Language Mixing in Northern and Western Belize: A Comparative Variationist Approach

Nicté Fuller Medina

A thesis submitted to the Faculty of Graduate and
Postdoctoral Studies in partial fulfillment of the requirements
for the PhD degree in linguistics.

Department of Linguistics
Faculty of Arts
University of Ottawa

ABSTRACT

This thesis examines the bilingual discourse of a cohort of Belizean Spanish speakers who engage in robust language mixing between Spanish, English and to a lesser extent Belize Kriol, an English-lexified creole. The speakers selected for the current study have been identified as the "highest language mixers" in a corpus of 51 interviews conducted in northern and western Belize, areas which have been classified as two distinct dialect regions (Cardona Ramírez 2010; Hagerty 1979). While an abundance of research exists on Spanish-English bilingualism in the U.S. (e.g. Torres Cacoullos and Travis 2014; Silva Corvalán 1994; Roca and Lipksi 1993) there is less research on non-U.S. varieties of Spanish in contact with English, in particular, Belizean varieties of Spanish. Thus, by appealing to the comparative variationist framework (Poplack and Tagliamonte 2001), the major aims of the study are: (i) to describe the major patterns of use among those speakers of Belizean varieties of Spanish who engage in language mixing and, (ii) to determine the status of the single and multiword English-origin fragments which comprise the majority of non-native material in Spanish discourse. In determining the status of the English-origin material with regard to borrowing and code-switching, not only are the specific linguistic mechanisms used by these speakers elucidated, but insights are gained as to whether code-switching and borrowing are distinct linguistic phenomena.

Diagnostics of subject position and gender and number agreement on English-origin nouns and verbal morphology and variable clitic placement for English-origin verbs revealed both these categories to pattern with Spanish suggesting that they are borrowings. The remaining one-third of the data, comprised of multiword fragments, consisted primarily of intrasentential and intersentential code-switching and a large category of multiword fragments which initially appeared to be neither code-switches nor borrowings. A comparative quantitative analysis revealed these items to be integrated into Spanish suggesting that they may be treated as single units of meaning.

Results, for the most part, are consonant with the literature on bilingual speech. Data consists mainly of lexical borrowing (Thomason 2001; Pfaff 1979; Berk-Seligson 1986), specifically nouns, the most borrowed category cross-linguistically (Muysken 2000; Poplack et al. 1988). Speakers engage in "skilled" or equivalence intrasentential code-switching consistent with other Spanish-English data (Poplack 1980). In addition, only those speakers who reported being equally dominant in the respective languages exhibited robust intrasentential code-switching, thus, concurring with the prevailing assertion that code-switching is the domain of fluent bilinguals (Bullock and Toribio 2012; Lipski 1985; Poplack 1980).

Some distinctive features of the language mixing employed by these speakers include the frequent and productive use of bilingual compound verbs (BCV) and a near categorical preference for BCVs as the mechanism for borrowing English-origin verbs. With regard to Spanish determiner marking on English-origin nouns, the masculine default is used almost exclusively, unlike the variability reported elsewhere (Dubord 2004; Smead 2000). By analyzing data from both dialect regions of Belize, this study provides insight both into the global picture of language mixing practices in Belize as well as regional patterns insofar as they are instantiated by the cohort analyzed.

In memoriam

Grace Meighan 1913-2015
Ida Fuller 1933-2015
Rosalía Medina Gil 1940-2014

Table of Contents

List of Tables

LIST OF GLOSSING ABBREVIATIONS

1	first person
2	second person
3	third person
DET	determiner
FEM	feminine
MASC	masculine
SG	singular
PL	plural
PRES	present
IMP	imperfect
PRET	preterit
FUT	future
INF	infinitive
PTCP	participle
CL	clitic
POSS	possessive
COMPL	completive marker
DOM	differential object marker

CHAPTER 1 INTRODUCTION

I have come to know something simple. Each sentence realised or
dreamed jumps like a pulse with history and takes a side.
- Dionne Brand*, No Language is Neutral*

This thesis has as its focus language as it used in Spanish-speaking regions in Belize. Drawing on

the comparative variationist framework (Poplack and Tagliamonte 2001), this thesis analyzes the

mixed speech of a cohort of 12 speakers who engage in robust language mixing between Belizean

varieties of Spanish, English, and to a lesser extent Belize Kriol, an English-lexified Creole. The

main objectives of the current study are: (i) to describe the major patterns of use employed by

speakers of Belizean varieties of Spanish who engage in language mixing and (ii) to determine the

status of the single and multiword English-origin fragments in the Spanish discourse of these

speakers. It is well known that bilingual speakers engage in both lexical borrowing (the adaptation

of lexical items from one language into another) and various types of code-switching (the

alternation between two or more languages in the same utterance). English-origin items in Spanish

discourse, for example, may hold the status of borrowing or code-switch (Poplack 2015, 2004;

Thomason 2001). Therefore, this study aims to identify these surface manifestations of language

mixing, a necessary task, though, not always a clear-cut one.

In fact, there is ongoing disagreement as to whether or not code-switches and borrowings

can be reliably distinguished with respect to single non-native items in mixed discourse (Poplack

2015; Poplack and Dion 2012; Stammers and Deuchar 2012; Winford 2009; Myers-Scotton

2002,1993), though a number of variationist studies advance the position that these phenomena

can, and should be, distinguished (Poplack 2015; Poplack and Dion 2012; Poplack and Meechan

1998a). Thus, as a consequence of the analysis carried out on all the English-origin items in the

current data, this thesis not only provides an account of language mixing by a cohort of Belizean

Spanish speakers, but also responds to the unresolved question regarding distinctions between borrowing and code-switching.

