	40,000,000	250,000 (approx.)	27.70
BNC subcorpus (Martínez-Insua 2004)	505,534	1,071	21.10
subtotal, native	*40,829,668*	*252,069*	*24.40*

the sentence is ungrammatical as it stands, and it is not at all clear what the learner actually meant. It should, however, be pointed out that the number of unclear cases was not high and did not affect our findings to any significant degree.

Analysis and discussion of results

Frequency of *there* constructions

The results in Table 1 show that *there* constructions tend to be more common in the Spanish learners' written interlanguage than in native speakers' writing (approximately 39 vs. 24 per 10,000). The figures for both ICLE and SULEC are strikingly similar, whereas there is wider variation found within the three corpora of native English.

In our view, three main observations could help explain the higher frequency of *there* constructions in this non-native versus native use:

a) *There* constructions are generally introduced at early stages of English language learning in Spain. An overview of textbooks for the teaching of English at elementary levels in Spain clearly shows that *there* constructions are usually presented in the first ten units.
b) We have observed that *there* constructions are learnt as prefabricated or formulaic language; that is, Spanish students may learn *there is/are* structures as chunks or fixed expressions.[6]
c) The expression of existence and the introduction of new entities into discourse are common communicative functions in everyday language in both Spanish and English. It could be the case that learners draw a direct and simple connection between English *there* constructions and Spanish existential *hay* constructions.

It would obviously be necessary to conduct further studies to be able to confirm these hypotheses.

The verb in *there* constructions

As can be seen in Table 2, which gives the most common verbs in *there* constructions in the corpora, the verb *be* is present in the vast majority in both native and non-native written English. The differences identified across the five corpora are not significant. Complex verb phrases (semimodals/catenatives) with *be* are the most common alternatives to *be* in the native-speaker data, although their frequency is very low.[7]

The verb *exis*t deserves an independent analysis since it is found five times in one non-native and one native corpus (ICLE and LOCNESS) but not in the other three. It might be the case that this use of *exist* is connected with the argumentative genre more than with any other specific factor. The figures also show, as could easily be expected, that native speakers of English have in their personal repertoire a wider variety of presentational or existential verbs than non-native learners of English. These presentational verbs are employed to introduce slight nuances of meaning. This explains why nine of the simple and complex verbs are not found in the Spanish learners' English.

The use of *have* as main verb instead of *be* in *there* constructions produced by non-native learners is also worth mentioning. The two occurrences of *have* are due to language transfer. The Spanish verb *haber* 'have' is used in existential constructions (as well as functioning as a tense auxiliary); the Spanish version of the existential sentence in (4) would have *habrá*, leading the learner to use *have* instead of *be*:

4) If I am not mistaken in the next few days there will **have** criminal sanctions against all these responsible of pollution. (SUL-429)

Language transfer may also be responsible for other special uses of *there* constructions in the learner data. Several examples were found where Spanish learners confused *it* and *there* as dummy subjects. A similar phenomenon was mentioned above for Norwegian learners (cf. example 1); however, this works in the opposite direction here as the Spanish learners tended to use dummy *there* where the correct form would be *it* (examples 5–6). This leads us to suspect that this kind of confusion may also occur in other non-native Englishes.

5) To sum up, we can say that rehabilitation system is an utopian and rather absurd project insofar as **there** is not clear who must be rehabilitation: whether criminals or society itself. (ICL-29)
6) the main reason is that **there** should be necessary a lot of money to maintein it, and not all nations can afford it. (ICL-81)

Table 2. Distribution of the most common verbs in *there* constructions

Verb	Non-native				Native					
	ICLE		SULEC		LOCNESS		LSWE		BNC subcorpus	
	no.	%	no.	%	no.	%	no.	%	no.	%
simple										
be	467	*98.3*	394	*99.8*	978	*98.0*	248,125	*99.25*	1,068	*99.73*
exist	5	*1.1*	–	–	5	*0.5*	–	–	–	–
have	1	*0.2*	1	*0.2*	–	–	–	–	–	–
cease	–	–	–	–	1	*0.1*	–	–	–	–
need	–	–	–	–	2	*0.2*	–	–	–	–
remain	–	–	–	–	1	*0.1*	–	–	1	*0.09*
follow	–	–	–	–	–	–	–	–	1	*0.09*
appear	–	–	–	–	–	–	375	*0.15*	–	–
develop	–	–	–	–	–	–	–	–	1	*0.09*
complex										
seem to be	1	*0.2*	–	–	7	*0.7*	875	*0.35*	–	–
used to be	1	*0.2*	–	–	–	–	375	*0.15*	–	–
have to be	–	–	–	–	2	*0.2*	–	–	–	–
tend to be	–	–	–	–	2	*0.2*	–	–	–	–
be supposed to	–	–	–	–	–	–	250	*0.1*	–	–

