Language contact in Gibraltar English: A pilot study with ICE-GBR*

Lucía Loureiro-Porto
lucia.loureiro@uib.es
Cristina Suárez-Gómez
cristina.suarez@uib.es
Universitat de les Illes Balears

ABSTRACT

The variety of English used in Gibraltar has been in contact with a number of European languages, such as Spanish, Italian, Hebrew and Arabic (Moyer, 1998: 216; Suárez-Gómez, 2012: 1746), for more than 300 years. Studies of this variety have traditionally been based on interviews and observation (e.g. Moyer, 1993, 1998; Cal Varela, 1996; Levey, 2008 2015; Weston, 2011, 2013, etc.), and a detailed morphosyntactic description is yet to be published. In this context, the compilation of a reliable Gibraltar corpus using the standards of the *International Corpus of English* (ICE) will constitute a landmark in the analysis of this lesser known variety of English. In the present paper we describe the ICE project and the current state of the compilation of ICE-GBR. In addition, we present a detailed comparison between the section on press news reports of ICE-GB (standard British English) and ICE-GBR, with the aim of identifying morphosyntactic features that reveal the influence of language contact with Spanish in this territory. We explore variables such as the choice of relativizer (assuming a higher preference for *that* in GBR, in agreement with Spanish *que*, the most frequent relativizer, Brucart, 1999: 490), the use of titles and pseudo-titles preceding proper names (which, as shown by Hundt and Kabatek, 2015, are very frequent in English journalese and extremely infrequent in Spanish), and the frequency of the passive voice (expected to be lower in ICE-GBR), among others. A preliminary analysis of these variables reveals that the influence of Spanish on the variety of English used in the Gibraltarian press, at the morphosyntactic level, is almost non-existent, limited to occasional cases of code-switching between the two varieties. We

hypothesize that a possible explanation for this strong exonormative allegiance to British English, at least in press news reports, can be found in a strong editorial pressure to reflect the prestigious parent-variety.

Keywords: Gibraltar English, morphosyntactic variation, language contact, International Corpus of English

1. Introduction

Gibraltar lies at the southern tip of the Iberian Peninsula, with a surface of some 6 km[2] and a population of 32,000 people (Census 2012). It is a British Overseas Territory, and has been subject to the sovereignty of the United Kingdom since the beginning of the 18th century, when Spain, following the War of the Spanish Succession, ceded the territory to Great Britain as part of the Treaty of Utrecht. From that moment, English became the official language, although, due to the Spanish roots of part of the population, as well as for obvious geographical reasons, Spanish remains one of the languages of Gibraltar, spoken with its own distinct Andalusian accent.

Gibraltar has historically been considered "a melting pot of peoples from different cultural backgrounds and with different languages who have settled there throughout the centuries to pursue various military, trade and commercial interests" (Moyer, 1998: 216),[1] which developed a pidgin language for communication. In the 20th century the local community has become ethnically more homogeneous, competent in both English and Spanish, and also speaking *Yanito*, the local vernacular language of Gibraltar, defined as "an Andalusian Spanish-dominant form of oral expression which integrates mainly English lexical and syntactic elements as well as some local vocabulary" (Levey, 2008: 3; see also Moyer, 1998: 216). These days an increase in the use of English at home (inter-parental situation) among youngest Gibraltarians is observed (Kellerman, 2001: 91-93; Levey, 2008: 58, 95-98; Weston, 2013).

The linguistic situation in Gibraltar has been described from different perspectives, such as the use of code-switching (e.g. Kramer, 1986; Moyer, 1993, 1998), its specific phonological features (e.g. Cal Varela, 1996, 2001), the status of Gibraltarian English as regards other Englishes according to Schneider's (2007) model (e.g. Weston, 2011), as well as in broader and more holistic ways (e.g. Levey, 2008). Nevertheless, all these studies are based on interviews and the observations of researchers, which are always potentially subject to a degree of subjectivity or unintended bias. In order to avoid this, a corpus-based methodology is the preferred option for studies on language variation (see Rissanen (2014) for a good review). The existence of a Gibraltarian corpus, then, would guarantee objectivity in the analysis of this linguistic variety. The team to which the authors belong is currently compiling the Gibraltar component of the *International Corpus of English* (ICE), an international project which seeks to provide corpora which allow for comparisons between different varieties of English.[2] Although the corpus is still in progress, we present here a preliminary study using one of the completed sections, which includes press news reports; in this context we provide a comparison of

the British and the Gibraltarian press, with the aim of detecting evidence of contact with Spanish. The paper is structured as follows. Section 2 accounts for the language contact situation in Gibraltar. Section 3 describes the ICE project and the current status of ICE-Gibraltar (ICE-GBR). Section 4 describes the data and methodology. Section 5 presents the corpus-based analysis of press news reports in Britain and in Gibraltar. Finally, Section 6 summarizes the results and reaches some conclusions.

2. Language contact in Gibraltar

The linguistic situation in Gibraltar has been the object of research since the 1980s, and has generated an extensive literature (cf. especially Ballantine, 1983, 2000; Cal Varela, 1996, 2001; Cavilla, 1990; Kellerman, 2001; Kramer, 1986; Levey, 2008, 2015; Lipski, 1986; Moyer, 1993, 1998; Weston, 2011, 2013; Suárez-Gómez, 2012). Most studies here have looked at the historical coexistence of English and Spanish and the resulting code-switching variety known as *Yanito*.

    As noted in the previous section, the community of Gibraltar has historically been defined as a cultural, military, commercial and linguistic "melting pot" (Moyer, 1998: 216). In fact, official documents report the following populations in the 18th century: "British 434; Genoese 597; Jews 575; Spaniards 185, and Portuguese 25", showing the coexistence of "peoples from different cultural backgrounds"[3] and with different languages. Hence, for communicative purposes it was necessary to develop some sort of pidgin language, in this case a mixture of Italian, Spanish, English, Arabic and Hebrew (Moyer, 1993: 85), which until the end of last century was frequently used by local speakers. By that time, Gibraltar had become a more homogeneous ethnic group, as reflected in the 2012 official census, which reports that 79% of the population is Gibraltarian (Gibraltarians 25,444; UK and Other British 4,249; Moroccan 522 and Other Nationalities, 1979) (<https://www.gibraltar.gov.gi>, last access 25 September 2017).

    As expected, the historical distribution of the population in Gibraltar correlates with the situation of language use. From a linguistic point of view, Gibraltarians in their 60s and older tend to be relatively competent in both English and Spanish, but currently speak *Yanito*. This is a linguistic variety which identifies Gibraltarians[4] and which has emerged as a result of "various patterns of code-switching with a proportionally small lexical substratum from Italian, Hebrew and Arabic and a local vernacular" (Moyer, 1998: 216). However, younger generations are more likely to exhibit linguistic code-alternation between *Yanito* and English, most of them with a passive familiarity of Spanish, but communicating mainly in English.[5]

    English has been the target language given that, in terms of socioeconomic dominance, it is more prestigious than Spanish. After so many years of contact with different languages, and especially intense between Spanish and English, a nativized variety is acknowledged, labelled Gibraltar(ian) English in the literature (Kellerman, 2001; Levey, 2008, 2015; Weston, 2011, 2015; Seoane et al., 2016). English is becoming not only the official language, but also the first language of most

Gibraltarians, and there is explicit recognition of an emerging variety of English which is in the process of becoming nativized. The English of Gibraltar is turning into a 'new New English' (Kellerman, 2001), this also known as Gibraltarian English. Although still absent from New Englishes studies, it is a variety developed to represent a local identity (Le Page and Tabouret-Keller, 1985) and is initially exonormative and with covert prestige, although now gradually turning into a more prestige variety used in formal situations. As a consequence of nativization, we assume that the resulting variety is affected by, and undergoes changes in comparison with, other languages that enter into the contact process (Schneider, 2007). A key element here is the degree of such interaction and the extent to which it has had an effect on the new variety of English.

As a result of contact between coexisting languages, "complex patterns of contact linguistics, including lexical transfer, code switching and code mixing, and discoursal and syntactic change and accommodation" (Bolton, 2006: 261) are very likely to occur (see also Thomason and Kaufmann, 1988). Therefore, following Thomason (2001: 63) we assume that in situations of language contact all language levels can be affected and "anything" can be adopted from the languages in contact, from vocabulary, to phonological or structural features. However, it is generally agreed that language contact is especially obvious in the case of lexis and phonology, but less so in syntax and pragmatics, which entails a very intense contact, as shown on Thomason and Kaufmann's borrowing scale (1988: 74-76). With Gibraltar English, the contact between the two languages has been so intense that the resulting variety has its own peculiar traits not only at the level of lexis, but also in terms of pronunciation and prosody, as observed by Levey (2015: 61), "it has a syllable-timed rhythm rather than a stress-timed one and weak forms are rarely used" (see also Kellerman, 2001: 307-308). Lexically speaking, the resulting variety reflects interference between Spanish and English,[6] and the semantic fields most affected are cultural terms relating to food (e.g., *greivi* 'gravy', *saltipina* 'salted peanuts'), specialized vocabulary related to docks and constructions (e.g., *cren* 'crane', *doquia* 'dockyard'), vocabulary associated with the classroom (English as the language of education; e.g., *cho* 'chalk') and the use of false friends (e.g., *aplicacion* 'job application').