Following Pfaff (1979), the term language mixing is used, here, as a cover term referring to any instance where more than one language is used in discourse by the same speaker. Accordingly, language mixing encompasses both borrowing and various types of code-switching, such as switching within a clause (intrasentential switching) or switching between independent clauses (intersentential switching). For emphasis and distinction from the various definitions of code-switching in the literature, in the current study, code-switching refers to instances where speakers alternate between two or more separate grammatical systems in the same utterance; these alternations are generally multiword fragments which exhibit internal consistency with the source language as evidenced, for example, by word order, the use of function words, and so forth (Sankoff et al. 1990; Poplack 1980). Borrowing, on the other hand, refers to instances where a single lexical item from a donor language is integrated into the grammar of a recipient language (Winford 2003:103; Poplack 1980; Haugen 1956). In some cases, multiword fragments may also function as lexical borrowings when they are treated as a single unit of meaning (e.g. compounds such as *credit card*) and, therefore, may also be integrated into the recipient language. Only the grammar of the recipient language is considered to be operative lexical items are borrowed. When code-switching, on the other hand, speakers are considered to alternate between distinct grammatical systems and, therefore, the respective grammars are considered to be operative since. Since a determination as to which language(s) might be operative in mixed discourse and the status of the surface manifestations of language mixing are made via quantitative comparative analysis, terms such as *fragments*, *items*, or *material* are used to refer to the data until the analysis reveals what their status might be with respect to code-switching or borrowing. Furthermore, these data are discussed in terms of *English origin* due to the overlap between the Belize Kriol and

English lexicons. At the level of single items (see *advantage* and *mix* in (2) and (3) below), for example, it is virtually impossible to determine, based on the surface form, if these items are Belizean English or Belize Kriol lexical items, though most non-Spanish items found in the corpus are etymologically English. Thus, the term *English-origin* refers to the etymological source. Where Kriol multiword fragments or lexical items were identified, these are noted throughout the thesis.

Data is drawn from both northern and western Belize as these regions have been identified as distinct dialect regions in Belize (Cardona Ramírez 2010; Hagerty 1979). Standard sociolinguistic interviews were carried out in these regions, resulting in a corpus of 51 interviews compiled in the Belizean Spanish, English, and Kriol Corpus (BSEK). All speakers were native speakers of Northern Belizean Spanish or Western Belizean Spanish and grew up in a Spanish-speaking home. Interviews from twelve speakers, identified as the highest language mixers, comprise the data analyzed in this thesis. Note that the focus of this thesis on language mixing should not be taken as a general reflection of the nature Spanish as spoken in Belize. As discussed in data collection, many speakers engage in very little language mixing and this thesis has simply targeted that cohort that does make regular use of mixed discourse.

The major tenets of the variationist method as applied to bilingual data (Poplack 1993) guide this thesis and consequently will be seen to underlie each chapter's method and analysis. First, the object of study must be clearly circumscribed so that all the relevant data is included in analyses responding to research questions, and by the same token, *only* data relevant to the research question is considered. Accomplishing this task requires setting appropriate benchmarks against which data can be compared or measured. A linguistically principled *a priori* definition, such as the one given for code-switching above, ensures that only those data corresponding to this independently motivated benchmark are included in the relevant analyses. More importantly, in the variationist approach, the communities' own vernaculars are used as benchmarks to tap into the

underlying grammars of the speakers so as to situate non-native material with respect to the recipient language. Thus, in the current study, prescriptive norms and theoretical analyses of grammatical systems of the respective languages are eschewed in favor of local varieties of Spanish as spoken by the speakers in the sample analyzed. In this way, the inherent variability of the local varieties is taken into account. This practice is derived from a third principle, and perhaps, the most fundamental one, which holds that language is inherently variable and, further, that this variability is systematic (Labov 1972). An accountable quantitative analysis is then applied to the data. In the case of bilingual data, this means that all relevant all data is presented, analyzed, and discussed, and any exclusion of data is based on a principled rationale and also reported. Additionally, by taking into account, and even making use of, the inherent variability in the local vernaculars, the analysis is also accountable to the variability in the data (Torres Cacoullos and Aaron 2003). Finally, the variationist method aims to uncover the communities' regular patterns of use, i.e., those patterns which reflect natural exchanges among speakers and which represent the majority of data (Poplack and Meechan 1998:262; Poplack 1993).

The first objective aims to describe language mixing patterns in varieties of Spanish that have gone largely unnoticed in Hispanic sociolinguistics and dialectology. In fact, language in Belize is highly understudied in general, and Spanish is no exception. Decades ago, Hagerty (1979:2) observed that the "meager documentation available on Belizean Spanish demonstrates a void in Hispanic dialectology which must be filled," and this remains true today as evidenced by only a handful of studies on Spanish in Belize (Fuller Medina 2015, 2005a,b; Balam et al. 2014; Balam 2013; Cardona Ramírez 2010; Quilis 1990; Hagerty 1996, 1979; Le Page and Tabouret Keller 1985). As an additional consequence, there are also gaps in the field of contact linguistics with regard to Spanish-English and Spanish-Kriol contact in the region despite the historical Anglophone and Creolophone presence in Central America (Holm 1983). Therefore, this thesis

aims to contribute to filling these gaps in the literature by documenting some features of varieties of Spanish in Belize and a detailed empirical description of language mixing as an outcome of language contact. Considering Himmelman's (1998:166) definition of documentation which is "to provide a comprehensive record of the linguistic practices characteristic of a given speech community," that is to say, "the observable linguistic behavior, manifest in everyday interaction between members of the speech community," the current study fulfills, in part, the work of documenting Spanish in Belize by presenting details of one aspect: the mixed speech of bilinguals and trilinguals.

Spanish-English contact has been widely studied, in particular, in the U.S. (Travis and Torres Cacoullos 2013; Lipski 2008, 1985; Mendieta 1999; Roca and Jensen 1996; Silva Corvalán 1994; Roca and Lipski 1993; Amastae and Elias Olivares 1984; Poplack 1980; Pfaff 1979; McClure 1977). The literature has examined diverse communities where different varieties of Spanish are in contact with English in the U.S. and where the length and type of contact also varies. Consequently, it is important to highlight that the outcomes of Spanish-English contact cannot be expected to be monolithic and that communities must be assessed on a case-by-case basis, particularly since extralinguistic factors interact with linguistic ones to shape these outcomes (Poplack 2004, 1993). Thus, despite the superficial similarities between contact Spanish in Belize and in the U.S., socio-historical circumstances differ between the two settings, warranting a detailed study of language mixing in the understudied context of Belize. This can consequently inform our understanding of Spanish-English mixing in various environments and, more generally, the linguistic mechanisms employed by bilingual speakers.

