Complex items and units in extra sentential code switching: Spanish and English in Gibraltar

(Article begins on next page)

26 December 2024

To whom this may concern,

This is to confirm that a paper by Eugenio Goria, titled *Complex items and units in extra-sentential code switching: Spanish and English in Gibraltar*, has been accepted for publication in an upcoming Special Issue of the *Journal of Language Contact* (titled: *Usage-based Contact Linguistics: Effects of frequency and similarity in language contact*), guest-edited by myself and Nikolay Hakimov. There are no bibliographical details for the journal issue yet, as we are in the final round of revisions, but we are on course to have the issue published in 2018.

Best wishes,

Prof. dr. Ad Backus

Complex items and units in extra-sentential code switching. Spanish and English in Gibraltar.

Abstract

As is well-known, code-mixing is particularly frequent at clause boundaries and with elements expressing pragmatic meaning. However, most of the literature has focussed on switching of morphologically simple elements such as conjunctions and discourse markers. This paper, in contrast, analyses clause peripheral switching involving two complex constructions: left dislocations and pseudo-clefts. The data are from English-Spanish bilingual conversations recorded in Gibraltar. A great majority of the bilingual constructions in the corpus belong to a few types occurring with a restricted set of lexical items. A vast amount of such highly recurrent strings in the data confirm the hypothesis that complex multi word strings that are switched together constitute units in code-mixing, i.e. they are processed together as single lexical items.

Keywords: extra-clausal code-mixing, alternation, Gibraltar, left dislocation, pseudo-clefts

0   Introduction

In this paper, I investigate the process of unitisation (Backus, 2003) as an alternative explanation for some aspects of English-Spanish code-mixing in Gibraltar that in previous analyses were mostly attributed to functional-pragmatic features (Author 2016). Coherently with other papers in the present special issue, I argue that the theoretical assumptions provided by Construction Grammar (Fillmore et al. 1988, Goldberg, 1995, 2006; Croft, 2001; Hilpert, 2014 *inter al.*) and by the usage based paradigm (Bybee, 2006, 2010; Traugott and Trousdale 2013 *inter al.*) can go a long way towards explaining the emergence of patterns in code-mixing, both providing independent alternative evidence to previous findings, and providing explanations for questions that remain unsolved by functional analyses.

Based on the view of constructions as syntactic patterns stored and processed at different levels of abstraction (Diessel, 2016), I will consider two cases of clause-peripheral code-mixing, each involving a specific syntactic construction: left dislocations and pseudo-clefts. Relying on the notion of lexical specificity (Croft and Cruse, 2004 *inter al.*), I will show that code-mixing is found more frequently with lexically specific instances of these constructions, that is cases in which the construction is associated with particular lexical items. Conversely, the phenomenon occurs less frequently with instances produced by using abstracted constructional schemas. To give an example, the transitive construction is schematic in that it involves abstract relations between predicate and arguments, regardless of the single lexical items occurring in context. However, transitive clauses like *spill the beans* or *kick the bucket* have an idiosyncratic meaning that is not predictable from the more general pattern and that depends on the lexical entries occurring in the construction; cases like these are instances of lexically specific constructions and correspond to Fillmore et al's (1988) substantive idioms (see Croft 2001: 15)*.* Moreover, constructions are often only partially filled. For example, even a highly idiomatic expression such as *kick the bucket* has a minimum degree of schematicity in that the argument slot for the subject is left open. In this paper, I argue that the different behaviour of (partially) lexically filled instances of a construction and its lexically unspecific instances in bilingual speech can be explained with reference to Backus's (2003) Unit Hypothesis: lexically filled instances of the scrutinised constructions are favoured in code-mixing because they are already available as conventionalised multi-word lexical units.

The behaviour of multi-word strings in code-mixing has been explored in the type of code-mixing that Muysken (2000, 2013) calls insertion. This type of mixing involves the presence of a matrix language providing the grammatical frame for the whole sentence, and of insertions of variable sizes whose internal structure is determined by the principles of another language, often referred to as the embedded language (see Myers-Scotton, 1993, 2006 *inter al*; Auer and Muhamedova, 2005). Backus (2003) argued that inserted multi word strings, or embedded

language islands in Myers-Scotton's framework, are often switched together because they are lexical units with a fixed form and a non-compositional meaning and emerge in the process of unitisation. In this paper, I argue that the same hypothesis applies to other types of code-mixing, namely what Muysken (2000, 2013) refers to as clause-peripheral alternation.

Given the well-known issues in distinguishing between insertion and alternation at clause boundaries, I will adopt an operational definition of alternational code-mixing. I will thus consider those cases in which a switch at a clause boundary involves a clause expressed in language A and an extra-clausal constituent (henceforth ECC) expressed in language B. The definition of ECC that I will rely on is based on Dik's (1997) account, one of the first attempts to deal systematically with this type of items. Crucially, he points out that ECCs may have different degrees of complexity, ranging from single lexical items to more complex and/or more abstract patterns. The same perspective has been further adopted recently by Kaltenböck et al. (2011), Heine et al. (2013), Heine (2013), Kaltenböck et al. (2016), who demonstrate that at the extra-clausal level, which they call *thetical grammar*, the same functional values can be equally expressed by fixed formulaic expressions, constructions and even instantaneously produced elements. Therefore, it comes as no surprise that clause-peripheral alternation may involve elements with different degrees of syntactic complexity and lexical specificity. Nonetheless, research on ECCs in code-mixing has been based on their pragmatic functions (hence their definition as "utterance modifiers", Matras, 1998), or on their social-indexical value (e.g. Poplack, 1980). Instead, I will argue in favour of an approach that takes into account both form and function of ECCs, which in this respect are not different from any other construction. I will look in particular at the frequency with which specific lexical items are used in ECC constructions.. This should help to provide a unified account of different types of alternation patterns, ranging from syntactically loose elements such as discourse markers or particles to more complex cases such as the left dislocation in (1), where the left-detached constituent, *les*

étrangers occurs in French, whereas the rest of the clause, including the co-referential resumptive pronoun *ze*, is in Dutch.

(1) French-Dutch (Treffers Daller 1994)[1]

Les	*étrangers*	ze	habben	geen	geld	he?
the	foreigners	3PL	have:PRS.3PL	no	money	DM

"foreigners, they have no money, isn't it?"

In *Author* (2016) I have highlighted the advantages of applying models such as Functional Grammar or Discourse Grammar to the qualitative analysis of clause-peripheral code-mixing, arguing that the same regularities that are observed in morphosyntactically simple constituents also hold in elements of greater complexity. This result is partly expected from a functional-pragmatic perspective because the presence of extra-clausal pragmatic functions (Dik, 1997) in ECCs urges the speakers to deal with both complex and simple items in the same way. Similarly, the *principle of pragmatic detachability* proposed by Matras (1998, 2009) states that those items that are more "detachable" from the propositional content of an utterance, regardless of their complexity, will constitute an ideal *locus* for code-mixing. However, this is not where the story ends. I suggest that even if relying on the pragmatic properties of ECCs allows us to investigate alternational code-mixing *regardless* of syntactic complexity, we still need to examine the status that complex items have in the speakers' mental representation of grammar. Pragmatic motivations do not explain why some morphosyntactically complex items behave like simpler

---

[1] To provide a better understanding of the non-English examples, I adopted a simplified version of the Leipzig Glossing Rules (https://www.eva.mpg.de/lingua/resources/glossing-rules.php). A list of the abbreviations is provided at the end of the article. The gloss was not provided in case of minimal insertions into an English clause.

ones, and some others do not. Therefore, the research question underlying this paper is the following: given the general tendency of clause-peripheral elements to be expressed in a different language than the language of the rest of the utterance, what are the cognitive factors that facilitate the switching of syntactically complex items of this type?

The rest of this article is structured in the following way. In Section 1 I briefly describe the most relevant aspects of bilingual speech in Gibraltar, in order to provide a general frame for the interpretation of the more specific issues related to the constructions under examination. In Section 2 I discuss the theoretical assumptions that underlie this study; it is argued that, the different status of lexically filled and lexically open instantiations of the two constructions might be the key to explaining the former type's higher usage frequency in code-mixing. Sections 3 and 4 describe the conducted case studies. In section 5, a final discussion summarises the results of this work.