3. The *International Corpus of English* project and ICE-Gibraltar

3.1. The ICE project

The *International Corpus of English* (ICE) project was born three decades ago, after a call in 1988 by Sidney Greenbaum (University College London) for English scholars to expand the scope of computerized corpora beyond Brown and the Lancaster-Oslo/Bergen (LOB) corpus, the American and British corpora from the 1960s that had set the standard for corpus-based linguistic studies. In an often quoted note published in *World Englishes*, he stated:

We should now be thinking of extending the scope for computerized comparative studies in three ways: (1) to sample standard varieties from other countries where English is the first language, for example Canada and Australia; (2) to sample national varieties from countries where English is an official additional language, for example India and Nigeria; and (3) to include spoken and manuscript English as well as printed English. (Greenbaum, 1988: 315)

Since then several ICE components have been launched: ICE-GB (1988), ICE-EA (1999), ICE-IND (2002), ICE-SIN (2002), ICE-PHI (2004) and ICE-HK (2006), among others. Each of them has the same design and consists of 1 million words (60%, spoken material, 40%, written material), distributed in text-types as follows:

SPOKEN (300)	Dialogues (180)	Private (100)
		Public (80)
	Monologues (120)	Unscripted (70)
		Scripted (50)
WRITTEN (200)	Non-printed (50)	Student writing (20)
		Letters (30)
	Printed (150)	Academic writing (40)
		Popular writing (40)
		Reportage (20)
		Instructional writing (20)
		Persuasive writing (10)
		Creative writing (20)

Table 1: ICE internal structure (numbers in brackets indicate number of 2,000-word texts)

Among the most remarkable strengths of the ICE project are that, firstly, 60% of each corpus comprises spoken material (much of it being spontaneous), as seen in Table 1, which allows for the study of less stylized varieties where more international variation is to be expected. Secondly, the project has allowed English corpus linguistics to become a global endeavour, instead of being restricted to the mainstream varieties, British and American English, this expansion being at the heart of Greenbaum's proposals (see also Nelson, 2006: 736-740). It contains parallel corpora of varieties of English as a Native Language (e.g., Canada, New Zealand, Ireland), English as an institutionalized Second Language (i.e., official or widely used language for intranational communication, such as education, media, administration, e.g., India, Singapore English), English as a Second Dialect (ESD, e.g., Jamaican English) and varieties of English spoken in places where its exact status is debatable (e.g., Maltese English). All the corpora are freely available for download following communication with the coordinator.[7]

The size of ICE, one million words per variety, may be considered small by comparison with current standards (see BNC, with 100 million words, COCA, with 520 million words, etc.). Yet it is sufficient for the study of grammatical variation, as shown by Seoane and Suárez-Gómez's (2013) study of the perfect meaning in Asian varieties, by Suárez-Gómez (2014), which analyses relative clauses, and by Loureiro-Porto (2016), which focuses on modal verbs. However, if the aim is to study low frequency items, such as lexical elements, specific collocations or contractions and even syntactic constructions (e.g., García-Castro 2017), ICE's size is clearly one of its main disadvantages. For such studies, the use of a larger corpus, such as the *Global Web-Based Corpus of English* (GloWbE, Davies, 2013), is required, although important differences regarding the degree of representativeness have been found between ICE and GloWbE (Loureiro-Porto, 2017).

Other caveats regarding ICE include the time lag between corpora (e.g., between ICE-GB and ICE-Nigeria there is a 20-year difference) and even within corpora (e.g., ICE-Fiji contains a 20-year time gap between some of its text types). Such gaps may involve the risk of interpreting diachronic differences as differences between the varieties themselves (cf. Hundt, 2015), especially in the study of rapidly-changing features, such as quotatives (e.g. *be like*) or intensifiers (*super, über*, etc.). In addition, the inter-corpora comparability of registers may be threatened by cultural differences, as explained by Hundt (2015: 384-385; see also Schaub, 2016), who cites the example of a student essay in ICE-PHI consisting of scattered thoughts and incomplete sentences. Spoken private registers also illustrate the effects of cultural differences in the compilation of the corpora, because very often the compiler has no choice but to interview speakers, with the spontaneous and natural character of the text-type then becoming questionable (cf. Hundt, 2015). Additionally, it is not always easy to decide exactly who qualifies as a speaker of the variety being compiled (Mukherjee and Schilk, 2012: 191).

In spite of all these caveats, ICE continues to be the only project that provides representative corpora of varieties of English, although excluding online texts (and important differences have been found between those two ways of compiling corpora, see Loureiro-Porto, 2017). In addition to the 12 corpora released thus far (India, New Zealand, Singapore, Australia, Canada, Great Britain, East Africa, Hong Kong, Ireland, Jamaica, Nigeria and Philippines), the written components of three varieties are also available (Ghana, Sri-Lanka and USA), and 12 international teams are working on the compilation of new members of the ICE family: Bahamas, Fiji, Gibraltar, Malaysia, Malta, Namibia, Pakistan, Puerto Rico, Scotland, South Africa, Trinidad & Tobago and Uganda. With this in mind, ICE-Gibraltar (ICE-GBR) will be described in the following section.

3.2. ICE-Gibraltar: current status

In 2014 the research unit *Variation in English Worldwide* (<http://view0.webs.uvigo.es>, Principal Investigator: Elena Seoane) was commissioned to compile the Gibraltar component of the ICE project. The current team

comprises three full-time members, three affiliated members, and six PhD students, one of whom also works as a research assistant. During the period 2014-2017 it received funding from the Spanish Ministry of Economy and Competitiveness (grant FFI2014-53930-P) to begin compiling the corpus. In accordance with other ICE corpora, we decided that the written part would be compiled first, and, for practical reasons, the first registers were Reportage (Press News Reports, W2C in Table 1 above), Persuasive Writing (Press editorials, W2E in Table 1) and Creative Writing (Novels and Short Stories, W2F in Table 1).

Following the ICE structure, 20 texts of 2,000 words each were necessary for Press News Reports, and 10 texts of 2,000 words for Press Editorials. Both of these registers have been extracted from online Gibraltarian newspapers such as *Gibraltar Chronicle* and *Panorama*. Nevertheless, the conversion of these electronic texts into part of the corpus requires much more work. To begin with, it is necessary to confirm that the journalists and writers in question are reliable Gibraltarian speakers and that they meet the criteria as laid down by ICE:

> The authors and speakers of the texts are aged 18 or above, were educated through the medium of English, and were either born in the country in whose corpus they are included, or moved there at an early age and received their education through the medium of English in the country concerned. [The ICE Project 2009]

Since Gibraltar is a very small community (see section 1), it is very common for its young citizens to study abroad, particularly in the UK (as noted by Jennifer Ballantine of the Postgraduate Institute of Gibraltar and Mediterranean Studies and also director of the Garrison Library, Seoane, 2017). Such a context would favour the accommodation of these Gibraltarian speakers into British English, as described in the literature on language contact (e.g. Bolton, 2006: 261). Obviously, we need to avoid including texts which do not portray the *real* Gibraltarian variety; yet, for the reasons explained by Jennifer Ballantine, it is highly likely that anyone with a university degree has spent some time in the UK. Therefore, in an attempt to find a balance between rigour and pragmatism, we decided that the conditions for anyone to be considered a *true* Gibraltar English speaker should be: (i) they cannot have lived in the UK longer than 10 years, and (ii) they must have been back in Gibraltar for more than 8 consecutive years. These periods would allow citizens to have obtained a university degree in the UK, but at the same time would have been expected to be fully readjusted to Gibraltarian speech and language use. These characteristics apply not only to the writers of the texts, but also to any speakers who may have produced utterances included as a quotation in any item of reportage included.

Having identified those news reports and editorials that satisfied the above criteria, the second step in the compilation of sections W2C and W2E involved a random selection of texts which reached 2,000 words, as specified by ICE. Because many news reports are shorter than this, in some cases several texts had to be used as sub-parts of the same file.