The most obvious similarity between the two contexts is the dominance and concomitant official status of English in the U.S. and Belize. Long-term contact with English is also a shared feature with some communities such as New Mexican Spanish speakers, as both the northern and

western varieties of Spanish in Belize were established over 150 years ago (Hagerty 1979). Additionally, Northern Belizean Spanish, like New Mexican Spanish, has a variety of Mexican Spanish as its input variety. One of the major differences between the two contact situations is the status of Spanish. Spanish in Belize is a de facto second official language, even if not formally recognized as such. In fact, one of the recommendations to emerge from the most recent national census is for the Government of Belize to consider making Spanish its second official language (SIB 2013:56). Furthermore, while English is the official language of Belize (as in the case of the U.S.), in fact, it is reserved for formal contexts such as education, whereas Belize Kriol is the lingua franca. Spanish is not considered a minority or immigrant language in the same sense as in the U.S.[1] even if it is stigmatized due to non-standard features and real or perceived language mixing. There is, however, no systematic impetus in Belize to impose English on Spanish-speaking communities[2]. One final difference worth noting between the U.S. and Belizean contexts, though there are many others, is that Spanish is also in sustained contact with Belize Kriol, an English-lexified Creole, a feature absent in the U.S. contexts. Ultimately, given that Belize provides a natural laboratory for studying Spanish in contact with English, Kriol, and Mayan languages as well as dialect contact amongst the different varieties of Spanish, it is essential to establish the features of the local varieties of Spanish and the outcomes of language contact, most notably language mixing.

This thesis is a first step in tackling this tremendous task and in closing current gaps in Hispanic linguistics and contact linguistics with regard to Spanish in Belize. Findings of the current study are contextualized in relation to the relevant literature on Spanish-English contact with a

---

[1] While Spanish in the U.S. cannot be described as an immigrant language since some varieties have been present in the U.S. for 400 years (Travis and Torres Cacoullos 2013), it is often perceived in this way.
[2] While the government and education system all operate in English, there are no English-only movements as in the U.S. and, in fact, there is evidence of commitment on the part of the government to preserve languages in Belize (NICH 2014).

focus on community patterns of language mixing. The possible role of Kriol in the linguistic interplay of language mixing is highlighted where relevant and where the data permits, even if not assessed via large-scale quantitative analyses. This sets the stage for further research with respect to this language in contact with Spanish.

The second objective of this thesis is achieved via a detailed study of English-origin material to determine the outcomes of language mixing in the respective communities of Belizean Spanish speakers. Community patterns of mixing may consist of borrowing, skilled intrasentential code-switching (unmarked by hesitations, false starts, or other flagging), flagged switching (marked by false starts, metalinguistic commentary, and so forth) (Poplack 2004, 1987). Thus, the underlying motivation in identifying the status of English-origin material is to ultimately elucidate he communities' major patterns of language mixing.

Multiword fragments which show internal consistency with the source language, as in (1), for example, are generally agreed to be code-switches and are easily identified. Examples such as *advantage* in (2), however, are ambiguous as they may be borrowings as evidenced by possible integration into Spanish via the use of the Spanish determiner. Alternatively, such items may be code-switches since *advantage* occurs at a site of equivalence, or area where the grammars coincide, for Spanish and English, specifically, word order.

1.   Habíamos dicho que **we want to downsize. So we are downsizing**.
        'We had said that we want to downsize. So we are downsizing.'
                        (BSEK/45NJ/45NJ_43NJ_3/01:33:13/)
2.   yo creo que ese es un **advantage** que tenemos.
        'I think that that is an advantage that we have.'
                        (BSEK/11WJ/11WJ_12WJ_13WJ_2/00:09:02)

3.   como allí en Corozal lo que hacen **mix** es el Creole y el English
        'like, there in Corozal, what they mix is Creole and English'
                        (BSEK/12WJ/11WJ_12WJ_13WJ/03:27.3)

Similarly, English-origin items such as *mix* in the construction *hacen mix* in (3), known in the literature as bilingual compound verbs[3] (BCVs), exhibit no inflectional morphology directly on *mix*. Thus, this may be an example of a code-switch between *hacer* and *mix* or, alternatively, *mix* may be a borrowed verb (or even noun) fully integrated into Spanish with no alternation between the Spanish and English grammatical systems. Other proposals consider these items to be neither code-switches nor borrowings and instead consider them to evidence a special bilingual grammar or third grammar (Wilson 2013; Muysken 2000; Romaine 1989, 1986) or to reflect processes of creolization (Balam et al. 2014; Gardner-Chloros 2009; Pfaff 1979). The current study addresses these disparate classifications in the literature with a detailed study of BCVs.

The ambiguity of the items in (2) and (3) derives mainly from the fact that both code-switches and borrowings may have the same surface form even if produced by two different underlying mechanisms (Poplack and Dion 2012). Consonant with the comparative variationist method, this study addresses this quandary by exploiting conflict sites (Poplack & Meechan 1998), or sites where the two respective grammars behave differently in terms of rates and/or linguistic conditioning. Single-items found at these conflict sites, and whose status is ambiguous with regard to code-switching and borrowing, are analyzed quantitatively. The corresponding categories in monolingual English and Spanish are also analyzed quantitatively the three independent analysis are then compared in order to determine which grammar. This is known as the comparative variationist method (Poplack and Tagliamonte 2001). In this way, the patterning of this ambiguous data vis-à-vis the conflict site can be assessed against the behavior of the corresponding categories in the monolingual grammars to determine the grammatical provenance of the ambiguous items. If the ambiguous items adhere to the grammatical patterns of the source-

---

[3] These are also known as *do*-verb constructions or simply bilingual verbs. Following Edwards and Gardner Chloros (2007), I adopt the term bilingual compound verb (BCV).

language vernacular, then such items are likely code-switches, but if they pattern with the recipient language, i.e., are fully integrated into the recipient language, then such evidence points to these items being borrowings. With regard to English-origin nouns, three conflict sites are identified and used as diagnostics: subject position, gender agreement, and number agreement. Verbal morphology and variable clitic position serve as diagnostics for assessing English-origin verbs. Thus, by employing the variationist method within a comparative approach, it becomes possible to identify the system membership of data that might would otherwise be ambiguous.

In addition to the ambiguity of the status of the English-origin item in BCVs, i.e., code-switch or borrowing, its categorial status (e.g. verb or noun) is also unclear. The English-origin item is variably considered an infinitive (Balam et al. 2014:245; Wilson 2013; Jenkins 2003), a bare stem (Pfaff 1979:300), a noun (Torres Cacoullos and Aaron 2003), or its categorial status remains ambiguous (Jenkins 2003). In order to accurately describe BCVs, and in keeping with the variationist approach of clearly defining or circumscribing the object of study, some methodological tools are suggested for the study of BCVs. A series of discourse-based tests are applied to distinguish non-verbal bilingual constructions with *hacer* from BCVs in order to include in the analysis only those constructions that are relevant to verbal borrowing and which constitute the major pattern of use. Following clear identification of the relevant data, the English-origin item can be assessed for system membership, and subsequently, its status with regard to code-switching and borrowing. This will then indicate whether speakers are borrowing verbs via *hacer* constructions, as seen above in (3) or if they are switching to English (code-switching) in order to use English-origin verbs in discourse.