1   Bilingual Speech in Gibraltar

Gibraltar is a small peninsula in the region of Cadiz, Southern Andalusia, which has been part of the British Overseas Territories since 1713, when the presence of English military forces dating back to a few years earlier was ratified by the Treaty of Utrecht. Since then, Gibraltar has been, more or less directly, under British rule[2]. Along with other social and political changes, English was introduced as the only official language of Gibraltar. However, the population was extremely mixed in the 18th century, with an Italian immigrant community from Liguria and Sephardic Jews constituting the majority (see the census data discussed in Levey, 2008*b*: Chapter 2). Although the distinctions between the various ethnic groups got blurred over the

---

[2] When this article was first written, no clear information was available about the future of Gibraltar after the British vote for leaving the European Union.

centuries, before the second half of the 20th century only a small part of the population had English as its first language. Instead, as Kramer (1986) argued, Andalusian Spanish was maintained as the major language of communication, mostly due to the intense commercial relationships with the neighbouring region of Spain, the presence of Spanish workers in Gibraltar, and the frequent mixed marriages.

Under these conditions, the linguistic repertoire of the past can be described as a case of extended diglossia (Fishman, 1967), with English being almost restricted to official and written uses and Spanish being used in all other contexts. This changed due to the historical events in the second half of the 20th century, which culminated in the closing of the border with Spain from 1969 to 1985. Under these circumstances, which had dramatic effects on the local population, a language shift towards English has been becoming more and more apparent, so that the linguistic repertoire at present can be best described as an "extended form of diglossia" in Auer's (2005a) terms: English has developed a local colloquial variety (Kellerman, 1996, 2001; Levey 2008*a*, 2014) and is crucially expanding its domains, including informal conversation, whereas local Spanish is more and more confined to the family and the spoken domain.

With English and Spanish being equally present in conversation, code-switching is an expected outcome of language contact. Its conversational aspects have been studied by Moyer (1998), whereas Weston (2013) focused on differences in code-switching patterns between the generations. In *Author* (2015, 2016) I showed the emergence of community-specific regularities in Gibraltar's bilingual speech: even considering the differences between three age brackets, code-mixing was found to be mainly alternational. Bilingual patterns involving different types of extra-clausal constituents (ECCs) are extremely frequent and, more importantly, they are clearly unidirectional, for example:

(2) *mira* i grew up in an environment where we had three television stations

"*look*, I grew up in an environment where we had three television stations"

(3) so you'll need to converse in spanish. *y claro ya*. that becomes a part of you and

your language your nationality

"so you'll need to converse in Spanish. *And of course,* that becomes a part of you,

and your language, your nationality"

(4) that was created *porque* the people of la_línea used to come to work in gibraltar

"that was created *because* the people of La Línea used to come to work to

Gibraltar"

Examples (2)-(4), show that alternation always involves a Spanish peripheral element (conjunctions, subordinators, discourse markers, and so on) with an English clause, whereas it almost never occurs the other way around. There are only a few isolated cases of Spanish clauses preceded by an English ECC. The emergence of this arbitrary and community-specific regularity was interpreted as a case of fusion in the sense of Auer (1999, 2014), that is, a situation where switching between the two languages loses its original pragmatic functional value and becomes only a residual behaviour, limited to a number of fixed patterns where the speakers' choice of language is heavily constrained. Single instances of (pragmatically motivated) code-switching become automatised routines and lose thus their conversational function, since switching is unavoidable at certain points. At the same time, these emerging fused lects (Auer 1999) generally acquire social-indexical values as related to particular social groups and globally opposed to monolingual practices; see O'Shannessy (2011 *inter al.*) for an extensive study of fused lects as related to younger generations.

.

One of the aims of the present study is to carry out a finer-grained analysis of the bilingual patterns mentioned above, which systematically involve a Spanish left- or right-

peripheral ECC and an English clause. Previous research has focussed on displaying and highlighting similarities between different types of peripheral elements, showing that:

(i)   regardless of their formal and functional differences, ECCs are easily switched even by speakers who do not code-switch a lot (Author 2015: Chapter 5);

(ii)   all ECCs are subject to the same constraints, since switching seems to occur only in one direction

Given these preliminary findings, I want to ask whether there are any differences between different types of ECC, and, more specifically, I will look at the way in which syntactic complexity can favour or disfavour the emergence of bilingual ECC-clause patterns. The data on which the analysis is based are described extensively in Author (2015). They were collected during two fieldwork sessions in 2013, yielding a total of over 20 hours of audio-taped semi-structured interviews. The data were initially meant to be used for qualitative analysis of the community-specific features of code-switching, so the sample was not built specifically for quantitative analysis. However, age was considered a relevant factor influencing code-mixing (see Kellermann, 2001; Levey, 2008*b*; Weston, 2013). For this reason, three age brackets were targeted during data collection: "young" speakers, below 30 years old, "adult" speakers, between 30 and 60 years old, and "old" speakers, over 60 years old[3].

2   Syntactic Complexity in Bilingual Speech

2.1 The unit hypothesis and alternational code mixing

According to the view adopted by usage-based linguistics (Croft 2001; Bybee 2006, 2010, 2013), constructions represent complex form-meaning associations whose meaning cannot be predicted by the semantics of the single elements involved. A distinction is made between

---

[3] Since the subcorpus from the "over 60" age bracket initially contained much more monolingual speech than expected, this part of the sample was further expanded during the second fieldwork session.

lexically open (or schematic) constructions, that can be filled with any contextually appropriate lexical material (e.g. the English causative construction), and lexically filled (or substantive) constructions, in which all the elements forming the construction are fixed (ex. *to cut a long story short*). However, "schematic [constructions][4] vary considerably in their schematicity" (Croft and Cruse 2004: 248), and in fact most constructions are partly schematic and partly fixed, in that they involve both lexically filled and lexically open slots. For example, the *way*-construction analysed in Goldberg (1995), as in *Frank kicked his way out of the bar*, has fixed slots represented by the lexical element *way*, the preceding possessive and the following prepositional phrase, but also open slots corresponding to the subject, the predicate and so on.

Research within the usage based paradigm (e.g. Bybee 2010) has provided a theory of how constructions, as elements of variable syntactic complexity, are dealt with in human cognition. Through repetition in use, linguistic representations of lexemes occurring in a given construction develop stable sequential relations, so that the whole string becomes stored in memory as a chunk. This holds true both for fully specific combinations, such as *break a leg!*, as well as for those that have an internal structure, such as *to lend X a hand*. Chunking is thus the process that continuously shapes the lexicon/grammar, allowing for multi-word elements of any complexity to be autonomously stored. This has important consequences on how constructions with different degrees of specificity are treated by bilingual speakers. Research in this direction led to the formulation of the Unit Hypothesis (Backus 2003), which states that "if two or more EL [embedded language] morphemes co-occur in CS [code-switching], and they form a conventional combination in the EL, then it would be too coincidental if the speaker had produced them as two or more independent switches" (p. 91). In other words, multi-word strings that constitute units in monolingual speech, and which have thus lexical status regardless of their internal structure, are expected to maintain this status also in bilingual speech: this

---

[4] I have replaced the term *idiom,* used in the original quote, with *construction*, to avoid terminological redundancy.

would make multi-word strings available for transfer from one language into the other under the same conditions as those that hold for simple lexical items. Now, the Unit Hypothesis has proven useful to describe different types of complex switches, involving elements such as nominal compounds (Backus 2003), plural nouns (Backus 2003, Hakimov 2016*a*), prepositional phrases (Hakimov 2016*b*). Even outside the usage based framework, studies like Poplack and Meechan (1995), among others have more or less intuitively referred to the notion of idiomaticity in meaning as a factor that may influence the switching of multi-word units. Most of case studies, though, describe cases of intra-clausal insertions in Muysken's (2000) taxonomy. In this paper, instead, I show that the notion of unit, being a feature of speech production in general, is also applicable to alternational code-mixing. Specifically, I will analyse the case of clause-peripheral alternation, which involves the juxtaposition of two structures, one of them typically being a clause and the other an extra-clausal constituent in the sense of Dik (1997). The extra-clausal constituent can thus be filled by units of variable complexity: monomorphemic elements such as discourse markers (see Heine 2016 for an overview) represent the best-known case, but complex elements are not infrequent, either as fixed multi-word expressions (such as the *second level discourse markers* analysed by Siepmann 2005), or as schematic patterns, such as the pseudo-cleft and left dislocation constructions, which are then filled by lexical elements during the conversation.

This paper thus seeks to bring evidence from new data in favour of the Unit Hypothesis in two ways: first, by applying the hypothesis to previously overlooked bilingual patterns involving alternation; second, by considering not only fully lexically specific instantiations of constructions, but also other instances displaying degrees of schematicity. This aspect will be discussed in detail in Section 2.2.