The third step involved the xml codification of the texts according to ICE norms. A comprehensive list of codes is provided for compilers, and includes issues such as metadata (e.g. author's age and sex), typographic markup (e.g. bold or italics), content markup (e.g. headings, paragraphs), extra-corpus material (e.g. pieces of text produced by interviewers), and normalization of the text (e.g. annotation of misspellings). This has been done manually with the annotation software Oxygen, as shown in Figure 1. In order to guarantee the correct and coherent annotation of the texts, two individual revisions were made of each transcription. This involved decisions such as the following:

- Distinguishing foreign text from an indigenous use of Spanish forms, as seen in examples (1) and (2) respectively:

(1) like El Mundo (<foreign>Radicales de Gibraltar invitan a hundir barcos españoles</foreign>) <*Gibraltar Chronicle* 2014-04-25>
(2) it appears that the plan was then that the two PSOE <indig>alcaldes</indig> from La Linea and San Roque who were due to make an appearance for a debate scheduled for today <*Gibraltar Chronicle* 2014-04-25>

- Identifying quotes introduced by reporting verbs, even when quotation marks are not used:

(3) <quote>Food for those who could afford it - officers and wealthy merchants </quote>, Bresciano **thought**. <*Fall of a Sparrow* 2010-11-18>
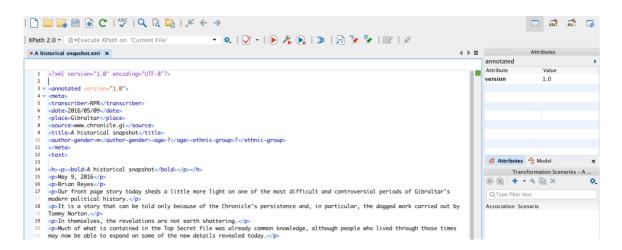

Figure 1: Oxygen xml editor: A snapshot

Regarding Novels and Short Stories, 20 further texts of 2,000 words each were needed. These were transcribed and codified applying the same criteria described above. Nonetheless, the idiosyncratic nature of each register involved a constant decision-making process which will probably only end once the final corpus is released.

As an example, when transcribing one of the novels selected for inclusion in the corpus (its writer having met all the criteria to be considered a *true* Gibraltarian), we realized that much of the text corresponded to the speech of a Scottish character, speaking with a Scottish variety. The difficulty in discriminating between vernacular forms representative of Scottish or Gibraltarian English led us to discard this novel as a source.

To sum up, ICE-GBR is still in progress and the challenges ahead are considerable. However, since the section on Reportage (Press News Reports, W2C) is now fully compiled and annotated, in what follows we present a corpus-based approach to the study of language contact by comparing this section with the corresponding part of the ICE-GB component.

4. Corpus and methodology

The data under analysis here are drawn from ICE corpora (*International Corpus of English*), and are intended to represent an educated standard variety of English spoken in the regions concerned. The ICE corpora include the speech of adult (over 18 years) males and females. As noted in the previous section, in the case of ICE-Gibraltar (ICE-GBR), the speakers must also have been educated in English and have spent the last eight years in Gibraltar. ICE-GBR is currently the only source of data available on Gibraltar English, and a distinct advantage of its use here is the availability of other corpora following the same compilation guidelines, making comparisons with other varieties included in the ICE project possible. In this case the comparison is with British English (ICE-GB), and we acknowledge that there is a 30 year gap between the dates of compilation of the two corpora. From these corpora, the whole section "press news reports" (W2C) was selected for analysis, which amounts to a total of 40,000 words per variety, that is, a total of 80,000 words.

The retrieval of examples proceeded in two stages. In order to ensure a maximally exhaustive search, we first retrieved examples automatically using AntConc. The examples thus selected were analysed manually in order to select relevant examples of each of the variables under analysis (see Section 5). Since different variables were analysed, the individual searches are detailed in the analysis of each variable in Section 5.

5. Comparing press news reports in GB and GBR: Evidence of language contact?

This section explores the W2C parts of ICE-GB and ICE-GBR with the aim of discovering whether the contact with Spanish plays any role in the language of the Gibraltarian press at the morphosyntactic and lexical levels. Since the style of English and Spanish journalistic registers differ in various respects, the variables selected in this pilot study include those in which a larger difference is expected to be found. Because of space constraints, they are reduced to: (1) frequency of the passive voice, (2) type of

relativizer in subject position, (3) frequency of titles and pseudo-titles, and (4) insertion of Spanish forms (code-switching).

5.1. Frequency of the passive voice

The use of the passive voice is very common in English in constructions where the agent is new information, as in (4):

> (4) *The mayor's term of office expires next month.* **She will be succeeded** *by George Hendricks*. [from Huddleston and Pullum 2002: 1444]

The passive voice as an information-rearranging strategy has been in use in English since the Middle English period (Seoane, 2006a: 361), as a result of the combination of grammatical factors, such as the levelling of inflections, and pragmatic ones, including the tendency for topics to appear in initial position (Seoane, 2006a: 368), and it is particularly common in registers where objectivity is the main aim, as is often the case with scientific texts (see Seoane, 2006b). This is probably one of the reasons why journalists use the passive voice so frequently, an example of which would be (4) above. However, this kind of construction is not the preferred option in Spanish, which, being an inflectional language with high word order flexibility, would more typically opt for a construction such as (5):[8]

> (5) *El mandato de alcaldesa termina el próximo mes.* **La sucederá** *George Hendricks*.
> her succeed.3p.sg.fut

In (5) the verb *sucederá* 'will succeed' is in the active voice, the subject being the following NP, *George Hendricks*. As we see, information packaging in this Spanish sentence only differs from (4) in terms of the morphosyntactic option taken (active vs. passive voice), because in both sentences the first element is a pronoun referring to the mayor, which is discourse-given.

The active voice, then, is the preferred option in spoken Spanish, and although its use in journalistic style is higher than in any other register (Rodríguez Pastor and Castro Verdala, 2013: 143), we begin with the hypothesis that it will be far more common in English press reportage than in Spanish, not least because its abuse in Spanish journalistic texts has been heavily criticized in style guides over recent decades (Rodríguez, 2001).

Thus, we have explored the 40,000 words of the W2C sections of ICE-GB and ICE-GBR by searching all possible forms of the verb BE, including both full and contracted forms. Only clear passives were included in the database; cases including an *-ed* form, which could be interpreted as an adjective, for example because it accepts intensification (see example (6)) or because it included a nearly lexicalized expression (7), were not considered actual passive constructions:

(6) <p>Alexia said a lot of young people were **scared** of marriage and taking their vows.</p><*Gibraltar Chronicle* 2014-02-15>

(7) Labour 's campaign co-ordinator , Mr Jack Cunningham , said : “ The Ribble Valley by-election result shows clearly that the Government is at the end of the road and that the poll tax has to go , ” while Shadow Environment Secretary Bryan Gould said : “ The British people **are fed up** with Tory excuses and with the continuing Cabinet disarray about the future of this hated tax . ” <ICE-GB:W2C-018 #26:1>

After having excluded such unclear cases, the final tally was that shown in Table 2:

	ICE-GB	ICE-GBR
BE-PASSIVES	521	353

Table 2: Number of passive structures with *be* in the W2C sections of ICE-GB and ICE-GBR

The use of passive constructions in the press texts of ICE-GBR is much less frequent than in ICE-GB, and according to the *z-test* score for the two population proportions, the difference is very significant.[9] This difference, however, must not immediately be taken as a result of contact with Spanish, and we should remember that ICE-GB was released in 1988, some 26 years before ICE-GBR was even conceived. Hence, we could be witnessing a case of stylistic change. In fact, a significant decrease in the frequency of the passive in another text-type where it formerly featured extensively, namely scientific registers, has been shown to have taken place over the course of the 20th century (see, for example, Seoane, 2006b). Westin (2002: 165) finds a similar pattern in the British press, where he notes not only a reduction in the number of passives, but also in the number of subordinate clauses, as well as an increase in the use of informal language, such as interactional features (questions and imperatives) and shorter sentences. This change in the style used in some text-types has often been referred to as informalization (Leech et al., 2009: 239; Farrelly and Seoane, 2012: 395).

Informalization is defined by Farrelly and Seoane (2012: 395) as "the process whereby the distance between addresser and addressee is shortened, probably in order to make the text more engaging, accessible, and reader-friendly". Such a process of change, which affects traditional formal registers, involves a high number of grammatical features, the exploration of which falls outside the scope of this paper, and thus will not be analysed in detail here. Nevertheless, Biber's (1988) proposal for multidimensional analysis of registers allows for a brief exploration. Thus, if we focus on his Dimension 1 "Informational vs. Involved Production" (1988: 102), we see that one of the negative features in this dimension, which implies a higher degree of formality, is the presence of passive constructions; another is the type/token ratio, that is, more involved (meaning 'more oral') texts will exhibit a lower type/token ratio, while more informational (i.e. 'less oral') texts will exhibit a higher type/token ratio. The differences between ICE-GB and ICE-GBR regarding this feature are shown in Table 3.

ICE-GBR	# Word Types: 5224 #Word Tokens: 39357	RATIO: 7.53
ICE-GB	#Word Types: 7569 #Word Tokens: 53543	RATIO: 7.07

Table 3: Type/token ratio in the W2C sections of ICE-GB and ICE-GBR

The higher type/token ratio of ICE-GBR is also found to be significant, which does not support the hypothesis that the differences between both varieties in terms of the frequency of the passive voice are due to a higher informalization of the journalistic style in the 2010's than in the 1980's.