Once the BCVs in this study have been clearly circumscribed and both the categorial status and system membership of the English-origin items has been determined vis-à-vis the application of tools of the variationist method, these very tools are brought to bear on disentangling

the question of potential aspectual restrictions on BCVs. While cross-linguistically some languages reserve a *do*-verb for with non-native verbs which are more active and a second verb for more stative verbs (Muysken 2000), it is unclear if both stative and dynamic English-origin verbs may occur with the *do*-verb *hacer* in BCVs as inconsistent results have been reported (Balam et al. 2014; Fuller Medina 2013, 2005a,b). Drawing on the resources of comparative variationist sociolinguistics, and following Walker (2010a), benchmarks are set for assessing the aspectual properties of the English-origin items in BCVs. Lexical aspect is operationalized by appealing to independent theoretically motivated criteria rather than relying on individual researcher intuitions or any presumptions regarding shared understanding of aspectual classes. In so doing, the hypothesis that BCVs are aspectually restricted at the level of lexical aspect is tested in order to further elucidate the manner in which BCVs operate in bilingual speech.

Special attention is paid to BCVs as in (3) above because it is an understudied structure in the Spanish-English literature and, to my knowledge, Spanish-English BCVs have not been previously analyzed from a variationist perspective in the Spanish-English literature with regard to the status of the English-origin items. BCVs have only recently become the focus of several studies on Spanish-English contact (Wilson and Dumont 2014; Balam et al. 2014; Wilson 2013; Fuller Medina 2013, 2010, 2005a,b). In addition, these constructions may be a distinctive feature of the language mixing found in Belize as they have been reported to be frequent (Balam at el. 2014; Fuller Medina 2005ab; Hagerty 1979). This is in contrast to large scale studies of other varieties of Spanish in contact with English where it is either not reported (e.g. Poplack 1980), or reported as a minor pattern (Jenkins 2003; Silva Corvalán 1994; Pfaff 1979), or at lower rates than those reported for Belize (Wilson and Dumont 2014).

The current study also assesses all multiword fragments in detail to determine if they are code-switches, and if so, what type of code-switches, or, alternatively, if they are being used as

single units of meaning (e.g. *credit card*, *professional makeup artist*) akin to loanwords. Once these data have been assessed and intrasentential code-switches identified, these are examined to determine if they are found at sites predicted by the equivalence constraint (EC) and the free morpheme constraint (FMC) (Poplack 1980) to favor for code-switching. According to the EC, code-switches will tend to occur at sites where the syntactic rules of the respective languages are not violated, i.e., sites where the grammars are homologous. Additionally, as per the FMC, there will be a tendency for code-switches to be avoided between bound morphemes (Poplack 1980). Both constraints are discussed in detail in Chapter 6.

There are two reasons for revisiting grammatical constraints. First, type of code-switching is an important element of community patterns of language mixing and the pattern for these varieties of Spanish in contact with English has not been identified. Second, while the EC and FMC received much criticism in the decades since they were first proposed, a number of methodological issues were insufficiently addressed in testing constraints and in the critiques put forward (Poplack and Dion 2012; Budzhak Jones 1998c; Eze 1997; Poplack and Meechan 1998; Poplack 1993; Sankoff et el. 1990). A major issue was the use of appropriate data, not only regarding the source of the data but also the type of data. For example, proposals based on judgments and intuitions were not always supported by actual bilingual use (Moro 2015; Poplack 1981), suggesting that there were some limitations to using this type of data. Likewise, even when production data was used, single non-native items with etymologically mixed morphologies were analyzed as code-switches without first empirically establishing whether they are relevant to testing constraints. Thus, the status of non-native items must be established before assessing the constraints governing language mixing. Those items that are borrowings and those that are code-switches must be distinguished and analyzed separately, which is precisely the approach used in the current study.

Subsequently, the EC and FMC are revisited in the mixed speech of Spanish speakers in Belize, taking into account all the developments in the field since these proposals were first made and making use of the methodological tools for distinguishing code-switching and borrowing (Poplack and Meechan 1998a). As a consequence, conceptually, multiword fragments and single non-native items are treated from the outset as two potentially distinct categories (Poplack and Dion 2012; Poplack and Meechan 1998a). Moreover, based on numerous quantitative studies (Poplack and Meechan 1998a; Sankoff et al. 1990; Poplack et al. 1988), which show borrowing to be an abrupt process independent of social and phonological integration, borrowing is now treated as a morphosyntactic process. Finally, despite the typological similarity between Spanish and English, which, according to the EC, would translate into many sites propitious to code-switching, these facts alone cannot predict which type of code-switching will constitute the community norm for Belizean Spanish speakers. Neither will these facts predict if the EC and FMC will be upheld in the data, pointing to the need for a detailed study of this new dataset. Finally, by isolating intrasentential switches, it also becomes possible to test the hypothesis that discourse markers, identified by some as code-switch pivots (Aaron 2004), introduce code-switches. In short, since the EC and FMC were first proposed, there have been advances in the field with regard to the conceptual and methodological tools for analyzing bilingual speech as well as the factors which play a role in the patterning of such phenomena.

The details of the research questions guiding this thesis and which respond to the aforementioned objectives are summarized below.

I. To describe the major patterns of language mixing employed by speakers of Northern Belizean Spanish and Western Belizean Spanish as revealed by quantitative analysis of English–origin fragments in mixed discourse.

II. To determine the status of single and multiword English-origin fragments in mixed discourse with respect to borrowing and code-switching.

In addition, to responding to the above questions, the comparative quantitative analysis of the data also enables additional questions to be addressed. In examining patterns of language mixing, for example, extralinguistic factors, such as language dominance and bilingual ability, which are reported to shape language mixing, can be assessed. The categorial status of the English item in BCVs is established and possible aspectual restrictions are also examined by appealing to the variationist approach.

The thesis is organized as follows: Chapter 2 presents an overview of the main issues in the literature regarding borrowing and code-switching as well as the most current accounts of grammatical constraints on code-switching. Chapter 3 provides an overview of Spanish in Belize and outlines the variationist method, which is employed in this thesis. Also detailed in this chapter are the special challenges posed by bilingual data, details of data collection and selection of participants, the linguistic profiles of the participants whose interviews are analyzed in the current study, and finally, the general distribution of the data. Chapters 4 and 5 assess the status of single English-origin verbs and nouns, respectively. Chapter 4 additionally presents methodological tools for clearly demarcating BCVs as well as for operationalizing lexical aspect. Chapter 6 has as its focus the analysis of multiword English-origin fragments and community patterns of language mixing. The different types of code-switching found in the data are described, and the grammaticality of code-switching as per the EC and FMC is also assessed as is the role, if any, of discourse markers in introducing switches. Finally, the patterns of language mixing in the two dialect regions are presented along with a discussion of bilingual proficiency and language dominance, both factors implicated in the formation of these patterns. Summarized in Chapter 7, the final chapter, are the conclusions as well as details of the contributions of this study.