Table 3. Extensions in the *there* constructions, in percentages

Corpus	ø	rel. cl.	pp	non-finite cl.	adj. p	*that* cl.
SULEC	21.3	45.1	21.5	8.20	2.50	1.4
ICLE	21.02	32.24	29.8	12.8	1.81	2.33
subtotal, non-native	*21.2*	*38.7*	*25.7*	*10.5*	*2.2*	*1.70*
LOCNESS	23.94	37.78	18.33	13.23	2.81	3.91
BNC (subcorpus)	29.42	13.17	32.31	18.21	1.97	4.92
subtotal, native	*26.68*	*25.49*	*25.32*	*15.72*	*2.39*	*4.4*

It may also be the case that learners do not make a clear distinction between locative and existential *there*, and this lack of distinction may be responsible for their confusion with *it*.

The structure of *there* constructions

We focus our attention now on the structure of the *there* constructions, considering the grammatical units that both precede and follow the existential marker *there*. The results obtained in our analysis of native and non-native data confirm that *there* generally occupies initial position in *there* clauses. However, *there* is not initial in cases of emphatic inversions produced by a negative or restrictive element, as in (7):

7) Only when this occurs **is there** any possibility that computers may make our . . . (LOC-11)

Regarding the complexity of the *there* constructions, we would expect learners to make use of more basic and less complex structures than native speakers. Analysis of the types of extensions in the notional subject of the *there* constructions indicates that in both the native and non-native data the patterns are similar, although they differ in frequency (Table 3). Such extensions can take the form of relative clauses, prepositional phrases, non-finite clauses, adjective phrases, and embedded *that* clauses. In contrast to these complex *there* constructions, we also found a number of minimal *there* constructions (ø in the table) in which the postverbal nominal phrase does not have any kind of postmodification or complementation.[8] It should be pointed out that Table 3 does not include data from LSWE as Biber et al. provide percentages only for prepositional phrases and relative clauses as structural expansions in the notional subject of existential clauses (1999: 949). Furthermore, in their account comparisons are made among different genres of speech and writing (conversation, fiction, news and academic prose), which do not make them comparable to the rest of our samples.

The following examples illustrate each of the categories in Table 3:

ø

8) . . . think that there are more disadvantages than there are advantages, thus they are likely to reject . . . (LOC-12)
9) the majority of people are christians but, there are a lot of criminals, so . . . (ICL-18)

relative clauses

10) Although far less popular, there are other sports **which are far more brutal**. (LOC-172)
11) I think, that nowadays there are a lot of crimes **that don't pay**. (ICL-17)

prepositional phrases

12) Throughout the first scenes, especially, there is a repetition **of "rien"**, perhaps reflecting . . . (LOC-468)
13) So there is a waste **of food** and a waste **of money**. (ICL-296)

non-finite clauses

14) There are a lot of things **to be taken into consideration**. (LOC-21)
15) . . . people being persuaded by the television, there is another point **to be considered**. (ICL-309)

adjective phrases

16) . . . to conceive a child in this manner and that there are other alternatives **available**, for example . . . (LOC-170)
17) There are fewer women **involuntarily unemployed**. (ICL-70)

that clauses

18) There is no way **that people should allow their** . . . (LOC-663)
19) It is obvious that in that places there is the thought **that the prison system is outdated** . . . (ICL-422)

The figures contained in Table 3 allow us to talk of certain similarities and differences between the native and non-native speakers' corpora. In both, relative clauses and prepositional phrases are the most common postverbal extensions of *there* constructions; however, the number of relative structures is considerably higher in the learner data. Next come non-finite clauses in both samples, although the proportion of these structures is slightly higher in the native-speaker data. Adjective phrases and *that* clauses have a very low frequency of occurrence in both sets.