Related to informalization is colloquialization (Farrelly and Seoane, 2012), a process characterized by "a tendency for written norms to become more informal and move closer to speech" (Leech et al., 2009: 20). Within the journalistic register colloquialization manifests itself, among other ways, in an increasing use of quotations, since, as Leech et al. (2009: 20) put it, "newspaper reports now come with more direct passages of direct quotation – whether real or fictitious – than they used to". In fact, a considerable increase in the frequency of quotations has been found in the Brown family of corpora between the 1960's and the 1990's (see Leech et al., 2009: 128). With the aim of determining whether the difference in the use of the passive voice shown in Table 2 can be interpreted as the result of the colloquialization of the English-speaking press from the 1980's to the 2010's, we have searched the quotations used in the W2C sections of ICE-GB and ICE-GBR.

Such a search is simple to conduct, in that quotations are annotated in ICE corpora, the codes being *“* (open quote) and *”* (end of quote) in ICE-GB, and <quote> and </quote> in ICE-GBR. Nevertheless, and interestingly enough, we have found that in ICE-GB more quotes are opened than actually closed, as seen in (8), and on occasions some quotes are closed, but not opened (9):

(8) <ICE-GB:W2C-001 #63:3>
    With the Dumfries inquiry mov<l>ing into its second month , Mr Kreindler **said** that there was nothing his group could do .
    <ICE-GB:W2C-001 #64:3>
    “We will have to wait and see if there is any substance to any of this .
    <ICE-GB:W2C-001 #65:3>
    “Meanwhile our work at the in<l>quiry will continue .
    <ICE-GB:W2C-001 #66:3>
    What has come out so far has been very im<l>portant to us .
    <ICE-GB:W2C-001 #67:3>
    We have been given new information into areas of cau<l>sation , and we will be able to du<l>plicate this evidence in our claim . **”**

(9) <ICE-GB:W2C-001 #93:5>
    A colleague of Francisco San<l>tos , news editor of the daily El Tiempo , seized on 19 September , **commented** : : This looks very bad , very dangerous for the hostages . **”**

This codification problem seen in ICE-GB is unlikely to be found in the new generation of ICE corpora, annotated in xml using the Oxygen software (as shown in Figure 1 above). Indeed, in our search for quotations in ICE-GBR there is no mismatch between the number of opening and closing quotes. That said, the number of quotations found in each corpora is shown in Table 4:

Corpus	Opening quote	Closing quote
ICE-GB	265	247
ICE-GBR	224	224

Table 4: Quotations in ICE-GB and ICE-GBR

In order to avoid any possible bias, and considering that the number of opening quotes is larger than that of closing quotes, we have taken the smaller figures of ICE-GB for the comparison between the two corpora, because in this way we make sure that we are not counting the same quotation twice, as would be the case in (8). Thus, we observe that quotations are less frequent in ICE-GBR than in ICE-GB (although the difference is not found to be significant at p<0.0.5, according to the *z-test* score), and thus we cannot conclude that a process of colloquialization is at work.

In summary, the frequency of passives is much lower in ICE-GBR than in ICE-GB and, having excluded the potential bias of the diachronic gap between both corpora by showing that this does not correlate with informalization or colloquialization features, the hypothesis that the Gibraltarian press might be influenced by the Spanish style is still valid. Let us turn now to other features.

5.2. Relative markers

Research on relative clauses and relativization processes has generated a substantial body of work over the years, especially with regard to the distribution of relative markers within the relative clause and the factors that condition this, mainly in British English (BrE), American English (AmE) (Tottie, 1997; Guy and Bayley, 1995; Hinrichs et al., 2015) and some regional varieties (Tottie and Rey, 1997; Tottie and Harvie, 2000; Tagliamonte, 2002; Herrmann, 2003). More recently, research on relative clauses has also extended to New Englishes (Gut and Coronel, 2012; Huber, 2012; Suárez-Gómez, 2014, 2015, 2017).

Adnominal relative clauses in standard PDE are introduced by different types of relative words or relativizers, the most frequent of which are the invariable relativizer *that* (10), the pronominal relativizers represented by *wh*-words, *who* (11), *whom* (12), *whose* (13), and *which* (14), and the zero relativizer (15), traditionally distributed according to the animacy of the antecedent (human (11) vs. non-human (10)), the syntactic function of the relativizer (subject (10), object (16), complement of a preposition (12) or possessive (13)) and the type of relative clause (restrictive, RRC

(10) vs. non-restrictive, NRRC (13)) (Biber et al., 1999: 608-631; Huddleston and Pullum, 2002: 1037-1057).

> (10) Many have lost one parent, some both, in the 15 years of civil war **that** have brought Angola, potentially one of the richest countries in Africa, to its knees. <ICE-GB:W2C-002 #40:2>
> (11) Other teachers believe the proposals will harm the pros<l>pects of those **who** want to specialise in traditional sci<l>ence subjects. <ICE-GB:W2C-002 #25:1>
> (12) Last season, he was on loan to Swansea, for **whom** he played in the European Cup. <ICE-GB:W2C-014 #28:1>
> (13) The 43-year-old man, **whose** name was withheld, will not face trial. <ICE-GB:W2C-019 #85:6>
> (14) The combination created a climate in **which** safety was not put first. <ICE-GB:W2C-007 #118:3>
> (15) Is that an irritation when you have a vague feeling **Ø** you've lent a book to somebody and you can't quite figure it out <ICE-GB:S1A-013 #92:1:A>
> (16) Publishing shares and Reuters featured on consideration of the savings **that** shareholders in the BSB satellite merger would expect to achieve. <ICE-GB:W2C-005 #61:3>

Research on the distribution of relativizers in emergent varieties of English or World Englishes, especially in Asian Englishes, shows that these varieties converge with Standard varieties of English in that *who* tends to be used as a subject with human antecedents (Levey, 2006), whereas *that* supersedes *which* in the same syntactic function with non-human antecedents, as a consequence of the general decrease of *which* at the expense of *that* (Leech et al., 2009: 227, 229). In varieties such as Indian English, however, *wh-* relativizers are clearly favoured, as is the case of *which* with non-human antecedents, irrespective of syntactic function, and this has been attributed to substratal influence, i.e. the non-reduction relativization strategy of Hindi, a language that uses pronominal relativizers to introduce the relative clause (Suárez-Gómez, 2014).

For this preliminary study on the distribution of relativizers in Gibraltar English compared to the same markers in ICE-GB, only adnominal subject relative markers in restrictive relative clauses were taken into account, that is, relative clauses introduced by a relative marker functioning as subject of the relative clause. For the search, cases of *that, who* and *which* were first automatically retrieved using AntConc. These searchers rendered 1,545 examples which were subsequently analysed manually in order to select the relevant examples. In this process, cases of *that* as a complementizer (17) or demonstrative (18) or in combinations such as *so that* (19), and cases of relativizers which neither functioned as subjects nor introduced restrictive relative clauses (20), were excluded. This yielded a sample of 355 examples of restrictive relative clauses introduced by *that, who* or *which*, as Table 5 below illustrates.

> (17) The Gibraltar Government acknowledged the concerns expressed by ACI Europe and echoed the British Government's view **that** Gibraltar must be included in any EU aviation measures. <*Gibraltar Chronicle* 2016-10-10>

(18) It becomes crystal clear that Gibraltar is not part of **that** negotiation. <*Gibraltar Chronicle* 2016-10-16>

(19) This memorial helps us to remember all those who were killed at work but at the same time it helps us to reflect on working procedures **so that** similar tragedies are not repeated <*Gibraltar Panorama* 2014-04-29>

(20) That is a position at odds with the British Government, **which** insists Gibraltar will be fully involved in preparations for Brexit <*Gibraltar Chronicle* 2016-11-15>

In Standard BrE, this type of relative clauses can be introduced by invariable *that* or by a *wh-* word, either *who*, if it refers back to a human antecedent, or *which*, if the antecedent is non-human. In Spanish, the invariable relativizer *que* "es el relativo de uso más general en español: puede aparecer tanto en las cláusulas especificativas como en explicativas y es la única que está capacitada para desempeñar en ambas cualquier función sintáctica (precedida, eventualmente, de la correspondiente preposición y del artículo determinado). …. Se trata, así pues, del nexo relativo por defecto en español".[10] (Brucart, 1999: 490); unlike English, the pronominal relativizer *quien*, only possible with human antecedents, is only possible in subject function in NRRCs (Brucart, 1999: 502).

Therefore, if language contact was operating in Gibraltar English in the choice of the relative marker functioning as subject in restrictive relative clauses, *that* would be expected to be used more frequently than in British English, which shows variation between pronominal *wh-* words and *that*.