2.2 Syntactic complexity in extra-clausal constituents

A usage-based approach to alternational code-mixing has several advantages. If one relies only on pragmatic analysis, functional accounts of ECCs leave a number of questions that cannot be fully answered. Dik's (1997) account of ECCs downplays the importance of syntactic complexity, since he treats both simple and complex items on an equal basis if they have the same function. More recently, works within the framework of Discourse Grammar as proposed in Kaltenbök *et al.* (2011), Heine *et al.* (2013), Heine (2013), Kaltenbök *et al.* (2016), have tried to reach a more accurate formalisation of the different types of phenomena occurring at the periphery of the clause, addressing syntactic complexity in a more specific way. Kaltenbök *et al.* (2011) argue for a major distinction between two domains of grammar, namely Sentence Grammar and Thetical Grammar. The latter is the one involved in the production of discourse-relevant entities, typically occurring at the periphery of the clause, which they call *theticals.* Three levels of complexity are considered in this sector of grammar. The first is the level of *formulaic theticals*, which are generally monomorphemic unanalysable units, as is the case with the discourse marker in example (5). The second is the level of *constructional theticals*, which are defined as "recurrent patterns or constructions of theticals, being compositional but having some schematic[5] structure and function" (Kaltenböck *et al.* 2011: 875); these elements, of which an example is given in (6), correspond to constructions in usage-based linguistics. The third is the level of spontaneous theticals, see example (7). These items are fully compositional and built with the principles of sentence grammar; they represent local interactional strategies and are used only occasionally.

(5) The therapist's level tone is bland and neutral — he has, *for example*, avoided

stressing 'you'*.*

(Huddleston and Pullum, 2002: 1355, quoted in Kaltenböck *et al.,* 2011)

---

[5] Note that *schematic* is not used here in exactly the same way as in Construction Grammar works, and it points to the presence of morphosyntactic specifications on the structure of theticals.

(6) *(As for) Peter*, he no longer is my friend. (Kaltenböck *et al.,* 2011: 856)

(7) Yes there was one woman particularly *I've forgotten what she was called* in this

group (ICE-GB:s1a-044-349; quoted in Kaltenböck *et al.* 2016*b*: 8)

In *Author* (2016) I have argued for the application of this framework to bilingual data, showing that it leads to interesting generalisations about alternational code-mixing. Namely, it allows to demonstrate that differently complex but functionally equivalent peripheral items behave in the same way in bilingual speech. Switches of different sizes can thus be considered instances of the same alternation pattern, while in the code-switching literature they have sometimes been labelled differently, for instance as cases of extrasentential and intersentential code-switching. However, Thetical Grammar alone cannot predict the probability with which theticals of a specific level occur in bilingual speech, nor to account for differences in behaviour of elements belonging to the same class. If we focus on constructions, a finer grained analysis is needed in order to explain why there are instances of a construction that are evidently more frequently switched than others and how this difference is related to the organisation of constructions in the bilingual's lexicon/grammar.

The main claim of this paper is that lexically specific instances of a given construction are likely to be stored as lexical chunks and thus more accessible in speech productionFor this reason, they represent better candidates for code-mixing than instances that involve filling in the schematic template. A good example of this distinction between specific and schematic is represented by pseudo-cleft constructions, whose behaviour in bilingual speech is illustrated in Section 4. Pseudo-clefts constitute instantiations of the schematic pattern represented in (8), which has open slots that can be filled by any appropriate element, with very few constraints, as is illustrated in (9) (see also Collins, 1991).

(8) *wh-* X *is/was Y*

(9)   a. what *I lost* was *my wallet*

b.what *Sean wanted* was *a new life*

c. what *the burglars took away* was *an old gift from my brother*

However, as pointed out among others by Günthner and Hopper (2010), lexically-specific instantiation of this general pattern have frozen into nearly-fixed lexical chunks such as the ones given in examples (10)-(12).

(10) and so *what happens is,* your mic runs into i:t, and then it plugs into your receiver

(adapted from Günthner and Hopper 2010: p.12)

(11) it's really central government really that (.) that is obviously all the time trying to make our life difficult really but (.) *the thing is that* we have to put up with it

(fieldwork data. Author 2015)

(12) *what I'm saying is*, don't get married in Vegas. Take your time.

 (2007 COCA: MAG quoted by Brinton 2014)

According to Backus's (2003) Unit Hypothesis, it is possible to assume that lexically specific instances of this construction are stored as units in the mental lexicon and are thus more accessible in speech production with respect to cases like (8) that represent productive instantiations of a schematic pattern. Due to this greater accessibility, they are also expected to occur more frequently in bilingual speech. The same reasoning may in principle be extended to partially specific constructions in that they may also count as units. This is the case, for example, with the analysis of bilingual left dislocation to be discussed in Section 3.1, where a distinction is drawn between productive uses of the schematic template and partially specific instantiations of the same construction. The latter case shows greater constraints in that only a limited set of

expressions may occupy the left-peripheral slot, but the whole construction is not fully lexically specific.

My claim is thus that in addition to single lexemes, two types of units are relevant in code-mixing: lexically specific chunks and partially specific constructions. This view is supported by the findings obtained from the two case studies presented in the following sections.

3   Bilingual left dislocations

Bilingual left dislocations (LDs henceforth) are an interesting, and yet neglected, case of clause-peripheral alternation involving a schematic construction. LDs are constructions in that they represent complex syntactic structures that express interpersonal pragmatic functions and clause-oriented informative functions in a non-compositional fashion. Moreover, LDs are interesting for the analyst because they allow two levels of generalisation corresponding to two different subtypes of the construction. The first one corresponds to what will be referred to as nominal LD, and has a fully schematic structure; the second one will be referred to as pronominal LD, and is partially lexically specific. In this Section I show that bilingual left dislocations occur more frequently as instances of codeswitching when the expression has a higher degree of specificity and corresponds to a lexical chunk, and less frequently when the construction is lexically open, and thus has a greater degree of schematicity. In fact, most bilingual LDs involve only a few highly frequent lexical items.

3.1 Formal and functional characteristics of left dislocation

The relevant constructions were identified in the Gibraltar corpus on the basis of both formal and functional features of LDs. Concerning formal parameters, Lambrecht (2001*a*) indicates four major features that cross-linguistically characterise LD. They are listed in (13) below:

(13)   Defining features of the dislocation construction

(i) Extra-clausal position of a constituent

(ii) Possible alternative intra-clausal position

(iii) Pronominal co-indexation

(iv) Presence of a special prosodic contour

According to this definition, LD involves thus the presence of a constituent, typically an NP, occupying a clause-peripheral slot, and coindexed by an anaphoric element inside the clause. In (14) below, thus, *al partido carlista* ("the Carlist party") is a dislocated NP that occurs outside the clause, and it is anaphorically coindexed by the weak pronoun *lo* inside the clause.

(14)   Al   partido carlista dicen   que

DOM   party   carlist   say:PRS.3PL   COMP

no   lo   legalizaron   para   las   elecciones

NEG   3SG.M.OBJ   legalise:PST.3PL   for   the   elections

'the Carlist party, they say that they did not legalise it for the elections' (Rivero 1980)

As for functional properties, LD has both informative and pragmatic-interactional values. Informatively, the construction marks the dislocated element as a Topic in the sense of Lambrecht (1994) and Dik (1997), and is often used to retrieve elements that are either textually or contextually accessible, but not topical at a certain point in the discourse (Hidalgo 2002). Furthermore, Duranti and Ochs (1979), and more recently Geluykens (1992), Hidalgo (2002) and Hidalgo Downing and Hidalgo Downing (2007), have argued that, particularly in Romance languages, LDs also have more global interactional properties, both interpersonal and metatextual. These constructions allow negotiating "difficult" and potentially face-threatening

operations such as keeping the floor, shifting the topic, and introducing the speaker's perspective.

It is crucial for the present analysis to observe that, even if informative and interactional functions cannot be completely disentangled, differences in function tend to be reflected in differences in form, so that it is possible to consider two subtypes of LD (see also Author 2016). The first one will be referred to as 'nominal LD', and it is characterised by clause-oriented informative functions, namely that of marking the sentence Topic, either introducing a new one or retrieving a previously activated referent. The dislocated constituent in this variant of the LD construction can be any NP; an example is given in (15).