The frequency of minimal and non-minimal *there* constructions provides a further source of difference between the native and non-native samples. The non-native speakers in the corpora tend to use non-minimal, extended *there* constructions more frequently than the native speakers do, which goes against our assumption that we would find more basic constructions in the learner data. This may be the result of the way in which *there* constructions

Table 4. Distribution of adverbials in the *there* constructions, in percentages

Corpus	Initial position	Final position
SULEC	62.90	37.10
ICLE	70.18	29.82
subtotal, non-native	*66.54*	*33.46*
LOCNESS	44.87	55.13
BNC (sub)	44.88	55.12
subtotal, native	*44.87*	*55.13*

are generally taught in the context exemplified in SULEC, i.e. as devices for drawing long and heavy subjects towards final position (in keeping with the End-Weight Principle) rather than as signalling a number of other pragmatic functions independent of the length and/or complexity of their notional subjects (cf. Martínez-Insua 2004 for a detailed account of the pragmatic values that *there* constructions can convey).

Turning our attention to the distribution of the adverbial expressions in *there* constructions in the different corpora, Table 4 shows that non-native speakers have a clear tendency to place adverbials in clause initial position rather than in final position, while native speakers do not show a strong preference either way. The figures for the two native-speaker corpora are strikingly similar. It might be inferred that these Spanish learners of English prefer to establish the circumstantial framework for the real subject before introducing it into the discourse. This could be conditioned by their native language, as there is a tendency in the use of Spanish existential sentences to state the contextual situation before referring to the topic in question.

Concord

The lack of concord between the verb and notional subject of *there* constructions has also been a recurrent issue lately in the specialized literature on *there* constructions. Some studies (Givón 1993; Meechan and Foley 1994; Biber et al. 1999; Martínez-Insua 2004; Martínez-Insua and Palacios-Martínez 2003; Crawford 2005) have shown a growing tendency for a lack of concord in present-day English among native speakers; this is particularly evident in the oral medium, as Table 5 illustrates. As mentioned above, a sample of 1,000,000 words drawn from the BNC was used as the basis for this analysis.

Biber et al. (1999: 186) do not provide concrete figures or data with regard to this issue, but they affirm that examples with a singular form of *be* followed by plural noun phrases "are somewhat more common in conversation than the standard constructions with plural verb plus plural noun phrase". Biber claims this is especially true with the contracted form *there's*, since in speech this is processed as if it were an individual item.

Table 5. Concord vs. non-concord in *there* constructions produced by native speakers of English in Martínez-Insua and Palacios-Martínez (2003)

Medium of expression	Concord	Non-concord
Writing	96.78%	3.22%
Speech	86.74%	13.26%
Total	*90.75%*	*9.25%*

In the LOCNESS native-speaker corpus, based exclusively on written language, non-concord is noticeably less frequent than in Martínez-Insua and Palacios-Martínez's data (2003). In fact, only nine examples of non-concord in *there* constructions were found out of a total of 997 tokens, which amounts to only 0.9%. The fact that this corpus consists only of written academic essays produced by students with a university education in a formal setting may explain this. It is interesting that six of these nine cases of non-concord occur with auxiliary *has*, e.g.:

20) Amongst recent issues **there has been** big **problems** in . . . (LOC-353)
21) **There has been** many **objections** to that . . . (LOC-360)

Quite surprisingly, the figures obtained for second language learners do not greatly differ from those attested in this sample of the BNC corpus. The contrast between writing and speech cannot be determined here since the corpora selected for this study only draw on written language.

Table 6 shows that the figures for concord in the ICLE and SULEC non-native speaker corpora are not the same: non-concord is more frequent in SULEC. The fact that some samples drawn from SULEC come from intermediate learners may explain this. It is logical to think that less advanced students of English should make more mistakes than advanced learners as regards concord in *there* constructions.