	Animate		**Inanimate**	
	Who	*That*	*Which*	*That*
ICE-GBR	67	2	59 (41%)	85 (59%)
ICE-GB	56	5[11]	39 (46.4%)	45 (53.6%)

Table 5: Subject relativizers in RRCs in press news reports in ICE-GBR and ICE-GB

The results from Gibraltar English show a different distribution between cases of animate and inanimate antecedents, although this difference is not significant at the 0.05 level. While *who* is the default marker with animate antecedents, irrespective of the variety, *that* is more frequently used than *which* with inanimate antecedents, especially in Gibraltar English. Although we can refer to the Spanish influence to account for this slightly higher use of *that*, it is difficult to give definite conclusions because, as seen above, a peculiarity of World Englishes is that *that* supersedes *which* as subject with inanimate antecedents, also associated with a general decrease of *which* at the expense of *that* (Leech et al., 2009: 227, 229). Apart from contact, the date of compilation of the corpora analyzed is different (see section 3.1), which may also influence the results.

Summarizing, language contact with Spanish would favour the frequent use of the invariable relativizer *that*, but this is also in line with the varied literature on the distribution of relative words which agree that the *wh*-words are contracting significantly in frequency and indeed may disappear in the near future, as Schneider (1992: 446-448) forecast.

5.3. Use of titles and pseudo-titles

The use of titles in the Anglophone press is much more common than in the Spanish-speaking press, as seen in the following two examples:

(21) Reports at the weekend claimed the Prime Minister, **Mr John Major**, would raise the issue with the United States president, **Mr George Bush**, at their summit meeting in Bermuda on Saturday , in the hope of softening US opposition to forced repatriation. <ICE-GB:W2C-019 #76:5>

(22) **John Major** se enfrenta a los peores días de su vida. […] Su buena relación con el presidente de Estados Unidos, **George Bush** (ambos eran los herederos grises de los carismáticos Reagan y Thatcher), su actitud sin fisuras ante el intento de golpe de estado en la antigua Unión Soviética, y su viaje a Pekín para pedir respeto a los derechos humanos, forjaron de él una interesante imagen de estadista. <*El País* 1992-04-10>[12]

Sentences (21) and (22) refer to two prominent politicians of the 1990's whose names are preceded by the title *Mr* in the English excerpt (taken from ICE-GB), while in the Spanish one the same public figures are referred to using their given names followed by their last names.

In addition, other nouns are also commonly used in English without a determiner and preceding a proper name, as in (23) with *President Macron*. Other cases include *lawyer, golfer* and *teenager* (Hundt and Kabatek, 2015) and are usually referred to as pseudo-titles (or bare noun phrases, see Kabatek and Wall, 2013):

(23) **President Macron** has pledged to clean up French politics and public life after a series of scandals that have damaged voter confidence in their elected representatives <*The Guardian* 2017-06-20>[13]

This use cannot be said to be unattested in Romance languages, since it is recorded in Brazilian Portuguese and American-Spanish varieties, as shown in (24), and has been analysed in detail in Sáez Rivera (2013):

(24) Eminente **juez** supremo Scarman (Argentina, <http://veanegociador.com.ar/Inseg.htm>, as quoted by Hundt and Kabatek 2015)

Despite being an attested use in the press of two Romance languages, the frequency of this construction is much lower than in English (Hundt and Kabatek, 2015) and, therefore, the hypothesis here is that if the Gibraltarian press is somewhat influenced by its contact with Spanish, titles and pseudo-titles will be less frequent in ICE-GBR than in ICE-GB.

In order to test this hypothesis, we explored the W2C sections of both corpora, first by identifying all proper names used in the text, and then by dividing them into names preceded by a title or pseudo-title and bare proper names. The results are summarized in Table 6:

Corpus	Preceded by titles	Preceded by pseudo-titles	Proper names without any title	% of (pseudo-) titles
ICE-GBR	179 (86.5%)	28 (13.5%)	441	(207) 31.9%
ICE-GB	223 (74.6%)	76 (25.4%)	843	(299) 26.2%

Table 6: Frequency of titles and pseudo-titles in ICE-GB and ICE-GBR

As we can see, the sum of titles and pseudo-titles represents a higher percentage of NPs involving proper names in ICE-GBR than in ICE-GB (right-most column), since 31.9% of all proper names are preceded by a noun without an article in the Gibraltar data, whereas the proportion in the British press is just 26.2% (this difference is significant according to the *z-test* score). This result contrasts with those in the previous sections, because ICE-GBR appears to be more British than ICE-GB itself. Such a finding is clearly not the result of language contact or of the potential informalization of the English-speaking press (Westin, 2002, as mentioned above). Two possibilities remain. The first concerns a potential diachronic change related to a specific type of title, because pseudo-titles have been observed to experience an increase in frequency in journalistic texts (Jucker, 1992; Meyer, 2002). However, a closer look at Table 6 reveals that pseudo-titles only represent 13.5% of all titles in ICE-GBR (example (25)), a number significantly lower than their percentage in ICE-GB (example (26)).

(25) <p>The agreement was swiftly dismissed by **Ukip leader** Nigel Farage as a <X>"truly pathetic deal"</X>.</p><*Gibraltar Chronicle* 2016-02-20>

(26) **Tory leader Cllr Bob Black<l>man**said :“ Brent Council must demonstrate to the charge-payer that it provides valuable ser<l>vices to the community. <ICE-GB:W2C-009 #54:3>

Therefore, the increasing use of pseudo-titles in global English cannot be responsible for the differences between ICE-GBR and ICE-GB. The Gibraltarian press in the 2010's shows a significant higher use of titles marking social differences, such as *Mr, Dr, Sir,* etc. than the English press in the 1980's, as illustrated by (27) and (28):

(27) <p>Local author and poet Jackie Anderson has organized a meeting for today for all those interested in writing. The first meeting of the Writer's Group or Writer's Circle will be held at the John Mackintosh Hall at 7.30pm today.</p>
<p>The idea came to **Mrs Anderson** after the interest shown in her workshops on World Book Day.</p><*Gibraltar Chronicle* 2016-04-19>

(28) THE Prime Minister was last night facing two crucial decisions as the shock result of the Ribble Valley by-election brought to an end his honeymoon period as **Mrs Thatcher 's** successor after only 100 days in office. <ICE-GB:W2C-018 #4:1>

Summing up, the use of titles and pseudo-titles in ICE-GB and ICE-GBR does not exhibit any evidence of contact with Spanish, nor does it show any global tendency reflecting recent historical changes regarding the use of bare noun phrases. The only

explanation we can propose, then, is that of a conscious use of titles such as those in (27) and (28) as a result of the prestige of British English. It is quite possible that Gibraltarian journalists use titles with this high frequency in an attempt to sound more decidedly British, a kind of stylistic hypercorrection, while concurrently they highlight the differences between themselves and the Spanish discursive tradition of not using titles so frequently. This hypothesis, we are aware, is not uncontroversial, and much more research would be needed to confirm it.

5.4. Code-switching

Having analysed three morphosyntactic features with diverging (inconclusive) results regarding the influence of Spanish as a contact language with Gibraltarian English (passives, relative markers and use of (pseudo-)titles), this section explores the lexis used in the news reports included in ICE-GBR, with the aim of finding a pattern in the use of Spanish words and expressions. Code-switching is one the common results of language contact (see section 2 and also Thomason and Kaufmann, 1988; Thomason, 2001; Bolton, 2006) and is one of the idiosyncratic features of Gibraltarian speech, as reported in Kramer (1986), Moyer (1993) and subsequent studies, because, as Gal (1988: 247) says:

> code-switching practices are symbolic creations concerned with the construction of the self and the other within a broader political, economic and historical context which reflect the way people respond symbolically to relations of domination between groups within the state, and how they understand this historic position and identity within a world capitalist system structure around dependency and unequal development (Gal, 1988: 247).

These "symbolic creations" are expected to be much more common in spoken registers than in written ones, but of all the types of switches identified by Van Hout and Muysken (1995) we have found that *insertions* do appear in our corpus, although at a low frequency. Of the total of 40,000 words in the W2C section of ICE-GBR, six Spanish forms are inserted into the English texts, as shown in Table 7:

Spanish form	Number of occurrences
Sr	5
Snr	3
alcalde(s)	2
mano a mano	1
a mí me pertenece	1
Pueblo	1

Table 7: Code-switching cases in ICE-GBR

Interestingly enough, the two most common Spanish forms are titles, namely abbreviations of Spanish *señor* 'mister', which exclusively precede names of Spanish-speaking people (*Sr Rajoy, Sr Margallo*, both Spanish politicians, and *Sr Morejon-Pazmiño*, an Ecuatorian diplomat).[14] None of these Spanish names is ever preceded by its English counterpart *Mr*, whose use is restricted to non-Spanish people, such as *Mr Cameron, Mr Clinton, Mr Cooper* and also *Mr Guerrero*, who, despite having a Spanish surname, is a Gibraltarian citizen, the CEO of the Gibraltar Tourist Board. The shortening *Snr* is also worth mentioning, since it is not a Standard Spanish abbreviation (as noted by the *Real Academia de la Lengua Española*).[15] Its use stands alongside that of *Sr*, even in the same text (e.g. *Snr Diego Morejon* and *Sr Morejon*), but it seems to be restricted to one author, Fabian Picardo, Gibraltar's chief minister.