(15) nominal left dislocation (Hidalgo Downing and Hidalgo Downing, 2007: 329)

oye	y	*la*	*revista*	*esta*	¿cada cuánto	*la*
DM	and	the	journal	this	how_often	3sg.F.OBJ

sacáis?
publish:PRS.2PL

"Listen, and this journal, how often do you publish it?"

The second type will be called 'pronominal LD'. This variant often expresses discourse-oriented interactive functions (see Duranti and Ochs 1979, Downing 1997, Hidalgo Downing and Hidalgo Downing 2007). In Spanish, pronominal LD involves only a restricted set of elements, namely personal (16a) and demonstrative (16b) pronouns.

(16) pronominal left dislocation (Hidalgo Downing and Hidalgo Downing, 2007: 342)

a. yo	mucha	gente	me	ha	comentado
1SG.SBJ	many	people	1SG.OBJ	have.AUX:3SG	tell

"As for me, many people told me"

b. eso   yo   creo   que   deberías

this.M   1SG   believe:PRS.1SG   COMP   must:COND.2SG

hacer=lo

do:INF=3sg.M.OBJ

"This, I think you should do it"

According to Hidalgo Downing and Hidalgo Downing (2007), pronominal LD is particularly frequent in Spanish: they observe that of all instances of dislocation in the *Corpus oral de referencia de la lengua española contemporánea* (Marin 1992), 40% is represented by pronominal LDs. Moreover, both Downing (1997) and Hidalgo Downing and Hidalgo Downing (2007) indicate the 1sg pronoun *yo* and demonstratives *eso* and *esto* are the most frequently found forms. In an exemplar model, as discussed in Bybee (2010), these would correspond to the prototype for pronominal LD. In contrast, no highly frequent lexical element was found for nominal LD, though the possibility cannot be ruled out that there are such frequent forms. This discrepancy between the two variants of the construction is of course biased by the different sizes of the class of nouns and pronouns, with possible effects also on their mental representations; nevertheless, this distribution suggests overall greater specificity of pronominal LD and greater schematicity of nominal LD.

Another issue that must be dealt with concerns the comparability of the LD construction across languages, and more specifically in the two contact languages considered here, English and Spanish. Differences between Romance languages and English were identified by Cinque (1990), Rizzi (1997) and Hidalgo (2000), who has argued that "[t]he form and use of [LD]s in English and Spanish is similar only in the prototypical form of the construction and in the main, general functions", while the two languages differ with regard to the functional values of the construction. In English, Pronominal LD has in fact a contrastive value (Prince 1981, 1984) that is not attested in Romance languages, as in example (17).

(17) *That silly season when everybody loves everybody else. Me - I'm different! Everybody hates me and I hate everybody* [Carl Barks, Christmas on Bear Mountain. "Four colour comics". Dec. 1947)]

To put it differently, LD has the same informative function in both English and Spanish, namely that of signalling a sentence Topic; the two languages differ, though, in the pragmatic uses of the construction. With respect to this issue, the analysis presented here is form-based in that the instances of the LD construction in the English-Spanish data have been considered in relation to Lambrecht's (2001) interlinguistic definition and without consideration of their pragmatic function.

To sum up, two variants of the LD construction have been identified: nominal and pronominal LD. From a constructionist perspective, a nominal LD represents a schematic instance of the construction in that the extra-clausal slot is lexically open and admits any informatively adequate NP.The pronominal LD on the contrary is more lexically specific in that it allows only a restricted set of items, namely personal and demonstrative pronouns, and has a few highly frequent instantiations. Moreover, the two different forms are also related to partially different functions: while nominal LD expresses a clause-oriented informative function, namely that of marking the Topic, pronominal LD has developed additional pragmatic and discursive values. Hence, greater lexical specificity goes along with functional specificity. For what concerns the behaviour of this construction in bilingual speech,

I will show that lexically specific instances of the LD construction that constitute lexicalised chunks are also more frequently switched from one language into the other.

3.2 Data analysis and results

The analysis of LDs in the corpus was preceded by manual annotation of the relevant constructions with the software ELAN[6] (Sloetjes and Wittenburg 2008). Based on the theoretical assumptions outlined above, two types of LD were identified: pronominal LDs, involving personal pronouns, demonstratives and possessives, and nominal LDs, involving any nominal constituent. With regard to the language of the analysed construction and the adjoined clause, four possible patterns were considered, assigning a separate language value for the dislocated constituent and the following clause. The bilingual speaker may produce both the dislocated constituent and the following clause in the same language, English or Spanish, and in this case the LD is labelled as monolingual, or she may switch the language at the juncture between the dislocated constituent and the clause. The monolingual patterns were labelled "e-e", or "s-s"; the bilingual patterns were labelled "s-e" if a Spanish dislocand precedes an English clause and "e-s" if an English dislocand precedes a Spanish clause. Therefore, eight possible combinations between types of dislocation and types of bilingual patterns are exemplified below:

(18) English monolingual (nominal) LD

you go there even if we're british [the cultural shock]$_i$ you have [it]$_i$

(19) Spanish monolingual (nominal) LD

[y	la	pequeña]$_i$	que	va	vení	dentro	una	semana
and the		youngest	COMP	go:3SG	come:INF	within	a	week

[ella]$_i$	tiene	tres	hijos
3SG.SBJ	have:3SG	three	son:PL

"And the youngest, who's going to come within a week, she has three sons"

(20) "s-e" bilingual nominal LD

---

[6] ELAN is a software provided by Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands (https://tla.mpi.nl/tools/tla-tools/elan/).

sin embargo		ahora	[mucho'	niño'	jovene']ᵢ
no doubt		now	many	children	young

[they]ᵢ 're writing with americanisms   tambien

also

"No doubt now *many young children* are writing also with americanisms"

(21) "s-e" bilingual pronominal LD

you 'd be really really struck *po'que* [yo]ₗ [i]ᵢ'm very innocent with those things

"You'd be really really struck, *because me*, I'm very innocent with those things"

(22) "e-s" bilingual nominal LD

and [the next child he caught]ᵢ

[eso]ᵢ	era	él	que	se	quedaba
this.M	be:PST.3SG	3SG:SBJ	REL	3SG.REFL	stop:PST.3SG

"And the next child he caught, *that was the one who would stop*"

(23) "e-s" bilingual pronominal LD

cosa como son	[us]ᵢ	economically	'tamo	demasiado	bien
as things stand			be:PRS.3PL	too	good

"As things stand, *us* economically *we are way too good*"

The total number of LDs found in the corpus is 81. The analysis of the distribution of the different constructions, which is given in Table 1, yielded the following results.

			Pattern		
	e-e	s-s	e-s	s-e	total
nominal LD	14	5	2	14	35
pronominal LD	2	23	2	19	46

Total	16	28	4	33	81

Table 1: Distribution of nominal and pronominal left dislocation

As for monolingual utterances, Spanish LDs appear to be more frequent than English ones, confirming the view expressed in Hidalgo (2000) about a broader diffusion of the construction in Spanish. Also, English shows a strong preference for nominal LDs (14 out of 16 tokens), whereas Spanish clauses more frequently contain pronominal LDs (23 out of 28 tokens). This too is partly expected in line with the literature discussed in Section 3.1, according to which Spanish pronominal LD was found, in some cases, to be predominant in number over nominal LD and associated with several interactive functions, whereas English LD is much more marginal. As for bilingual LDs, the "s-e" type, occurring 33 times, outnumbers the "e-s" type. This tendency is also observable in all cases of clause-peripheral alternation in the Gibraltar corpus: with very few exceptions, bilingual patterns involve a Spanish extra-clausal constituent and an English clause. In Author (2015, 2016) I have argued that the same unidirectionality applies also to discourse markers, conjunctions and subordinators[7]. We also observe that pronominal LDs with a Spanish dislocand are more frequent than nominal LDs both in monolingual and bilingual sentences.