Although the results do not reveal important differences between native and non-native use with respect to the concord of the notional subject of the *there* constructions with the verb, the explanations for this phenomenon are of a different nature in each situation. In the case of Spanish learners of English, the loss of verbal agreement has to do with their incomplete knowledge of the subject–verb agreement rule. It could also be related to a lack of enough practice with these structures, or even to the fact that it takes studens a long time to assimilate *there* constructions even though they are usually introduced quite early on in instruction. It should be borne in mind that English in Spain, as in most Spanish-speaking countries, is a foreign language rather than a second one. This means that we cannot speak of natural or spontaneous acquisition.

Table 6. Concord vs. non-concord in *there* constructions produced by Spanish learners in the ICLE and SULEC corpora

Corpus	Concord		Non-concord	
	no.	**%**	**no.**	**%**
ICLE	465	97.89	10	2.11
SULEC	377	95.40	18	4.60
Total	*842*	*96.60*	*28*	*3.40*

In several examples it is clear that the student does not know that the notional subject noun phrase is uncountable and, consequently, that the verb should be singular, e.g.:

22) In public places **there are** so much **smoke**. (SUL-196)
23) In the world **there are liberty**. (SUL-99)
24) In all places **there are** very **contamination**: cars (SUL-234)

At times learners have problems because they are not familiar with the rule that governs verbal concord with certain nouns. This is the case, for example, with *people*, which takes plural concord in English. In Spanish, in contrast, the corresponding word *gente* takes singular concord. The following two examples exemplify this problem:

25) In my friend's group, **there is** some **people** that smoke. (SUL-333)
26) **There is** still so many **people** who smoke. (SUL-335)

In the advanced (university-level) group of learners in the ICLE corpus, we also find instances of non-concord with quantifiers, e.g.:

27) And **there are a few soldier** who works . . . (ICL-12)
28) Long ago **there was plenty of "ladrones"**. . . (ICL-473)
29) On the house **there weren't any book**. (SUL-422)

This tendency for a lack of concord is particularly noticeable in the case of negative polarity (as in example 29). This is especially so in the ICLE corpus, where one third of the *there* constructions with non-concord are negative, e.g.:

30) It's not because **there aren't money** but because . . . (ICL-206)
31) but this doesn't mean that **there aren't place** for dreams. (ICL-207)
32) . . . result useless and too that **there aren't place** for them in this. (ICL-208)

If we look at these last three examples carefully, we see that they have a feature in common: the *there* constructions are part of an embedded sentence (introduced by *because* or *that* here), which in Spanish would require a subjunctive verb. The lack of a direct structural correspondence between the Spanish subjunctive and English may make it more difficult for the Spanish learners to express this meaning in English and could thus contribute to concord errors.

Although the number of examples studied in this preliminary analysis is not sufficient to reach categorical conclusions, we could hypothesize from the two last sets of examples that there is a connection between the lack of concord and the negative polarity factor for these learners. This could be explained by the well-known fact that it is more complex to process negative than positive sentences, to which is added the difficulty of getting concord correct when producing *there* constructions.

In contrast, for the native speakers the examples with lack of concord cannot be accounted for by an incomplete or imperfect knowledge of a grammatical rule; in this case we have to refer to the tendency of language users to simplify the linguistic code.[9] SMS language and the code used in Internet communication are good examples of this. We could also hypothesize that non-concord within *there* constructions, which seems to be relatively common in speech, is being extrapolated to writing as well. Such a development would not be new, as written language is becoming more and more influenced by spoken language, especially in informal registers (Biber 1988; Cornbleet and Carter 2001).

Polarity

The figures for the number of positive and negative polarity *there* constructions in both groups of corpora are quite similar, so we cannot speak of significant differences between native and non-native use regarding this feature. However, an issue that deserves our attention is the high number of *there is no X* combinations in both the native and non-native data, as illustrated in the following examples:

33) **There is no doubt** that theoretical classes are necessary. (SUL-307)
34) It is said that nowadays **there is no place** for imagination. (ICL-352)
35) **There is no reason** for fear. (ICL-358)
36) **There is no point** on spending five years at university . . . (SUL-311)

Curiously enough, lack of concord was not recorded with any of these negative collocations. This finding seems to reinforce the hypothesis that students may learn them as prefabricated or fixed expressions (see note 6).