The word *alcalde(s)* is used to refer to Spanish mayors, although the use of the English word is also possible when referring to Spanish politicians, as seen in (29) and (30):

> (29) <p>Whatever evolution that moot invitation took, if at all, it appears that the plan was then that the two PSOE <indig>**alcaldes**</indig>from La Linea and San Roque who were due to make an appearance for a debate scheduled for today, Saturday, asked for this to be changed for diary reasons.</p><*Gibraltar Chronicle* 2014-04-25>
>
> (30) <p>Speaking later to Spanish journalists Mr Feetham said he was saddened that he was invited to sit in a debate with the mayors and the La Linea **mayor** had then pulled out <quote>"not to offend Mr Picardo...who was in the Caymans anyway"</quote>.</p><*Gibraltar Chronicle* 2014-04-25>

The reason for this code-switching in (29) is, therefore, not justified by a linguistic gap, but by its expressive function (Poplack, 1980). Expressivity also seems to be the reason behind the use of the sequences *mano a mano* and *a mí me pertenece*, as shown in (31) and (32):

> (31) <p>Locally nothing changed on the running <indig>**mano a mano**</indig> between Fabian and Daniel, the latter who rose to Leader of the Opposition whilst his predecessor was in June to arise as Sir Peter.</p><*Gibraltar Chronicle* 2014-01-23>
>
> (32) <p>Mr Feetham describes his project as <quote>"seeking to break the mould of Gibraltar politics".</quote></p>
> <p><quote>"I want to break from promising people things, of promising everything to everyone. Gibraltar cannot continue with the culture of <punctuation>'</punctuation><indig>**a   mi   me pertenece**</indig><punctuation>'</punctuation>. It will catch up on us,"</quote> he says warning that if Gibraltar were hit by difficulties the issues he is warning about will become critical.</p><*Gibraltar Chronicle* 2014-01-23>

In sentences (31) and (32), two Spanish expressions are inserted in the discourse, although their equivalent English forms would suit the communicative purposes perfectly: *hand in hand* and *it belongs to me*. The reason for code-switching, then, has

to do with a high degree of expressiveness and the construction of a Gibraltarian identity, in the sense described by Gal (1988), above.

Summing up, these cases of code-switching constitute strong evidence for the influence of Spanish on the Gibraltarian press. Although the native code-switching variety *Yanito* exists mainly in the spoken mode, Gibraltarian speakers seem to have a tendency to insert occasionally Spanish words and expressions into written registers, such as press news reports.

6. Summary and conclusions

This paper has presented a preliminary study of the effects of language contact in Gibraltar by exploring the reportage sections of two components of the *International Corpus of English*, namely ICE-GB (Great Britain) and ICE-GBR (Gibraltar). The ICE project was born in the 1980's with the idea of providing comparable one-million-word corpora of different varieties of English, and the first member of the family was ICE-GB, released in 1988. The team to which the authors belong was commissioned in 2014 to compile the ICE-GBR, and currently 100,000 words have been completed. For this preliminary study only 40,000 words were used from each corpus, namely those including press news reports. With the aim of investigating whether the Gibraltarian press variety is influenced to some extent by its contact language, Spanish, four different features were scrutinized, the frequency of the passive voice, the type of relativizer in subject position, the use of (pseudo-)titles, and code-switching. The results are summarized in Table 8.

Feature	Evidence of contact with Spanish	Other possible explanation
Frequency of the passive voice	Possibly	No. Informalization and colloquialization have been ruled out.
Type of relativizer in subject position	Difficult to determine	Global tendency in World Englishes.
Use of (pseudo-)titles	No	Prestige of British English?
Code-switching	Yes	No

Table 8: Summary of results

It was hypothesized that, if Spanish does influence the Gibraltarian press, then the frequency of the passive voice would be lower in ICE-GBR than in ICE-GB. This was indeed seen to be the case, and was statistically significant. Further tests were conducted to rule out a potential diachronic change (owing to the 30-year gap between the compilation of the two corpora), but the exploration of the type/token ratio (in line with Biber, 1988) and the frequency of quotations (following Leech et al., 2009)

suggest that ICE-GBR does not exhibit features related to informalization or colloquialization.

The lower frequency of the passive voice in the Gibraltarian press is, therefore, not the result of either of these two processes, which implies that contact with Spanish may possibly play a role in this morphosyntactic feature.

Regarding the type of relativizer, the hypothesis was that the overwhelming presence of Spanish *que* would favour a similar use of English *that*. This was found to be the case with inanimate antecedents (but not with animate ones), although asserting that this is the result of the Spanish influence is premature, since other varieties of English worldwide have witnessed the same development recently, especially in the low frequency of *which* with inanimate antecedents, and this seems to be the trend in native varieties of English, as reported by Leech et al. (2009: 227, 229).

The third variable studied was the use of titles (e.g. *Mr, Dr, Sir*) and pseudo-titles (e.g. *Tory leader, therapist*, etc.), with the hypothesis that the Spanish influence would render a much lower frequency of these items in ICE-GBR than in ICE-GB. The results, in this case, clearly go against the hypothesis, since the Gibraltarian press includes (pseudo-)titles with a significantly higher frequency than ICE-GB. This constitutes another piece of evidence for the lack of informalization and colloquialization of this text-type in ICE-GBR; it was also suggested that the reason for this unexpected result could be rooted in a voluntary attempt to distance the text from Spanish discourse traditions and, concomitantly, a desire to invoke the higher prestige of British English.

Finally, cases of code-switching were also examined, since the combination of English and Spanish is common in the spoken mode (and the source for the Gibraltarian variety *Yanito*, see Section 2). The press news reports included in ICE-GBR exhibit several cases of insertion of Spanish forms for expressive purposes, such as the title *Sr*, the noun *alcalde*, and constructions such as *mano a mano*. This final variable is, undoubtedly, the most conspicuous piece of evidence for the contact between English and Spanish in Gibraltar.

All in all, we have shown that the compilation of comparable corpora constitutes a necessary tool for the study of variation in general, and for the analysis of language contact phenomena in particular. Although the compilation of ICE-GBR is ongoing, this preliminary study demonstrates that, despite the fact that the English spoken in Gibraltar exhibits a strong exonormative allegiance to British English, the language used by the Gibraltarian press exhibits evidence of the influence of Spanish, as well as some idiosyncratic features which require further analysis. The findings in this study will have to be compared with the exploration of more text-types and, finally, with the whole ICE-GBR corpus, which will provide a comprehensive picture of Gibraltarian English at the beginning of the 21st century.

Notes

1. MG Sanchez, a Gibraltarian writer, reflects this idea in *The Escape Artist*: "Most Gibraltarians have [...] mixed bloodlines. This is because most of us are the sons of immigrants who came to Gibraltar from other parts of the Mediterranean from the mid-1750s onwards. In fact, in the nineteenth-century everybody used to call us 'mongrels' because we were neither Spanish nor British, but a strange composite of the two with a bit of Genoese, Maltese, Irish, Portuguese and Jewish through into the melting pot as well" (2013: 13).

2. See <http://view0.webs.uvigo.es/ice-gibraltar> for further details, last access 25 September 2017.

3. Official census:<https://www.gibraltar.gov.gi/new/gibraltar-census-history>, last access 25 September 2017.

4. The issue of identity underlies most references to the linguistic situation of Gibraltar. Kellerman (2001: 411) defines Gibraltar as a region in search of an identity, a language and a culture. For Moyer (1993: 236), "Gibraltarians avoid identifying themselves with a given group, and at the same time they are affirming a local identity". This is also clearly shown in the novels of MG Sanchez, one of the most prolific local writers. In *The Escape Artist* we find the following excerpts: "'I am Gibraltarian,' he said in a tone which made it clear that he was no longer kidding". (Sanchez, 2013: 5) and "No, I'm not Spanish. I'm from Gibraltar, mate" (Sanchez, 2013: 14). Finally, Jennifer Ballantine in a recent interview states: "I was away from Gibraltar for a period of years, and from when I left in 1994 to when I came back in 2007, in that period of time this question of identity had somehow erupted in a way that I hadn't heard before, certainly not when I left initially. I suspect that the tercentenary commemoration of the taking of Gibraltar in 2004 would have focused everybody's minds towards the fact that it was three hundred years since Gibraltar was taken from Spain, and this must mean something; at that time a number of publications were produced which aimed to look at the question of Gibraltar" (Seoane, 2017: 222).