The picture drawn so far suggests a slight preference for pronominal LDs when the dislocand is Spanish, and the absence of a similar pattern with English dislocands. This tendency can be better understood if one considers the single lexical items involved in LD, as represented in Table 2.

	type of pattern		
dislocand	**Monolingual**	**bilingual**	**total**
yo	6	12	18

---

[7] Unidirectionality in code-mixing patterns is also a diagnostic feature of incipient fusion in the sense of Auer (1999, 2014), in that it shows that code-mixing is restricted to a number of fixed patterns.

eso	11	5	16
people	4	0	4
ello'	2	1	3
nosotros	2	1	3
proper nouns	3	0	3
la gente	0	2	2
la' persona'	0	2	2
other spa	7	10	17
other eng	9	4	13
total	**44**	**37**	**81**

Table 2: Monolingual and bilingual dislocations for single lexical items occurring more than once

The table lists the different instantiations of the first slot in the construction and their frequencies in monolingual and in bilingual combinations. The non-recurring lexical items were coalesced into two categories "other Spanish" and "other English". It can be observed that English elements are underrepresented in LD constructions, with only the semantically general noun *people* occurring more than once. In contrast, the Spanish types that recur in LDs, according to Table 2, belong to a few well identifiable classes: personal pronouns (*yo, ello', nosotro'*), demonstratives (*eso*), semantically general nouns (*gente, personas* "people") and proper nouns. It is also important to note that all these elements represent prototypical sentence Topics. Furthermore, the two most frequent lexical types that enter the construction are the Spanish pronouns *yo* and *eso*, which together represent 42% of all LDs and 46% of all bilingual LDs. This distribution seems thus to mirror the distribution found in monolingual Spanish: Hidalgo Downing and Hidalgo Downing (2007: 333) observe that: "[i]n Marcos Marín's *Corpus oral de referencia de español contemporáneo* (1992), pronouns amount to the 40% of

detachments. Of these pronominal detachments, 68% corresponds to personal pronouns, whereas the remaining 31.8% is composed of demonstrative pronouns".

This parallelism between monolingual and bilingual data supports the hypothesis that lexical specificity of particular instances of the LD construction is one of the major factors influencing the production of bilingual LDs. The items that occur more frequently in bilingual speech, namely the 1st person singular pronoun *yo* and the proximal deictic *eso* are highly frequent items also in monolingual speech.

To conclude, the analysis of bilingual left dislocation in my corpus supports the view that code-mixing, and in particular the emergence of patterns such as "*yo* + English clause" or "*eso* + English clause", is highly affected by the frequency with which lexical items appear in the construction. Given a schematic pattern such as the LD construction, it has been demonstrated that those subtypes of the construction that are attested in monolingual Spanish as lexicalised routines are also the ones more frequently used in code-mixing.An assumption behind this effect is that lexicalised intantiations are more accessible to the speakers than purely schematic patterns." To conclude, a constructionist approach to bilingual LDs is complementary to the observations made on the same topic (among others, Author, 2016). Besides being more frequent in monolingual speech and more lexically specific, the subtypes of LD that are switched more frequently are also the ones that typically express interpersonal and discourse-oriented functions, and that are thus more pragmatically detachable (Matras, 1998, 2009) from the propositional content of the utterance.

4 Bilingual pseudo-clefts

Pseudo-cleft sentences are another example of a partially schematic construction that can have lexically filled and lexically open slots. These constructions are used to express pragmatic and

informative values of an utterance in a non-compositional way, i.e. through the realisation of a simple propositional content in a biclausal construction involving a relative clause and a copular clause. As occurs with other elements with similar functions, pseudo-clefts in my corpus often appear in bilingual utterances, with one element of the construction being realised in a different language than the other, which results in language alternation. In this section I analyse this type of code-mixing in the light of the Unit Hypothesis, and I argue that clause-peripheral alternation is more frequent with those instances of pseudo-clefts that have a higher degree of lexical specificity, and which therefore constitute lexical chunks, whereas more schematic instantiations of the construction will tend to be monolingual.

4.1 Form and function of pseudo-clefts

Pseudo-clefts, as part of a broader family of cleft constructions (e.g. Collins 1991, Lambrecht 2001*b*, Traugott, 2008, Traugott and Trousdale 2013, Patten 2012), are characterised by a number of formal and functional properties; some of them shared with related constructions, some others specific to this particular type. Here, I will only concentrate on those aspects of pseudo-clefts that are relevant for their identification in the bilingual corpus.

Pseudo-clefts are biclausal constructions with a headless relative clause, often introduced by a *wh-* element, in first position, followed by a copular clause. The complement of the copula is coreferent with the relative clause and often expresses the pragmatic function of Focus, whereas the relative clause contains presupposed content, for example:

 (24) a. what we are talking about is your future

      b. we are talking about your future

As can be seen, (24a) has the same propositional content as (24b), but the two have different informative values. The pseudo-cleft in (24a) has the function of marking the constituent *your future*, which occurs as the complement of the copula, as the Focus of the sentence, and at the same time it signals that the content of the relative clause is presupposed or retrievable from the context (see Lambrecht 1994). From this definition, pseudo-clefts are thus schematic form-function pairings involving two fixed structural positions, one for the subordinator introducing the relative clause and one for the copula, and two open slots, one for the relative clause and one for the argument of the copula, as represented below:

(25) WH- x BE y

The element indicated as WH- in (25) corresponds to a relative-like subordinator. As recently argued by Patten (2012), this slot has three main realisations, corresponding to three subtypes of the pseudo-cleft construction, which are referred to as WH-clefts, TH-clefts and ALL-clefts; see examples (26a-c).

(26) a. What I like best is grape soda (Patten 2012: 62) [WH-cleft]

　　b. The thing that I like best is grape soda (Patten 2012: 62) [TH-cleft]

　　c. All he drinks is grape soda (Patten 2012: 62) [ALL-cleft]

However, since ALL-clefts have to be regarded as a more specific and in a way a less prototypical instance of the construction (see Traugott 2008), my analysis will address only the first two types. The pseudo-cleft constructions exhibits obvious similarities across English and Spanish, both in form and in function; the structure is thus highly comparable between the two languages in contact. As for *wh*-clefts, the most important difference between English and Spanish is the form of the *wh*- element. In English, a personal pronoun and a relative marker are fused into a single portmanteau form, which is commonly referred to as fused relative (Agar Marco 2014);

in Spanish, like in most Romance languages, the *wh*-element is expressed analytically by a personal pronoun followed by a relative pronoun, with the neuter form *lo que* being the default case; for instance:

(27) lo   que   quiero   es   un   coche

3SG   REL   want:PRS.1SG be:PRS.3SG   a   car

"what I want is a car" (Fernández Soriano 2009: 88; my glosses)

Th-clefts have the same structure in English and Spanish. According to Collins (1991), the *wh*- element is formed by a semantically general noun such as *thing*, *time, place, one, way* and so on, followed by a relative clause. Examples (26b) and (28) show that TH-clefts are structured similarly in English and Spanish.

(28)   *la persona* a la que mas echa de menos es a su madre

"*the person* that he misses most is his mother" [Agar Marco (2014: 180)]

Both in English and Spanish, the second argument of the copular clause indicated as *y* in (25), can be a nominal element, as in (26a-c), as well as a clausal one, as in (29) below.

(29) all the companies are here. and what happens is that they have (.) a lot of (.) workers that

do a lot of work for customer services o:r (.) or IT work [fieldwork data. Author 2015]

It must be noted that pseudo-clefts that introduce a nominal constituent typically express a clause-internal function of focus-marking, while the ones introducing a clausal complement do not have this function. As Hopper and Thompson (2008) and Traugott and Trousdale (2013) have

26

argued, this variant of the construction expresses a discourse-oriented cataphoric function, in that it is used to project (Auer 2005) an upcoming unit of talk. This variant of the construction has been also identified for Spanish pseudo-clefts, for example by Curnow and Travis (2004) and Fernández Soriano (2009). Interestingly, works on both languages (see Fernández Soriano 2009, Reig Alamillo 2011 and Pannunzi 2016 for Spanish and Hopper and Thompson 2008 for English) point out that constructions expressing a discourse-oriented function are strongly associated to a few highly frequent lexical types, such as *what happens is* in English and *lo que pasa es (que)* in Spanish. Therefore, a difference in function is reflected by greater specifications on one of the constructional slots, in that the first member of the construction is often represented by a lexicalised chunk that constitutes a unit in Backus's (2003) sense.