CHAPTER 2 REVIEW OF THE LITERATURE

2.1 Introduction

Borrowing and code-switching are two of the most frequently observed manifestations of language mixing, however, disagreement remains in the field with respect to distinguishing the two phenomena and, as a consequence, which linguistic phenomena should be counted as code-switches and which should be counted as borrowings is also unclear. Moreover, identifying universal grammatical constraints on intrasentential code-switching continues to be elusive (Poplack and Dion 2012). Complicating matters further is that there is little consistency in the literature regarding the terms used to refer to surface manifestations of mixed speech (Moro 2015; Poplack and Dion 2012; Winford 2009:280; Muysken 2000). Code-mixing has been used to refer to any type of alternation between two languages (Muysken 2000), while code-switching has been used variably to refer to non-native multiword fragments (Poplack 1980; Pfaff 1979; Lipski 1985) or to both multiword fragments and single words (Myers-Scotton 2002; Treffers-Daller 2001:82). Additionally, terms such as classic code-switching and composite code-switching are deemed to refer to distinct manifestations of language mixing (Myers Scotton 2002). This chapter provides an overview of the points of contention in the language mixing literature, namely the question of borrowing and code-switching being distinct phenomena. Also discussed is the role of linguistic and social factors in the patterns that can be observed in diverse communities.

Patterns of language mixing are addressed first, followed by an outline of the main points of disagreement relating to distinguishing borrowing and code-switching. Some current proposals which attempt to account for the rule-governed behavior of code-switching are discussed briefly. Special attention is paid to variationist approaches to bilingual data which set forth the equivalence constraint, the free morpheme constraint (Poplack 1980), and later the Nonce Borrowing Hypothesis (Sankoff et al. 1990).

2.2 Patterns of Language mixing: Community Patterns and Typological Constraints

When bilingual speakers use two languages simultaneously in discourse, "various problems of incompatibility may arise" (Poplack et al. 1989:389). How the speakers resolve these to build bilingual structures depends on: (i) typological constraints and (ii) community norms (Poplack 2015, 2004; Muysken 2000:221-249; Budzhak Jones 1998c, 1996; Backus 1992; Poplack 1987). Some researchers lend more weight to factors such as length and intensity of contact between the two languages in determining outcomes of contact, including but not limited to mixing (e.g. Thomason and Kaufman 1988), while others favour the structures of the languages in contact in determining emergent patterns (e.g. Treffers-Daller 2001; Woolford 1983). In fact, it has been found that both typological constraints and community norms play key roles in shaping patterns of language mixing when two languages are in sustained contact (Torres Cacoullos and Travis 2014; Parafita Couto et al. 2011; Poplack 2004, 1987; Muysken 2000:221-249; Budzhak Jones 1998, 1996; Bentahila and Davis 1995). Typologically dissimilar languages, for example, are considered to have fewer permissible switch points than a language pair that shares typologically similarity (Poplack 2015, 2004). As a case in point, high levels of intrasentential code-switching have been reported for Spanish-English bilinguals (Poplack 1987,1980; Lipski 1985). On the other hand, typologically dissimilar language pairs, such as Ukrainian-English have been found to exhibit less intrasentential code-switching (Poplack 2004; Budzhak Jones 1998c, 1996; Sankoff et al. 1990; Poplack et al. 1989).

Nevertheless, typology appears to be limited in its predictive power. In Canadian French-English data, for example, intrasentential code-switching was found to be minimal despite numerous permissible switch sites (Poplack 1987). Instead, speakers were reported to exhibit what Poplack (1987) terms "flagged" switching where speakers "flag" the switch with hesitation, pauses, or metalinguistic commentary. Poplack (1987) correlated this type of flagging with the social stigma

attached to the use of English in French discourse in these communities. This contrasts with Puerto-Rican Spanish-English speakers where code-switching was reported to be an overall discourse strategy, used in in-group interactions and emblematic of Puerto Rican identity (Poplack 1987:71-72)[4]. In the case of typologically dissimilar languages such as Finnish-English and Ukrainian-English, speakers employed code-switching at very low rates, if at all, even at the junctions where the grammars presented no conflict and would allow for uninhibited switching (Budzhak Jones 1998c, 1996; Poplack et al. 1989). Furthermore, among the Ukrainian speakers, there appeared to be a tendency to avoid multiword English stretches and when such fragments were used, they were flagged with metalinguistic commentary, as reported for the French-English speakers noted above. Thus, code-switching did not appear to be part of the community norm (Budzhak-Jones 1996:33).

Even where different communities share the same language pair, distinct patterns can emerge. In Moroccan Arabic-French discourse (Nait M'Barek and Sankoff 1988), the dominant pattern of language mixing was found to be constituent insertion, which is where the constituent shows internal consistency with the donor language but occurs at syntactic boundaries consistent with the recipient language (Poplack 2004). On the other hand, this strategy was almost absent in the discourse of Lebanese French-Arabic speakers (Ibid).

Borrowing patterns are subject to similar variability across bilingual communities (Muysken 2000; Poplack et al. 1988; Poplack 1987). In Dutch-Malay language mixing, for example, patterns of verbal borrowing are found to be community-specific rather than uniform across communities (Muysken 2000). Similarly, Spanish-English Bilingual Compound Verbs (BCVs) are not reported for all Spanish-English bilingual communities. Furthermore, even within the same bilingual

---

[4] See also McClure (1977:27) on language mixing among Mexican-American children where language mixing may serve the purpose of marking identity of bilingual speakers as members of a community.

communities, rates of borrowing may differ based on extralinguistic factors, such as social class, among others (Poplack et al. 1988). What these studies show is that simply because code-switching (or borrowing) is possible does not mean that it will constitute a major pattern. On the other hand, community norms generally do not override typological constraints; in other words, if the linguistic constraints prohibit a particular pattern, then social motivations cannot override it (Budzhak Jones 1998c, 1997, 1996)[5]. Typology appears to be a stronger indicator of what type of language mixing is highly unlikely and limited in predicting the patterns will definitively develop in any given bilingual community.