5. This age-grading pattern (Weston, 2011; Seoane et al., 2016) is also reinforced by MG Sanchez, who agrees with the assumption that the choice of language is directly related to age (personal communication). Nevertheless, according to Jennifer Ballantine (cf Seoane, 2017: 223) we cannot generalize. In her opinion, "I think that social class and educational background has a certain amount of influence in this question," although she agrees that, before WWII, Spanish was very widespread as shown by the existence of Gibraltarian newspapers other than the *Gibraltar Chronicle*, such as *El Calpense* or *El Anunciador*, both written in Spanish produced for a local readership. Nowadays, however, the youngest speakers are mostly monolingual in English ("I think that English is becoming the main language for this generation, and code switching is probably very difficult; only certain terms enter their register but not every term that perhaps an older generation would be able to use".). The closure of the Instituto Cervantes in Gibraltar in 2015 is further evidence for this increasing gap between the Spanish language and the Gibraltarian population.

6. As acknowledged by MG Sanchez, "[t]o understand modern-day Gibraltar you have to understand this sense of hybridity and all the apparent craziness that comes with it – blond-haired people with Spanish surnames, red pillar boxes beside *chiringuitos* serving *gambas a la plancha,* British bobbies speaking Spanish, et cetera, et cetera" (Seoane, 2016: 255).

7. Even as this article is being written, changes are taking place regarding ICE's website and its coordination. The former coordinator, Gerald Nelson (Chinese University of Hong

Kong), has just been succeeded by Marianne Hundt (Univ. of Zürich) and a board is in the process of being constituted, as announced at a recent ICAME conference (Kirk and Nelson 2017). Likewise, ICE's webpage is in the process of being moved from its present hosting site to an institutional URL at the University of Zürich.

8. Spanish has a passive construction with the pronoun *se* (e.g. *Se venden limones* 'Lemons are sold'; see Mendikoetxea (1999: 1631-1722)), which falls out of our study here.

9. The *z-test* score was obtained using the online calculator available at <http://www.socscistatistics.com/tests/ztest/Default2.aspx>, last access 25 September 2017.

10. The invariable relativizer *que* "is the most common relativizer in Spanish: it can appear both in restrictive and non-restrictive relative clauses, and it is the only relativizer that can play any syntactic function (preceded, if needed, by the corresponding preposition and the definite article). … It is, therefore, the default relativizer in Spanish" [our translation].

11. Here examples with antecedents such as *army* or *government* were found, as in *When Beirut and its environs will be declared officially free of mili<l>tias and under the control of a reunited Lebanese army that has, for the first time, the germ of sectarian balance* <ICE-GB:W2C-010 #77:2>. They refer to a group of people, therefore they have an animate antecedent, but they also allow relative clauses introduced by *which*.

12. This sentence was taken from:
<http://elpais.com/diario/1992/04/10/internacional/702856801_850215.html>, last access 25 September 2017.

13. From <https://www.theguardian.com/world/2017/jun/20/french-defence-minister-sylvie-goulard-resigns-after-a-month>, last access 25 September 2017.

14. Spelled *Morejon-Pazmino* in the Gibraltarian press.

15. See <http://buscon.rae.es/dpd/apendices/apendice2.html> for a full list of accepted Spanish abbreviations, last access 25 September 2017.

References

Ballantine, Sergius J. (1983): "A study of the effects of English-medium education on initially monoglot Spanish-speaking Gibraltarian children". MA dissertation, University of Valencia.

Ballantine, Sergius J. (2000): "English and Spanish in Gibraltar: Development and characteristics of two languages in Gibraltar". *Gibraltar Heritage Journal*, 7: 115–124.

Biber, Douglas (1988): *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan (1999): *Longman Grammar of Spoken and Written English*. Harlow: Longman.

BNC = *The British National Corpus*, version 3 (BNC XML Edition) (2007): Distributed by Oxford University Computing Services on behalf of the BNC Consortium. URL: <http://www.natcorp.ox.ac.uk/>[accessed 10 June 2017].

Bolton, Kingsley (2006): "World Englishes Today". In Braj B. Kachru, Yamuna Kachru and Cecil L. Nelson, eds. *The Handbook of World Englishes*. Oxford: Blackwell, 240-269.

Brown = A Standard Corpus of Present-Day Edited American English, for use with Digital Computers. Compiled by W. Nelson Francis and Henry Kučera (Brown University. Providence, Rhode Island).

Brucart, José Ma. (1999): "La estructura del sintagma nominal: las oraciones de relativo". In Ignacio Bosque and Violeta Demonte, coords., *Gramática descriptiva de la lengua española. Vol. I Sintaxis básica de las clases de palabras*. Madrid: Espasa Calpe, 395-522.

Cal Varela, Mario (1996): "Hacia una concepción prototípica de comunidad de habla: Gibraltar". *Atlantis*, XVIII(1–2): 37–52.

Cal Varela, Mario (2001): *Algunos aspectos sociolingüísticos del inglés gibraltareño: Análisis cuantitativo de tres variables a nivel fónico*. Santiago de Compostela: Servizo de Publicacións, Universidade de Santiago de Compostela.

Cavilla, Manuel ([1978] 1990): *Diccionario Yanito*. 2nd edition. Gibraltar: MedSUN.

Census of Gibraltar (2012): https://www.gibraltar.gov.gi/new/census [accessed 15 June 2017].

Davies, Mark (2008-): *The Corpus of Contemporary American English: 520 Million Words, 1990-present*. Available at <http://corpus.byu.edu/coca/> [accessed 15 May 2017].

Davies, Mark (2013): *Corpus of Global Web-Based English: 1.9 Billion Words from Speakers in 20 Countries*. Available at <http://corpus.byu.edu/glowbe/> [accessed 15 May 2017].

Farrelly, Michael and Elena Seoane (2012): "Democratisation". In Terttu Nevalainen and Elizabeth C. Traugott, eds., *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 392-401.

Gal, Susan (1988): "The political economy of code choice". In Monica Heller, ed., *Code-Switching: Anthropological and Sociolinguistic Perspectives*. Berlin: De Gruyter, 245-263.

García-Castro, Laura (2017): "'I remember to read a statement from somebody on net': clausal complementation in L2 varieties of English". Paper presented at the *Santiago-Leuven-Edinburgh seminar on Grammatical Variation and Change in English*. University of Leuven, 4-5 April 2017.

Greenbaum, Sidney (1988): "A proposal for an international computerized corpus of English". *World Englishes*,7(3): 315.

Gut, Ulrike and Lilian Coronel (2012): "Relatives worldwide". In Marianne Hundt and Ulrike Gut, eds., *Mapping Unity and Diversity Worldwide: Corpus-based Studies of New Englishes*. Amsterdam: John Benjamins, 215-242.

Guy, Gregory R. and Robert Bayley (1995). "On the choice of relative pronouns in English". *American Speech*, 70(2): 148-172.

Herrmann, Tanja (2003): *Relative Clauses in Dialects of English: A Typological Approach*. Freiburg, Germany: University of Freiburg dissertation.

Hinrichs, Lars, Benedikt Szmrecsanyi and Axel Bohmann (2015): "*Which*-hunting and the Standard English relative clause". *Language*, 91(4): 806-836.

Huber, Magnus (2012): "Syntactic and variational complexity in British and Ghanaian English relative clause formation in the written parts of the International Corpus of English". In Bernd Kortmann and Benedikt Szmrecsanyi, eds., *Linguistic Complexity*. Berlin and Boston: De Gruyter, 218-242.

Huddleston, Rodney and Geoffrey K. Pullum (2002): *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hundt, Marianne (2015): "World Englishes". In Douglas Biber and Randi Reppen, eds., *The Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, 381-400.

Hundt, Marianne and Johannes Kabatek (2015): "Pseudo-titles, global language contact and discourse traditions". Paper presented at the *48th Meeting of the Societas Linguistica Europaea*, Leiden, Netherlands, 2-5 September 2015.

Jucker, Andreas (1992): *Social Stylistics. Syntactic Variation in British Newspapers*. Berlin: De Gruyter.

Kabatek, Johannes and Albert Wall, eds. (2013): *New Perspectives on Bare Noun Phrases in Romance and Beyond*. Amsterdam: John Benjamins.

Kellerman, Anja (2001): *A New New English. Language, Politics and Identity in Gibraltar*. Heidelberg: Heidelberg Schriften zur Sprache und Kultur.

Kirk, John and Gerald Nelson (2017): "The International Corpus of English (ICE) Project: The next 25 years". Plenary Lecture delivered at the *38 ICAME Conference*. Prague, Czeck Republic, 24-28 May 2017.

Kramer, Johannes (1986): *English and Spanish in Gibraltar*. Hamburg: Helmut Buske Verlag.

Le Page, Robert B. and Andreé Tabouret-Keller (1985): *Acts of Identity: Creole-based Approaches to Language and Ethnicity*. Cambridge: Cambridge University Press.

Leech, Geoffrey, Marianne Hundt, Christian Mair and Nicholas Smith (2009): *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.

Levey, David (2008): *Language Change and Variation in Gibraltar*. Amsterdam and Philadelphia: John Benjamins.