Moreover, an even greater level of fixedness is found in connectives such as *es que* in Spanish and *(the) thing is* in English, as in examples below:

(30) it's really central government really that that is obviously all the time trying to make our life difficult really but *the thing is that* we have to put up with it [Fieldwork data. Author 2015]

(31) that's what i wanted to say *é que   no   me   acuerdaba*

                              it's that   NEG   REFL.1SG   remember:PST.1SG

     "that's what I wanted to say, *it's that I didn't remember*" [Fieldwork data. Author 2015]

As discussed for Spanish by Travis (2005), these items are diachronically related to pseudo-clefts and in particular they retain the same projecting function that has been identified also in other languages (Günthner 2011). From a formal perspective, they represent an outcome of chunking, in that they are lexical units that no longer admit internal lexical variation.   To sum up, two

variants of the pseudo-cleft construction can be distinguished. On the one hand, there are schematic instantiations resulting from the pairing of a headless relative clause with a copular clause, with few constraints on the elements that can enter the construction. These constructions are related to a local function of focus marking and typically take a nominal argument. On the other hand, lexically specific instances of the pseudo-cleft construction involve the presence of a few highly frequent verbs in the relative clause, meaning "do", "say" or "happen". Such constructions typically take a clausal argument and express global discourse-oriented functions, in that they serve to project an upcoming propositional content. The same holds for the two fixed expressions *es que* and *the thing is*. In the following Section I will give empirical evidence that alternational code-mixing between the relative clause and the copular clause is more likely to occur if the first member of the construction is a lexical chunk, i.e. code-mixing tends to involve lexically specific instances of the construction.

4.2 Data analysis and results

Relying on the description of the pseudo-cleft constructions given in Section 4.1, I have considered in this study the cases involving a switch between the first and the second member of the construction, i.e. between the relative clause and the copular clause, as in (32).

(32) *lo   que   pasó   era   que*

    3SG   REL   happen:PST.3SG   be:PST.3SG   COMP

  they started the youth as a normal thing

   "what happened was that they started the youth (club) as a normal thing"

I have hypothesised that the occurrence of a bilingual pattern is favourably influenced by the high frequency of the verb occurring in the relative clause. In other words, highly frequent verbs would give rise to lexically specific instances of the construction; in these cases, the first member corresponding to the relative clause is a nearly fixed lexical chunk. According to Backus (2003),

these units should be considered the best candidates for switching. Such lexically specific instances of the construction are also typically related to pragmatic, or discourse-oriented, functions, whereas schematic instantiations frequently express clause-internal functions.

The analysis followed the same methodology as was used for left dislocations. According to the language in which each component of the pseudo-cleft construction was realised, the constructions were given different labels. Patterns where both the relative and the copular clause were in the same language were labelled "e-e" (English monolingual) or "s-s" (Spanish monolingual); see (33a) and (33b). They were labelled as "s-e" in cases were the relative clause occurs in Spanish and the copular clause in English, as in (34a) and as "e-s" when the relative clause occurs in English and the copular clause in Spanish, as in (34b).

(33a) Spanish monolingual pseudo-cleft

tu	*lo*	*que*	*tiene*	*decí*	*a*	*la*	*gente*
2SG:SBJ	3SG	REL	have:PRS.2SG	say:INF	to	the	people

é	*mira*	*e'to*	*é*	*lo*	*que*	*tu*
be:PRS.3SG	DM	this	be:PRS.3SG	it	REL	2SG.SBJ

puede'	*hacé*
can:PRS.2SG	do:INF

"You, what you have to tell the people is: «look, this is what you can do»"

(33b) English monolingual pseudo-cleft

what they sell to their nation is "no no, gibraltar is a tax heaven"

(34a) | *la* | *gracia* | *y* | *la* | *suerte* | *que* | *tenemo'* | *é que* |
|------|----------|-----|------|----------|-------|-----------|--------|
| the | flair | and | the | luck | REL | have:PRS.1PL | is that |

el llanito is funny in itself

"*And the flair and the luck that we have is that the* llanito is funny in itself"

(34b) so basically what you do is

te	queda'	en	casemates	and
2SG:REFL	stay:PRS.2SG	in	Casemates	

te	toma'	una	cerveza
2SG:REFL	take:PRS.2SG	a	beer

"So basically what you do is, *you stay in Casemates and you have yourself a beer*"

In the next step, the constructions were divided according to whether they introduced a nominal or a clausal argument. The frequency distribution of each type is given in Table 3.

Pattern

ARGUMENT	?	s-s	e-e	e-s	s-e	TOTAL
CLAUSAL	2	51	25	3	32	113
NOMINAL	0	26	3	0	1	30
TOTAL	2	77	28	3	33	143

Table 3: Distrbution of pseudo-clefts taking clausal and nominal arguments.

As can be observed, constructions taking clausal arguments are clearly predominant in the data, and that bilingual constructions represent 25% of the observed constructions (36 out of 143 tokens). As for bilingual patterns, two clear tendencies can be identified. The first tendency is the general preference for "s-e" over "e-s" (33 *vs* 3 tokens): recall that this was found to be the dominant pattern for all types of extra-clausal constituents (see Author 2015, 2016). What makes pseudo-clefts particularly interesting is that they are not fully-fledged extra-clausal

constituents like discourse markers or conjunctions, but constructions with a complex syntactic structure and exhibiting varying degrees of schematicity. This shows that the first member of a pseudo-cleft, the relative clause, is treated in the same way as other lexical elements and shows the same regularities. The second interesting finding in the distribution represented in Table 3 is that bilingual patterns are observed almost only in the case of constructions taking a clausal argument (35 out of 36 tokens). In other words, those instances of pseudo-clefts where the relative clause has the function of projecting an upcoming full utterance are more likely to have a switch between the projector and the following clause. This would confirm what has been argued in functional analyses (e.g. Matras 1998), namely, that elements expressing pragmatic functions are more available for switching than those expressing clause-internal functions. Again, pragmatic motivations that facilitate the occurrence of particular items in bilingual speech seem to go along with structural motivations, and particularly with their degree of lexical specificity. The following analysis supports this claim.

PATTERN

LEXICAL TYPES	?	MONOLINGUAL	BILINGUAL	TOTAL
es que	2	23	16	41
pasar	0	16	8	24
tener	0	1	3	4
thing is	0	17	2	19
hacer	0	6	2	8
do	0	1	1	2
decir	0	6	0	6
afectar	0	2	0	2
happen	0	2	0	2
querer	0	2	0	2
ser	0	2	0	2

| other | 0 | 16 | 1 | 17 |
| Total | 2 | 102 | 33 | 137 |

Table 4: Frequency distribution of single lexical items occurring more than once in monolingual and bilingual pseudo-clefts.

Table 4 shows the frequency distribution of all the lexical types[8] occurring more than once as the predicate of the relative clause. The data show that both monolingual and bilingual instances of the construction are strongly associated either to the connectives deriving from reduced pseudoclefts, such as *es que* ("it is that") and *the thing is,* or to a small set of recurrent verbs, which correspond to lexically specific instantiations of the construction. Bilingual constructions, in fact, are found most frequently with the *es que* connective (16 over 33 tokens), and with the lexical type *pasar* "happen" (8 over 33 tokens), all of which correspond to the lexical chunk *lo que pasa es* "what happens is". This finding is thus in line with the examined hypothesis that those instances of the construction that are more directly associated with single lexical elements, or that have already developed into formulaic expressions in monolingual speech, are more likely to function as units in code-mixing, as opposed to their schematic counterparts that require the productive application of a lexically open template.

5   Discussion

The present case studies allow to make a few more general remarks. The aim behind the quantitative evaluation of the data of the two constructions was to provide further evidence in favour of the Unit Hypothesis (Backus 2003), showing that it applies not only to lexical insertions but also to specific types of alternational code-mixing. Positive evidence from Spanish-English

---

[8] I chose to consider lexical types instead of surface forms in order to consider predicates occurring in the present tense and in the past tense as instances of the same type.

code-mixing in Gibraltar was presented in both cases. The bilingual constructions that rank higher in the frequency lists discussed in Sections 3 and 4 are related to a few lexical items, corresponding in the case of dislocations to the 1st person pronoun *yo* and the deictic *eso,* and in the case of pseudo-clefts to the lexicalised chunks *es que* "it is that" and *lo que pasa es que / lo que pasó era que* "what happens is that / what happened was that".

This analysis proposes a novel approach to explaining regularities emerging in bilingual speech. As outlined in the Introduction, extra-sentential switching has been frequently explained on a pragmatic basis, under the assumption that linguistic items expressing pragmatic and discourse-related meanings are more frequently switched from one language into the other *regardless* of their syntactic complexity and lexical specifity. In this study, I have adopted a different perspective, arguing that alternational switching of syntactically complex items is heavily constrained by the nature of the involved constructions, as determined by their mental representation. Specifically, those items that are available as lexical units in monolingual speech represent the best candidates for switching. By adopting as a major criterion the distinction between schematic and lexically specific instances of a construction, we do not merely find that code-mixing involves to the same extent simple and complex units, but we also can predict which instances of a construction will be more likely to be selected in bilingual patterns. In fact, even in typically schematic constructions, like dislocations and pseudo-clefts, switching was found to involve only a few lexical types which share the independently established feature that they are highly frequent in monolingual corpora.