Yet another factor that plays a role in language mixing patterns is bilingual ability (Bullock and Toribio 2012; Treffers-Daller 1994; Romaine 1989; Poplack et al. 1988; Berk-Seligson 1986; Lipski 1985; Poplack 1980; McClure 1977). Intrasentential switching is correlated with high bilingual ability, while intersentential switching, tag switching[6], and borrowings are correlated with less fluent bilingual ability (Bullock and Toribio 2012; Lipski 1985; Poplack 1980). Spanish-dominant speakers who are less proficient bilinguals, for example, will tend to code-switch less and use more single non-native items while the English-dominant speakers, who know little Spanish, are said to utilize Spanish items as identity markers (Lipski 1985:18; McClure 1977).

The above-mentioned patterns are summarized in Poplack's typology of language mixing patterns:

(i) Smooth Intrasentential switching or equivalence based
(ii) Flagged Switching
(iii) Tag switching (also referred to as backflagging (Muysken 2015))
(iv) Constituent insertion
(v) Borrowing

---

[5] Some researchers hold that given the right circumstances in a contact situation, "Anything goes, including structural borrowing that results in major typological changes in the borrowing language." (Thomason 2001:71). Nonetheless, it has been shown that structural change is unlikely in cases of contact (Poplack and Torres Cacoullos 2015).
[6] Tag switches generally occur outside the clause and include items such as *you know*, *right*, and so forth.

Bearing in mind that bilingual ability appears to be a determining factor in some of these patterns, it is important to note that they are not necessarily mutually exclusive. More than one pattern can be found in the same community as a result of differing levels of bilingualism and, alternatively, the identical pattern can be found in disparate communities but at different rates. In Puerto-Rican Spanish in New York, for example, Poplack (1980) found the dominant pattern to be (i), smooth intrasentential switching, but also found (iii), tag switching, and (v), borrowing, particularly among speakers with lower bilingual ability, while in the Canadian French-English data, patterns (ii), flagged switching, and (v), borrowing, were found to be dominant with intrasentential switching occurring at low rates (Poplack 1987).

To summarize, patterns need to be established on a community-by-community basis given that not all bilingual phenomena that are possible for a language pair will be present, or present at comparable rates, in all communities of speakers of that language pair. And, at the same time, where the same bilingual phenomena are present in disparate communities, they may not hold the same status in each of those communities (Poplack 1987:73). The current study aims to ascertain the regular patterns of use of speakers of Belizean varieties of Spanish by means of quantitative analysis of the mixed discourse of speakers of Belizean varieties of Spanish. In this way, how speakers resolve the "various issues of incompatibility" associated with the simultaneous use of two or more languages in discourse can be elucidated (Poplack et al. 1989:389).

2.3 Code-switching and Borrowing

There is general agreement that code-switching and borrowing constitute different manifestations of language mixing (Torres Cacoullos and Aaron 2003); however, there is little consensus beyond this generalization. Uniformity in terminology for both borrowing, code-switching, and types of code-switching remains lacking. Disagreement also centers on the potential distinctions between

these two phenomena. Yet in order to understand patterns of bilingual language use, it is important to have clear working definitions of bilingual phenomena and appropriate methodologies for identifying them. Being able to isolate data that is true code-switching, as opposed to other phenomena such as borrowing or constituent insertion, for example, permits the assessment of grammatical constraints on code-switching, whereas lexical borrowings would not be relevant data to include since they are not governed by code-switching constraints (Poplack 2015).

Early in the literature, attempts have been made to distinguish code-switching and borrowing. Haugen (1956) considered the level of linguistic adaptation of non-native items into the recipient language to be the main criteria where non-native items could be considered borrowings if they were fully integrated (morphologically, syntactically, and phonologically) into the recipient language. Social integration was also considered integral to the definition of a borrowing (Poplack 1980; Hasselmo 1970). The English word *spray,* for example, becomes *espray* in Spanish via the process of prosthesis where an *e* is inserted word-initially to conform to Spanish syllable structure and, in addition, it is a dictionary attested word assigned (masculine) gender[7], all of which point to full linguistic and social integration. As will be discussed further below, however, the definition of borrowing evolved to show that both phonological and social integration are optional processes (Poplack et al. 1988), leaving morphosyntactic integration as the major identifying characteristic of borrowings.

A code-switch would be characterized as maximally distinct from the recipient language (Haugen 1956). Thus, code-switching entails the alternation "between two different grammatical systems or subsystems" (Gumperz 1982:59). In the now famous example taken from the title of Poplack's (1980) seminal paper, seen below in (4), neither the English fragment nor the Spanish fragment is considered a borrowing but rather, each fragment would generally be considered a

[7] Diccionario de la Real Academia Espanola. www.rae.es

code-switch since there is evidence that there are two grammatical systems present. Lexical items and verbal morphology in each fragment, for example, are consistent with the grammars of English and Spanish respectively. In contrast, the English-origin lexical item *espray,* for example, shows indices of integration into Spanish.

4. **Sometimes I start a sentence** in Spanish y termino en español.
   'Sometimes I start a sentence in Spanish and I finish in Spanish.'

General consensus remains elusive with regard to the distinction between code-switches and borrowings and which of the manifestations of bilingual speech should be included in each category (Poplack and Dion 2012; Winford 2009). Code-switching and borrowing are considered to be indistinguishable by some (Winford 2009; Myers-Scotton 2002,1993; McClure 2001; Bentahila and Davies 1991; Gardner-Chloros 1987; Eliasson 1989) or distinguishable only to a limited extent based on frequency criteria (Thomason 2001:134). Furthermore, the task of distinguishing between the two is often presented as unfeasible because the requisite criteria for operationalizing a distinction between the two phenomena are deemed lacking (Gardner-Chloros 2009; Winford 2009, 2003; Thomason 2001; Eastman 1992; Eliasson 1989). It is also presumed that there is no inherent clear-cut boundary between the two phenomena (Haspelmath 2009; Myers-Scotton 2002,1993; Thomason 2001:133; Eliasson 1989) and, instead, they are considered to lie on a continuum (Myers-Scotton 2002; Thomason 2001:135-136). Related to the notion of a continuum is the idea that non-native single-word items may be at some intermediate stage between code-switching and borrowing (Stammers and Deuchar 2012:242; McClure 2001:163) and that at least some types of borrowings start out as single-word code-switches, becoming borrowings over time with repeated use (Myers-Scotton 2002:239; Thomason 2001:135-136).