Levey, David (2015): "Gibraltar English". In Daniel Schreier, Peter Trudgill, Edgar W. Schneider and Jeffrey P. Williams, eds., *The Lesser-Known Varieties of English.* Vol 2. Cambridge: Cambridge University Press, 51-69.

Levey, Stephen (2006): "Visiting London relatives". *English World Wide*, 27: 45-70.

Lipski, John (1986): "Sobre el bilingüismo anglo-hispánico en Gibraltar". *Neuphilologische Mitteilungen*, LXXXVIII: 414–427.

LOB Corpus = Lancaster-Oslo/Bergen Corpus. Compiled by Geoffrey Leech (Lancaster University), Stig Johansson (University of Oslo) and Knut Hofland (University of Bergen).

Loureiro-Porto, Lucía (2016): "(Semi-)modals of necessity in Hong Kong and Indian Englishes". In Elena Seoane and Cristina Suárez-Gómez, eds., *World Englishes: New Theoretical and Methodological Considerations* (Varieties of English Around the World, G57). Amsterdam: John Benjamins, 143-172.

Loureiro-Porto, Lucía (2017): "ICE vs. GloWbE: Big data and corpus compilation". *World Englishes*, 36(3): 448-470.

Mendikoetxea, Amaya (1999): "Construccctiones con "se": medias, pasivas e impersonales". In Ignacio Bosque and Violeta Demonte, coords. *Gramática descriptiva de la lengua española. Vol. II Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales*. Madrid: Espasa Calpe, 1631-1722.

Meyer, Charles F. (2002): "Pseudo-titles in the press genre of various components of the International Corpus of English". In Randi Reppen, Susan M. Fitzmaurice and Douglas Biber, eds., *Using Corpora to Explore Linguistic Variation*. Amsterdam: John Benjamins, 147-166.

Moyer, Melissa G. (1993): *Analysis of Code-Switching in Gibraltar*. Barcelona: Publicacions de la Universitat Autònoma de Barcelona.

Moyer, Melissa G. (1998): "Bilingual conversation strategies in Gibraltar". In Peter Auer, ed., *Code-Switching in Conversation: Language, Interaction and Identity*. New York: Routledge, 215-234.

Mukherjee, Joybrato and Marco Schilk (2012): "Exploring variation and change in New Englishes. Looking into the *International Corpus of English* (ICE) and beyond". In Terttu Nevalainen and Elizabeth Closs Traugott, eds., *The Oxford Handbook of the History of English*. Oxford: Oxford University Press, 189-199.

Nelson, Gerald (2006): "World Englishes and corpora studies". In Braj Kachru, Yamuna Kachru and Cecil L. Nelson, eds., *The Handbook of World Englishes*. Oxford: Blackwell, 733-750.

Poplack, Shana (1980): "Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching". *Linguistics*, 18: 581-618.

Rissanen, Matti (2014): "On English historical corpora, with notes on the development of adverbial connectives". In Alejandro Alcaraz-Sintes and Salvador Valera-Hernández, eds., *Diachrony and Synchrony in English Corpus Linguistics* (Linguistic Insights 181). Bern: Peter Lang, 109-139.

Rodríguez, José Manuel (2001): "Errores comunes en el lenguaje periodístico: Invasión pasiva". *Revista Latinoamericana de Comunicación CHASQUI* 76. Available at <http://www.redalyc.org/pdf/160/16007609.pdf> [accessed 10 June 2017].

Rodríguez Pastor, César Luis and Verónica Castro Verdala (2013): "Voz pasiva perifrástica. Una propuesta didáctica". *Foro de Profesores de E/LE*,9: 143-151.

Sáez Rivera, Daniel M. (2013): "Causes and conditions for the lack of the definite article in American-Spanish headlines". In Johannes Kabatek and Albert Wall, eds., *New Perspectives on Bare Noun Phrases in Romance and Beyond*. Amsterdam: John Benjamins, 157-188.

Sanchez, Mark G. (2013): *The Escape Artist. A Gibraltarian Novel*. United Kingdom: Rock Scorpion Books.

Schaub, Steffen (2016): "The influence of register on noun phrase complexity on varieties of English". In Christoph Schubert and Christina Sanchez-Stockhammer, eds., *Variational Text Linguistics: Revisiting Register in English*. Berlin: De Gruyter, 251–270.

Schneider, Edgar (1992): "*Who(m)*? Constraints on the loss of case marking of *wh*-pronouns in the English of Shakespeare and other poets of the Early Modern English period". In Matti Rissanen, Ossi Ihalainen, Terttu Nevalainen and Irma Taavitsainen, eds., *History of Englishes: New Methods and Interpretations in Historical Linguistics*. Berlin: De Gruyter, 437-452.

Schneider, Edgar (2007): *Postcolonial English: Varieties around the World*. Cambridge: Cambridge University Press.

Seoane, Elena (2006a): "Information structure and word order change: The passive as an information rearranging strategy in the history of English". In Ans van Kemenade and Bettelou Los, eds., *The Handbook of the History of English*. Oxford: Blackwell, 360-391.

Seoane, Elena (2006b): "Changing styles: On the recent evolution of scientific British and American English". In Christiane Dalton-Puffer, Dieter Kastovsky and Nikolaus Ritt, eds., *Syntax, Style and Grammatical Norms: English from 1500-2000*. Bern: Peter Lang, 191-211.

Seoane, Elena (2016): "Telling the true Gibraltarian story: an interview with Gibraltarian Writer M.G. Sanchez". *Revista Alicantina de Estudios Ingleses,* 29: 251-258.

Seoane, Elena (2017): "A Gibraltar in the making: Interview with Dr. Jennifer Ballantine Perera". *Canarian Journal of English Studies,* 75: 217-225.

Seoane, Elena and Cristina Suárez-Gómez (2013): "The expression of the perfect in East and South-East Asian Englishes". *English World-Wide,* 34(1): 1-25.

Seoane, Elena, Cristina Suárez-Gómez and Lucía Loureiro-Porto (2016): "The ICE project looks at Iberia: The International Corpus of Gibraltar English". Paper presented at the *40th AEDEAN Conference*. Universidad de Zaragoza, 8-10 November 2016.

Suárez-Gómez, Cristina (2012): "English in contact with other European languages (Italian, Spanish, Slavic)". In Alexander Bergs and Laurel J. Brinton, eds., *Historical Linguistics in English: An International Handbook.* Vol 2. Berlin: De Gruyter, 1738-1753.

Suárez-Gómez, Cristina (2014): "Relative clauses in Asian English". *Journal of English Linguistics,* 42(3): 245-268.

Suárez-Gómez, Cristina (2015): "*The places where English is spoken*: adverbial relative clauses in World Englishes". *World Englishes,* 34(4): 620-635.

Suárez-Gómez, Cristina (2017): "The role of transparency and language contact in the structural nativization of relative clauses in New Englishes". *English World-Wide,* 38(2): 212-238.

Tagliamonte, Sali (2002): Variation and change in the British relative marker system. In P. Poussa, ed., *Dialect Contact on the North Sea Littoral*. Munich: Lincom Europa, 147-165.

The ICE Project. 2009. The design of ICE corpora. Available online at <http://ice-corpora.net/ice> [accessed April 2016].

Thomason, Sarah G. (2001): *Language Contact: An Introduction.* Edinburgh: Edinburgh University Press.

Thomason, Sarah G. and Terrence Kaufmann (1988): *Language Contact, Creolization, and Genetic Linguistics*. Berkeley and Los Angeles: University of California Press.

Tottie, Gunnel (1997): "Relatively speaking. Relative marker usage in the British National Corpus". In Terttu Nevalainen and Leena Kahlas-Tarkka, eds., *To Explain the Present: Studies in the Changing English Language in Honour of Matti Rissanen*. Helsinki: Societe Neophilologique, 465-481.

Tottie, Gunnel and Michael Rey (1997): "Relativization strategies in earlier African American Vernacular English". *Language Variation and Change*, 9(2): 219-247.

Tottie, Gunnel and Dawn Harvie (2000): "It's all relative: Relativization strategies in early African American English". In Shanna Poplack, ed., *The English History of African American English*. Oxford: Blackwell, 198-230.

Van Hout, Roeland and Pieter Muysken (1995): "Alternation, insertion, congruent lexicalizations: Corpus based and structural approaches". *Summer School on Code-Switching and Language Contact. Ljouwert/Leeuwarden, 14-17 September 1994.* Ljouwert/Leeuwarden: Fryske Akademy, 302-306.

Westin, Ingrid (2002): *Language Change in English Newspaper Editorials*. Amsterdam: Rodopi.

Weston, Daniel (2011): "Gibraltar's position in the Dynamic Model of Postcolonial English". *English World-Wide,* 32(3): 338-367.

Weston, Daniel (2013): "Code-switching variation in Gibraltar". *International Journal of Bilingualism,* 17(1): 3-22.

Weston, Daniel (2015): "The lesser of two evils: Atypical trajectories in English dialect evolution". *Journal of Sociolinguistics,* 19(5): 671-687.