Table 7. Word combinations with *no* in *there* constructions in the native and non-native data

Collocation	Non-native data	Native data
no doubt	7	3
no place	6	7
no value	6	5
no (good) reason	4	10
no point	2	–
no need	2	8
no sense	1	–
no use	1	–
no comparison	1	–
no difference	1	–
no (clear-cut) way	–	7
no evidence	–	3
no mention	–	3
no black or white area	–	2
no advantage	–	2
no chance	–	2
no room	–	2
no cure	–	2
no guarantee	–	2
no hope	–	2
no possibility	–	2
no proof	–	2
no control	–	2
nothing wrong	–	2
no attempt	–	1
no sign	–	1
no link	–	1
no fear	–	1
no interest	–	1
no wonder	–	1
no excuse	–	1

Table 7 lists the relevant collocations identified as well as the number of occurrences in the different corpora. As expected, the number of such word combinations with existential *there* constructions is much higher in the native-speaker data, since native speakers possess a richer and more varied repertoire of vocabulary. Note that these combinations do not occur with *not*, only with *no*. This may be explained by the fact that such *not* constructions (e.g. *there can not be any doubt*, *there is not a reason*) are syntactically more complex, needing an auxiliary and sometimes a quantifier, or both, and consequently they present more problems for Spanish learners than their counterparts with *no*.

Conclusions

The information gathered here will inform our teaching and can be used to provide teachers of English with concrete guidelines they can follow in the presentation of these constructions.

As regards the frequency of *there* constructions, the data clearly show that these structures are more frequent in the written English of these Spanish students than in the native English written discourse. The introduction of *there* constructions at an early stage of the learning process, the fact that the expression of existence is basic in any language, the assumption that these constructions could be learnt as prefabricated chunks, and language transfer were hypothesized to be possible factors that may account for this.

The verb *be* was the most common, in terms of frequency, in *there* constructions in both the native and non-native data. As expected, native speakers exhibited a wider range of other presentational verbs. It is evident that teachers, especially at intermediate and advanced levels, should introduce a wider variety of existential or presentational verbs apart from *be*, and they should also refer to the specific nuances of meaning conveyed by them as well as to the registers and styles where they would be more suitable. Concrete instances extracted from corpora or reflecting real language use could also be very helpful for that purpose.

From this data we can also conclude that Spanish teachers of English should pay more attention to the pragmatic meanings expressed by existential *there* structures, that is, they should make their students aware of the different discourse functions of *there* constructions apart from the typical one of introducing new elements in the discourse, such as focusing on the occurrence or non-occurrence of something, preparing the ground for what is coming next in the sentence, and anticipating a series or list of items. In this vein, it would be useful to work with corpora data to exemplify how native speakers of English use *there* constructions according to their pragmatic needs.

These conclusions are not definitive. Further research still needs to be carried out to investigate questions like the pragmatic value of *there* constructions in native and non-native language, the stages of acquisition of these constructions longitudinally, their particular characteristics in speech and in different registers and genres. It would also be interesting in the future to consider spoken non-native data to see whether the tendency towards lack of agreement between the notional subject and the main verb identified for native speakers has its counterpart in Spanish learners of English. We also presume that important differences could also be identified if we contrasted second language versus foreign language data.

There are some obvious pedagogical implications of these findings. Spanish teachers of English could, for example, draw students' attention to the most common mistakes identified in the use and form of *there* constructions, such as the lack of concord, problems with word order, and confusion between *have* and *be* and between *it* and *there* as notional subjects.

Learners should be made aware of the fact that existential *there* as a grammatical subject has no lexical content. In addition, teachers of English could focus in particular on the two main uses of *there* as locative and existential. Examples taken from real data could be presented; students could then be asked to classify the examples into two different groups according to the function and meaning of *there*.

We also believe that some of these findings could be applied to other languages and learners (cf. the Norwegian parallel of *it*/*there* confusion). In addition, we hypothesize that some of the features related to the pragmatic use of *there* constructions may also apply to other thematically marked structures, such as dislocations, *it* and *when* clefts, extraposed and passive sentences. Finally, we contend that producers and developers of English language materials should take this information into account in the preparation and design of grammars, textbooks, dictionaries, glossaries, and readers. The inclusion of examples of learner data could certainly have a very beneficial pedagogical function.

We hope this study has contributed to a better and more comprehensive characterization and view of the learning of English as a foreign language by Spanish students. We also hope the completion of the learner corpus we are currently compiling will help us to conduct other projects along these directions in the future.