Interestingly, this tendency does not contradict previous findings based on functional-pragmatic distinctions between clause-internal and clause-external functions, but goes hand in hand with them. In both case studies, the elements that were recognised as lexical chunks also convey discursive and pragmatic meaning, whereas schematic instances of the constructions were more typically associated with local clause-oriented functions.

33

To conclude, these findings also suggest a more general conclusion about language contact itself. It was shown that lexically-filled slots of particular constructions may have the status of lexical chunks in monolingual speech; the two case studies presented in the previous section demonstrate that these items are more available in code-mixing than schematic instantiations of the same constructions. Now, at a more general level, this could mean that code-mixing is cognitively easier to perform than interference, or pattern-replication in the sense of Matras and Sakel (2007). In fact, the first involves the use of surface forms whose mental representations are already stored in the lexicon/grammar, while the latter requires to pick up an empty (schematic) template in language A and fill it with surface material from language B (or both). Further evidence in favour of this view could be found by adopting the same perspective in the analysis of similar constructions, characterised by some degree of internal complexity and by the possibility to have both lexically-filled and lexically-open instantiations.

6   References

Agar Marco, Rocío. 2014. Pseudo-cleft senteces. Italian-Spanish in contrast. In Anna-Maria De Cesare (ed.) *Frequency, forms and functions of cleft constructions in Romance and Germanic,* 177-216. Berlin, De Gruyter.

Auer, Peter. 1999. From Code-switching via Language Mixing to Fused Lects: Toward a Dynamic Typology of Bilingual Speech. *International Journal of Bilingualism* 3 (4): 309-332.

Auer, Peter. 2005a. "Europe's Sociolinguistic Unity, or: A Typology of European Dialect / Standard Constellations". In Nicole Delbecque, Johan van der Auwera, Dirk Geeraerts (eds.), *Perspectives on Variation* (Trends in Linguistics; 163), 7-42. Berlin: de Gruyter.

Auer, Peter. 2005*b*. Projection in interaction and projection in grammar. *Text* 25: 7–36.

Auer, Peter. 2014. Language mixing and language fusion: when bilingual talk becomes monolingual. In Juliane Besters-Dilger, Cynthia Dermarkar, Stefan Pfänder, Achim Rabus (eds.), *Congruence in Contact-Induced Language Change*, 294-336. Berlin: de Gruyter.

Author. 2015. *Il discorso bilingue a Gibilterra. L'emergere di schemi regolari nella commutazione di codice*. PhD Thesis, Pavia University.

Author. 2016. The role of extra-clausal constituents in bilingual speech. The emerging of regular patterns in a bilingual corpus. In Gunther Kaltenböck, Evelien Keizer, Arne Lohmann (eds.) *Outside the clause: form and function of extra-clausal constituents.* Studies in Language Companion Series 178, 273-302. Amsterdam, Benjamins.

Backus, Ad. 2003. Units in codeswitching: evidence for multimorphemic elements in the lexicon. *Linguistics*, 41: 83-132.

Brinton, Laurel. 2014. The recent rise of "saying" ECCs: *(I'm) just saying*, *All I'm saying (is)*, and *What I'm saying (is)*. Paper presented at the *Outside the Clause* Workshop. University of Wien, 4-5 July 2014.

Bybee, Joan. 2002. Sequentiality as the basis of constituent structure. In Talmy Givón and Bertram Malle (eds.), *The evolution of language from pre-language*, 109–32. Amsterdam /Philadelphia: John Benjamins.

Bybee, Joan. 2006. From usage to grammar: the mind's response to repetition. *Language* 82: 711–733.

Bybee, Joan. 2010. *Language, Usage and Cognition.* Cambridge: Cambridge University Press.

CA: Stanford University Press.

Bybee, Joan and David Eddington. 2006. A usage-based approach to Spanish verbs of 'becoming'. *Language* 82: 323–55.

Cinque, Guglielmo. 1990. *Types of A' Dependencies*, Cambridge (MA), MIT Press.

Collins, Peter C. 1991. *Cleft and Pseudo-cleft Constructions in English*. London: Routledge.

Croft, William and D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Curnow, Timothy Jowan, Catherine E. Travis. 2004. The emphatic *es* construction of Colombian Spanish. In Christo Moskovsky (ed.), *Proceedings of the 2003 Conference of the Australian Linguistic Society*. Downloadable at http://www.als.asn.au/proceedings/als2003.html. Last access: 1st July 2016.

Diessel, Holger. 2016. "Frequency and Lexical Specificity. A Critical Review." In *Experience Counts: Frequency Effects in Language*, edited by Heike Behrens and Stefan Pfänder. Linguae & Litterae, Freiburg Institute for Advanced Studies. Berlin: de Gruyter.

Dik, Simon. 1997. *The theory of Functional Grammar. Vol 2, Complex and derived constructions*. Dordrecht, Foris.

Downing, Angela. 1997. Discourse and pragmatic function of the Theme constituent in spoken European Spanish. In John Connolly, Roel Vismans, Chris Butler, Richard Gatward (eds.) *Discourse and Pragmatics in Functional Grammar*, 137-161. Berlin: de Gruyter.

Duranti, Alessandro and Elinor Ochs. 1979. Left-dislocation in Italian conversation. In Talmy Givón (ed.) *Syntax and semantics. Discourse and syntax*, vol. 12 , 377–416. New York: Academic Press.

Ellis, Nick C. 1996. Sequencing in SLA: phonological memory, chunking and points of order. *Studies in Second Language Acquisition* 18: 91–126.

Fernández Soriano, Olga. 2008. On the nature of covert operations. The case of focus in Spanish pseudoclefts. In Enoch O. Aboh, Elisabeth van der Linden, Josep Quer and Petra Sleeman (eds.) *Romance Languages and Linguistic Theory. Selected papers from 'Going Romance' Amsterdam 2007*, 87-108. Amsterdam: Benjamins.

Fillmore, Charles, Paul Kay and Catherine O'Connor (1988). Regularity and Idiomaticity in Grammatical Constructions: The Case of let alone. *Language* 64: 501–38.

Fishman, Joshua. 1967. Bilingualism with and without diglossia; diglossia with and without bilingualism". *Journal of Social Issues* 23 (2): 29-38.

Geluykens, Ronald. 1992. *From discourse process to grammatical construction. On left-dislocation in English*. Amsterdam: Benjamins.

Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Gregory, Michelle L. and Laura A. Michaelis. 2001. Topicalization and Left-Dislocation: A Functional Opposition Revisited. *Journal of Pragmatics* 33: 1665-1706

Günthner, Susanne and Paul Hopper. 2010. Zeitlichkein & sprachliche Strukturen: Pseudoclefts im Englishen und Deutschen. *Gesprächsforschung. Online Zeitschrift zur verbale Interaktion: 1-28.* Downloadable at http://www.gespraechsforschung-online.de/heft2010/heft2010.html. Last access: 1st July 2016.

Günthner, Susanne. 2011. "N be that-constructions in Everyday German Conversation. A Reanalysis of 'die Sache ist/das Ding ist' ('the thing is')-Clauses as Projector Phrases." In Ritva

Laury and Ryoko Suzuki *Subordination in Conversation*, 11–36. Amsterdam: John Benjamins. DOI: 10.1075/slsi.24.03gun

Hakimov, Nikolay. 2016*a.* Explaining variation in plural marking of single German noun insertions in Russian sentences. In Heike Beherens and Stefan Pfänder (eds.) *Experience counts: Frequency effects in language,* Linguae et litterae 54. Berlin, de Gruyter: 21-60.

Hakimov, Nikolay. 2016*b*. Effects of frequency and word repetition on switch-placement: evidence from Russian-German code-mixing. In Justyna Robinson and Monika Reif (eds.) *Cognitive perspectives on bilingualism,* Trends in Applied Linguistics 17. Berlin, de Gruyter: 91-125.

Haiman, John. 1994. Ritualization and the development of language. In William Pagliuca (ed.), *Perspectives on grammaticalization*, 3–28. Amsterdam/Philadelphia: John Benjamins.

Heine, Bernd, Gunther Kaltenböck, Tania Kuteva, and Haiping Long. 2013. An outline of discourse grammar. In Bischoff, Shannon and Carmen Jany (eds.), *Functional Approaches to Language,* 175-233. Berlin: de Gruyter.

Heine, Bernd. 2013. On discourse markers. Grammaticalization, pragmaticalization, or something else? In *Linguistics* 51: 1205-1247.