In contrast, other researchers have shown that code-switching and borrowing can and should be classified as distinct (Poplack and Dion 2012; Poplack and Meechan 1998a; Budzhak

Jones 1998abc; Eze 1998). Additionally, data from such studies indicate that the two respective grammars in a language pair are operative only where code-switching is present, whereas only the grammar of the recipient language is operative when borrowing takes place (Poplack and Dion 2012; Poplack and Meechan 1998a: 127-128). Poplack and associates have applied the variationist comparative framework to the analysis of bilingual data to show that the methodological tools are available for distinguishing borrowing from code-switching (Torres Cacoullos and Aaron 2003; Sankoff et al. 1990; Poplack and Meechan 1998a; Budzhak Jones 1998a,b,c; Eze 1998; Turpin 1998). One major methodological imperative in operationalizing the distinction has been to clearly define *a priori* what will be considered an unambiguous code-switch and what can be considered an unambiguous borrowing (usually established loanwords), thus setting a yardstick for what will be considered a code-switch and a borrowing. Since ambiguous non-native items are compared to corresponding categories in monolingual speech in order to assess the integration into the recipient language, a second imperative, then, is to establish the appropriate comparative benchmark for such an analysis. In keeping with the usage-based approach of the variationist method, comparative monolingual data is drawn from the speakers' own varieties, rather than standard varieties or formal grammars (Poplack and Meechan 1998; Sankoff et al. 1990; Poplack et al. 1988).

When the data is subjected to quantitative comparative analysis using the appropriate benchmarks, the distribution of single and multiword non-native items differ in discourse as well as across functional and lexical categories (Poplack and Dion 2012). Cross-linguistically, borrowings tend to be nouns (Muysken 2000; Poplack et al. 1988; Matras 2009:146-187) and therefore appear in syntactic slots allocated for nouns such as in DPs in object or subject position (Torres Cacoullos and Aaron 2003; Sankoff et al. 1990). Multiword fragments, however, appear in much more varied

syntactic positions, and the onsets of these fragments are distributed more widely across functional and lexical categories (Poplack and Dion 2011; Poplack 1980; Eze 1997; Pfaff 1979).

Furthermore, single non-native items which are ambiguous as to their status of code-switch or borrowing, have been shown to be conditioned by the same constraints as their monolingual counterparts given that they pattern quantitatively with the equivalent categories in the recipient language (Poplack and Meechan 1998a,b; Sankoff et al. 1990). With regard to overt morphological marking, for example, single English-origin nouns in both Tamil (Sankoff et al. 1990) and Ukrainian (Budzhak Jones 1998a,c; Budzhak Jones and Poplack 1997) were case marked with the recipient language case markers in the same ways as native nouns in the corresponding recipient languages (Tamil, Ukrainian). Similarly, gender assignment, plural marking, and noun-adjective word order of English-origin nouns in French all mirror the patterning of French nouns in French discourse with regard to these areas of grammar (Turpin 1998; Poplack et al. 1988). This finding is repeated in French-Wolof where French items in Wolof followed Wolof patterns for determiner marking and adjectival placement (Poplack and Meechan 1995).

Even where items appeared without overt morphological marking, they also patterned with the recipient language. Zero determination, for example, on non-native items paralleled the recipient language rates for French-English (Poplack et al. 1988), Wolof French (Poplack and Meechan 1995), and Igbo-English (Eze 1997)[8]. Likewise, in Spanish-English data, it was found that even when rates were similar for determiner marking on nouns, the grammatical constraints governing zero determination for English-origin nouns in Spanish data were parallel with monolingual Spanish nouns but different from monolingual English (Torres Cacoullos and Aaron 2003). Furthermore, where a language pair such as Tamil-English does not share the same word order, the English-origin items appeared in the Tamil syntactic slot for nouns conforming to Tamil

---

[8] With the exception of indefinite marking of French nouns in Wolof discourse (Poplack and Meechan 1995:211-212)

word order even if not morphologically marked as exemplified in (5) below (Sankoff et al. 1990:80)[9].

5. *Oru*   seal   *pooTTu koDuppaanga*
    one   put (inf)   give (3p-pl-fut)
    'They will put on the seal and give it.' (1855/M)

Additionally, where English-origin nouns surfaced without overt case marking, which according to prescriptive norms are said to be obligatory in Ukrainian and Tamil, it was found that the rates of zero case marking paralleled the zero marking in monolingual Ukrainian (Budzhak Jones 1998c) and Tamil (Sankoff et al. 1990) respectively. In each study where multiword switches were also examined, all of the above patterns differed from multiword fragments in which nouns behaved in ways consistent with the grammar of the multiword switch. In the English portion below, for example, the object is post-verbal and has no overt case marker, both of which are consistent with English grammar but not with Tamil (Sankoff et al. 1990:84).

6. *batilaa immediate-oa*   they take the easiest route
    instead (adv)
    'Instead, they immediately take the easiest route'. (254/M)

These studies also showed that morphosyntactic integration was evidenced on the non-native single items, in their majority, regardless of frequency in the corpus (Torres Cacoullos and Aaron 2003; Budzhak Jones 1998c; Eze 1998; Sankoff et al. 1990) or phonological integration (Poplack et al. 1988). And while the results of these studies overwhelmingly point to non-native single items in a host language being borrowings, the methodology employed also facilitated the identification of potential single-word code-switches. Meechan and Poplack (1995:186-187), for example, found that French adjectives in Fongbe [ɖó +ADJ] constructions did not appear with Fongbe morphology

---

[9] Where examples are cited from the literature, translations as well as transcription and glossing conventions are retained from the source.

but rather with French morphology, as shown below. These adjectives also appear at equivalence sites between the two languages, sites propitious to code-switches, and did not pattern with native Fongbe adjectival constructions.

7. Donc   ɔ̀   nyɛ́   mɔ̀. ɖɔ̀ que langue   ɔ̀   é ɖó importante.
    So   TOP I   see tell that language DEF   she be important
    'So, me I see that language is important. '(Fongbe 4:5pkr 04:480)

French adjectives in Fongbe, in contrast to French nouns in Fongbe discourse, were consequently considered to be best analyzed as code-switches rather than borrowings[10]. Thus, the variationist comparative method does not simply aim to identify non-native items in a host language as borrowings, but rather to identify which might be borrowings and which might be code-switches when these are present (Poplack 1993; Poplack and Meechan 1998a; Poplack et al. 1988:53).

The distinction between the underlying processes that give rise to manifestations of language mixing, which on the surface may appear identical, is an important one. If these surface forms are produced from different underlying processes and are, in fact, *different linguistic elements,* then treating them as the same obscures the facts of bilingual behavior, is descriptively inadequate, and will have little explanatory power (Poplack and Dion 2012; Poplack 1993; Budzhak Jones 1998abc, 1996; Eze 1997). If borrowings are produced via a mechanism that is different from one which produces code-switching, for example, then borrowings are not governed by the same grammatical constraints as code-switches and are therefore irrelevant data for testing constraints on code-switching and should be excluded from any such analysis.