Notes

1. We would like to express our gratitude to Professor Leiv Egil Breivik, who supplied us with the contrastive Norwegian reference materials, and to both him and an anonymous referee for their useful suggestions, comments and recommendations, most of which have been incorporated in the final version of this article. We also want to thank Jean Hannah for her careful and thorough revision of the original manuscript. The limitations that remain are all our own responsibility. The research reported on in this article was funded by the Galician Ministry of Education (grant no. PGIDIT05PXIB20401 PR) and is gratefully acknowledged.
2. See Ellis (1994: 73–117) for a thorough account of this.
3. Although some scholars, such as Krashen (1983, 1988), make a clear distinction between the concepts of acquisition and learning, we will use them as synonyms here.
4. More detailed information about the purpose and characteristics of this corpus can be obtained at the following web pages: <htttp://www.usc.ia303. SULEC/SULeC.htm> and <http://sulec.cesga.es>. The corpus currently contains approximately 400,000 words of written and spoken learner material, but the aim is to collect 1,000,000 words.
5. The abbreviations and numbers in brackets after each of the examples stand for the corpus from which they were extracted. Thus LOC, ICL and SUL represent LOCNESS, ICLE and SULEC, respectively. The numbers give the document reference. All the examples are given as found in the original database.

6. The study of the acquisition of formulae or formulaic language has attracted growing attention in the last few years, as it has practical implications for both language learning theories and language teaching methodology. For a complete review of this area, see Nattinger and Decarrico (1992), Weinert (1995) and Granger (1998b). De Cock et al.'s study (1998) is quite innovative, as they investigate formulaic language in French EFL learners by automatically extracting recurrent word combinations from corpora of informal speech.
7. Following Biber et al. (1999), complex verb phrases like *there will be* and *there have been* were listed along with *be* as main verb rather than in the group of semimodals and catenatives.
8. A distinction should be drawn between 'bare existentials' (Quirk et al. 1985; Givón 1993; Huddleston and Pullum 2002) and minimal *there* constructions (Martínez-Insua 2004). 'Bare existentials' refers to *there* constructions that lack an SVX counterpart. In contrast, minimal *there* constructions are those lacking any kind of postmodification or complementation in their postverbal nominal structure.
9. For a review of the factors that favour non-concord in native *there* constructions, see Martínez-Insua and Palacios-Martínez (2003: 280).

References

Aniya, S. (1992) The semantics and the syntax of the existential *there*-construction. *Linguistic Analysis* 22.3/4: 154–84.

Biber, D. (1988) *Variation across speech and writing*. Cambridge University Press.

—— , S. Johansson, G. Leech, S. Conrad and E. Finegan (1999) *The Longman grammar of written and spoken English*. London: Longman.

Breivik, L.E. (1990 [1983]) *Existential there. A synchronic and diachronic study*. (Studia Anglistica Norvegica.) Bergen University: Department of English.

Butler, C. (1985) *Statistics in linguistics*. Oxford: Blackwell.

Cornbleet, S. and R. Carter (2001) *The language of speech and writing*. London: Routledge.

Crawford, W.J. (2005) Verb agreement and disagreement. A corpus investigation of concord variation in existential *there + be* constructions. *Journal of English Linguistics* 33.1: 35–61.

De Cock, S., S. Granger, G. Leech and T. McEnery (1998) An automated approach to the phrasicon of EFL learners. In S. Granger (ed.), *Learner English on computer*. London/New York: Longman. 67–79.

Downing, A. and P. Locke (1995) *A university course in English grammar*. London: Phonenix Elt.

Ellis, R. (1994) *The study of second language acquisition*. Oxford University Press.

Erdman, P. (1976) *The sentences in English*. München: Tuduv Studie.

Givón, T. (1993) *English grammar. A function-based introduction*. Amsterdam/ Philadelphia: John Benjamins.

Gómez-González, A. (2000) *The theme-topic interface: evidence from English*. Amsterdam: John Benjamins.

Granger, S. (1998a) *Learner English on computer*. London/New York: Longman.