Hidalgo Downing, Raquel and Laura Hidalgo Downing. 2007. Metadiscursive and interpersonal values of pronominal topics in spoken Spanish. In: Christopher S. Butler, Raquel Hidalgo Downing and Julia Lavid (eds.) *Functional perspectives on grammar and discourse. In honour of Angela Downing*, 327-348. Amsterdam/Philadelphia: Benjamins.

Hidalgo, Raquel. 2000. Establishing Topic in conversation: a contrastive study of left-dislocation in English and Spanish. In Angela Downing, Jesus Moya, José I. Albentosa (eds.), *Talk and text. Studies on spoken and written discourse*, 31-50. Cuenca: Ediciones de la Universidad de Castilla-La Mancha.

Higgins, Roger. 1979. *The Pseudo-cleft construction in English.* New York: Garland Publishing.

Hilpert, Martin. 2014. *Construction Grammar and its Application to English*. Edinburgh: Edinburgh University Press.

Holzinger, Daniel. 1993. *Das Romanes: Grammatik und Diskursanalyse der Sprache der Sinte*. Innsbruck: Institut für Sprachwissenschaft.

Hopper, Paul J. and Sandra A. Thompson. 2008. Projectability and clause combining in interaction. In Ritva Laury (ed.), *Crosslinguistic Studies of Clause Combining: The Multifunctionality of Conjunctions*, 99–123. Amsterdam: Benjamins.

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

Kaltenböck, Gunther, Bernd Heine, Tanja Kuteva. 2011. On thetical grammar. *Studies in language* 35: 852-897.

Kaltenböck, Gunther, Evelien Keizer and Arne Lohmann. 2016. *Outside the Clause. Form and Function of Extra-clausal constituents.* Amsterdam: Benjamins.

Kellerman, Anja. 1996. When Gibraltarians speak, we're quite unique! Constructing Gibraltarian Identity with the Help of English, Spanish and their respective local varieties. In: Sandi Michele De Olivera (ed.) *The linguistic construction of social and personal identity. First international conference of sociolinguistics in Portugal*, 73-78. Evora: Cidehus.

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

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

Lambrecht, Knud. 1994. *Information Structure and Sentence Form. Topic, Focus and the Mental Representation of Discourse Referents*. Cambridge: Cambridge University Press.

Lambrecht, Knud. 2001*a*. Dislocation. In: Martin Haspelmath (ed.) *Typologie des langues et les universaux linguistiques: Manuel international*. Cap. 80, 1056–1078. Berlin: De Gruyter.

Lambrecht, Knud. 2001*b*. A framework for the analysis of cleft constructions. *Linguistics*, 39(3):463-516.

Langacker, Ronald. 1987. *Foundations of cognitive grammar: theoretical prerequisites*. Vol.I. Stanford,

Levey, David. 2008*a*. The changing face of Gibraltarian English. Th-fronting on the Rock. In: Ewa Waniek-Klimczak (ed.), *Issues in accents of English*, 91-102. Newcastle: Cambridge Scholar Press.

Levey, David. 2008*b*. *Language change and variation in Gibraltar*, Amsterdam: Benjamins.

Levey, David. 2015. Gibraltar English. In Williams, Jeffrey P., Edgar Schneider, Peter Trudgill and Daniel Schreier (eds.), *Further Studies in Lesser-known varieties of English*, 51-69. Cambridge: Cambridge University Press.

Marcos Marín, Francisco. 1992. *Corpus oral de referencia de español contemporáneo*. Madrid:

Matras, Yaron and Jeanette Sakel. 2007. *Grammatical borrowing in cross-linguistic perspective.* Berlin: De Gruyter.

Matras, Yaron. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36(2): 281-331.

Matras, Yaron. 2000. Fusion and the cognitive basis for bilingual discourse markers. In *International Journal of Bilingualism* 4(4): 505-528.

Matras, Yaron. 2009. *Language Contact*. Cambridge: Cambridge University Press.

Moyer, Melissa. 1998. Bilingual conversation strategies in Gibraltar. In Peter Auer (ed.) *Codeswitching in conversation. Language, interaction and identity*, 215-234. New York, Routledge.

O'Shannessy, Carmel. 2011. Young Children's Social Meaning-Making in a New Mixed Language. In Ute Eickelkamp (ed.) *Growing Up in Central Australia - Indigenous Experiences of Childhood, Youth and Transformations*. New York: Berghahn Books pp 131-155

Pannunzi, Alessandro. 2016. Pseudo-cleft sentences in spoken Italian. Syntax, semantics, information structure and a comparison with Spanish. In Anna-Maria De Cesare and Davide Garassino (eds.) *Current issues in Italian, Romance and Germanic non canonical word-orders*, 227-250. Frankfurt am Main: Peter Lang.

Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish y termino en español. *Linguistics* 18: 581-618.

Poplack, Shana and Marjory Meechan. 1995. Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In Pieter Muysken and Leslie Milroy (eds.) *One speaker, two languages.* Cambridge, University Press: 199-232.

Prince, Ellen. 1981. Toward a taxonomy of given-new information. In Peter Cole (ed.) *Radical Pragmatics*, 223-255. Amsterdam: Benjamins.

Prince, Ellen. 1984. Topicalization and left-dislocation: A functional analysis. In Sheila White, Virginia Teller (eds.), *Discourses in reading and linguistics. Annals of the New York Academy of Sciences* 433, 213-225.

Reig Alamillo, Asela. 2011. "The pragmatic meaning of the Spanish construction *lo que pasa es que*". *Journal of Pragmatics 43(5),* 1435-1450.

Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". In Liliane Haegeman (ed.), *Elements of Gramma*r, 281-337. Kluwer, Dordrecht.

Sloetjes, Han, & Peter Wittenburg. 2008. "Annotation by category – ELAN and ISO DCR". In: *Proceedings of the 6th International Conference on Language Resources and Evaluation* (LREC 2008).

Traugott , Elizabeth Closs and Graeme Trousdale. 2013. *Constructionalization and Constructional Changes.* Oxford: Oxford University Press.

Traugott, Elizabeth. 2008. "All that he endeavoured to prove was …": On the emergence of grammatical constructions in dialogual and dialogic contexts. In Robin Cooper and Ruth Kempson, (eds.), *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*, 143-177. London: Kings College Publications.

Travis, Catherine E. 2005. *Discourse Markers in Colombian Spanish: A Study in Polysemy*. Berlin, De Gruyter.

Treffers-Daller, Jeanine. 1994. *Mixing two Languages: French-Dutch Contact in a Comparative Perspective.* Berlin: De Gruyter.

Universidad Autónoma de Madrid.

Ward, Gregory, Betty Birner and Rodney Huddleston. 2002. Information packaging. In Rodney Huddleston and Geoffrey Pullum (eds.) *The Cambridge Gramar of the English Language,* 1363-1447. Cambridge: Cambridge University Press.

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

Table 1

		Pattern				
		e-e	s-s	e-s	s-e	total
nominal ld		14	5	2	14	35
pronominal ld		2	23	2	19	46
total		16	28	4	33	81

Table 2

Pattern

Dislocand	Monolingual	Bilingual	Total
Yo	6	12	18
Eso	11	5	16
people	4	0	4
ello'	2	1	3
nosotro	2	1	3
name	3	0	3
la gente	0	2	2
la' persona'	0	2	2
other spa	7	10	17
other eng	9	4	13
Total	44	37	81

Table 3

Pattern

Argument	?	s-s	e-e	e-s	s-e	total
Clausal	2	51	25	3	32	113
Nominal	0	26	3	0	1	30
Total	2	77	28	3	33	143

Table 4

<p style="text-align:center">Pattern</p>

Lexical types	?	Monolingual	Bilingual	Total
es que	2	23	16	41
pasar	0	16	8	24
tener	0	1	3	4
thing is	0	17	2	19
hacer	0	6	2	8
do	0	1	1	2
decir	0	6	0	6
afectar	0	2	0	2
happen	0	2	0	2
querer	0	2	0	2
ser	0	2	0	2
other	0	16	1	17
Total	2	102	33	137

SG	Singular
PL	Plural
AUX	Auxiliary
COMP	Complementiser
COND	Conditional
DM	Discourse marker
DOM	Differential object marking
F	Feminine

INF	Infinite
M	Masculine
NEG	Negation
OBJ	Object
PRS	Present
PST	Past
REFL	Reflexive
REL	Relative
SBJ	Subject
1	1st person
2	2nd person
3	3rd person