—— (1998b) Prefabricated patterns in advanced EFL writing: collocations and formulae. In A.P. Cowie (ed.), *Phraseology. Theory, analysis and application*. Oxford: Clarendon Press. 145–60.

Hannay, M. (1985) *English existentials in functional grammar*. Dordrecht-Holland: Foris.
Hasselgård, H., S. Johansson and P. Lysvåg (1998) *English grammar: theory and use*. Oslo: Universitetsforlaget.
Huddleston, R. (1971) *The sentence in written English*. Cambridge University Press.
—— (1988) *Introduction to the grammar of English*. Cambridge University Press.
—— and G. Pullum (2002) *The Cambridge grammar of the English language.* Cambridge University Press.
Johansson, S. (1997) A corpus study of English existential clauses: register variation and discourse function. In T. Nevalainen and L. Kahlas-Tarkka (eds.), *To explain the present. Studies in the changing English language in honour of Matti Rissanen.* Helsinki: Société Néophilologique. 303–18.
—— and P. Lysvåg (1987) *Understanding English grammar*. Oslo: Universitetsforlaget.
Johnson, C. (2001) Constructional grounding: on the relation between deictic and existential *there*-constructions in acquisition. In A. Cienki, B. Luka and M.B. Smith (eds.) *Conceptual and discourse factors in linguistic structure.* Stanford: CSLI Publications. 123–36.
Krashen, S.D. (1983) *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
—— (1988) *Second language acquisition and second language learning*. New York: Prentice Hall.
MacWhinney, B. (1995) *The CHILDES project. Tools for analyzing talk*. Hillsdale, NJ: Lawrence Erlbaum Associates.
Martínez-Insua, A. (2004) *Existential there-constructions in contemporary British English*. Muenchen: Lincom Europa.
—— and I. Palacios-Martínez (2003) Non-concord in existential *there*-constructions. *English Studies* 84.3: 262–83.
Meechan, M. and M. Foley (1994) On resolving disagreement: linguistic theory and variation. *There's bridges. Language variation and change* 6: 63–85.
Milsark, G.L. (1979) *Existential sentences in English*. New York: Garland.
Nattinger, J.R. and J.J. Decarrico (1992) *Lexical phrases and language teaching*. Oxford University Press.
Pérez-Guerra, J. (1999) *Historical English syntax. A statistical corpus-based study on the organisation of early modern English sentences*. Muenchen: Lincom Europa.
Quirk, R., G. Leech, S. Greenbaum and J. Svartvik (1985) *A comprehensive grammar of the English language*. London/New York: Longman.
Radford, A. (1997) *Syntactic theory and the structure of English. A minimalist approach*. Cambridge University Press.
Weinert, R. (1995) The role of formulaic language in second language acquisition: a review. *Applied Linguistics* 16.2: 180–205.

e-mails: iafeans@usc.es, minsua@uvigo.es *[Final version received 22/3/06]*

MARKED PROOF

Please correct and return this set

Please use the proof correction marks shown below for all alterations and corrections. If you wish to return your proof by fax you should ensure that all amendments are written clearly in dark ink and are made well within the page margins.

Instruction to printer	*Textual mark*	*Marginal mark*
Leave unchanged	··· under matter to remain	Stet
Insert in text the matter indicated in the margin		New matter followed by
Delete	through matter to be deleted	
Delete and close up	through matter to be deleted	
Substitute character or substitute part of one or more word(s)	/ through letter or through word	New letter or new word
Change to italics	— under matter to be changed	
Change to capitals	≡ under matter to be changed	
Change to small capitals	= under matter to be changed	
Change to bold type	~ under matter to be changed	
Change to bold italic	≈ under matter to be changed	
Change to lower case	Encircle matter to be changed	
Change italic to upright type	(As above)	
Insert 'superior' character	/ through character or where required	under character e.g.
Insert 'inferior' character	(As above)	over character e.g.
Insert full stop	(As above)	
Insert comma	(As above)	
Insert single quotation marks	(As above)	and/or
Insert double quotation marks	(As above)	and/or
Insert hyphen	(As above)	
Start new paragraph		
No new paragraph		
Transpose		
Close up	linking letters	
Insert space between letters	between letters affected	
Insert space between words	between words affected	
Reduce space between letters	between letters affected	
Reduce space between words	between words affected	