---

[10] On other studies that revealed single word switches see Turpin (1998:230), Meechan and Poplack (1995:186-187), and Poplack and Meechan (1995:212).

2.3.1 Grammatical constraints on code-switching

While the question of borrowing versus code-switching remains open, it continues to be relevant for modeling language mixing (Poplack and Dion 2012; Stammers and Deuchar 2012), the once hotly debated issue of constraints on intrasentential code-switching is now less controversial and less discussed. More researchers are taking the position that there are no universal constraints on intrasentential code-switching (Moro 2015; Stell and Yakpo 2015; Eppler 1999:286) or that there continue to be impediments to arriving at a set of universal constraints (Poplack 2015, 2004). However, attempts to account for the rule-governed behaviour of code-switching (alternation between two languages' grammatical systems within the same sentence), that might be universally applicable, have not ceased. Some of the most current approaches include MacSwan's (2014, 2000, 1999) minimalist approach, Myers-Scotton's Matrix Language Frame Model (Myers-Scotton 2002; Myers-Scotton and Jake 2001), Muysken's (2013, 2000) typology of language mixing, and variationist approaches (Poplack 1993, 1980; Sankoff and Poplack 1981), all of which continue to be both critiqued and applied to the analysis of language mixing[11].

In the Matrix Language Frame model (Myers Scotton 2002, 1993; Myers-Scotton and Jake 2001) a central principle is the assumption of inherent asymmetry in bilingual speech. One language is considered to dominate, acting as a matrix language (ML) for embedded morphemes or islands from the other embedded language (EL). Additionally, there is morpheme asymmetry in that the embedded language contributes more content morphemes than system morphemes[12].

---

[11] Recent applications of MacSwan's research program can be found in MacSwan (2014); Parafita Couto et al. (2014); of Muysken's typology in Hoi Ying Chen (2014), Lipski (2009) and Deuchar et al. (2007); of the MLF model Parafita Couto et al. (2014); Herring et al. (2010) and Callahan (2004); and of Poplack's FMC and EC in Schmidt (2014). For a recent critique of the equivalence constraint and the free morpheme constraint see Moro (2015). Other approaches that adopt the minimalist framework include Fernández-Fuertes et al. (2016) on subject-verb alternations when the subject is a pronoun or an NP.

[12] System morphemes do not assign thematic roles but rather indicate the relations between content morphemes; these are roughly analogous to functional categories. Content morphemes, which are the main elements that convey semantic and pragmatic aspects, do assign or receive thematic roles (nouns, verbs) (Myers Scotton 2002, 1993).

Furthermore, morpheme order is determined by the matrix language (the language that contributes the system morphemes), and grammatical constraints on code-switching arise largely from restrictions on embedded language morphemes. However, despite later revisions of the model, the ML cannot always reliably be identified[13] and remains a source of controversy. Some even argue that the ML construct itself is unnecessary for an adequate analysis of code-switching (MacSwan 2010, 2005), particularly if the only requirement for switching is homologous grammars at the switch site (Poplack 1980). In addition, under the ML model, it is difficult to identify the manifestations of language mixing. Code-switches and borrowings are distinguished, in part, based on frequency and whether or not the items are tagged in the bilingual's repertoire for the embedded or matrix language. Code-switches do not recur, and the lemmas underlying code-switches are tagged for the EL only, whereas borrowed forms recur and have "forms tagged for both donor and recipient languages" (Myers-Scotton 2002:153). The distinction of tagging is at the level of the mental lexicon and thus difficult to clearly detect in discourse. In addition, in this model, at least some code-switches may become borrowings over time with frequent use (Myers-Scotton 2002:41). It appears, then, that while there is recognition that these two phenomena are distinct, the model does not offer a concrete distinction between the two phenomena. Additionally, in keeping with this approach to borrowing and code-switching, both languages in the language pair are considered to be activated for the speaker for both borrowing and code-switching. This underlying assumption does not find support in the literature given that borrowing and code-switching have been found to be correlated with lower and higher levels of bilingualism respectively (Bullock and Toribio 2012; Poplack 1980).

Another approach to modeling bilingual data divides the surface phenomena of language mixing into three main types or categories: insertion, alternation, congruent lexicalization (Muysken

---

[13] Herring et al. (2010), for example, found that they could only identify a matrix language for 70% of their data.

2000) and a fourth, more recently added category, called backflagging (Muysken 2015, 2013). All these categories fall under Muysken's (2000) umbrella term code-mixing. *Insertion* refers to material from one language appearing, or as the name suggests, "inserted" in the structure of another. This material may "involve single elements, akin to borrowing, as well as more complex constituents" (Muysken 2000:33). Borrowing is analyzed as a special case of insertion. *Alternation* involves switching between structures (between clauses or utterances) from the respective languages (Muysken 2000:3,33). The third type of code-mixing is *congruent lexicalization* where material from different lexicons appears in a shared grammatical structure (Ibid). Finally, *backflagging* is analogous to tag-switching which refers to the use of English-origin items clause–externally, often at the end of a clause (e.g. *right*, *¿verdad*?). These categories are descriptively adequate in many ways since they capture distinctions between the different manifestations of language mixing, setting them apart from those theories, such as the MLF, which do not. Another contribution of this approach is (limited) quantitative analysis of data and consideration of both linguistic and extralinguistic constraints (Muysken 2000:30). Additionally, Muysken advocates against the analysis of individual examples on a case by case basis, instead privileging the most frequent forms found in data as the main body of evidence (Muysken 2000:29).

The difficulties of this model lie in the challenge of determining which of the four categories the data fall into; equally difficult is validating the model empirically, no doubt due, in part, to sometimes unclear criteria (Muysken 2015; Muysken 2000:29). Morphological integration, for example, is a feature of congruent lexicalization due to the similarity between the two grammatical systems, but this same criterion is also listed as a diagnostic of insertion (Muysken 2000:231). Borrowing on the other hand, is not in a separate category as it occurs under both insertion and congruent lexicalization (Muysken 2000:16). Another drawback is that the analysis is based on a meta-analysis of the literature (Muysken 2015, 2000) and is therefore dependent on reported data

as well as on formal grammars which do not take into account the inherent variability of language. What formal grammars hold as ungrammatical, for example, may appear as productive patterns in spontaneous speech.

The most recent approach to code-switching within the generativist tradition is MacSwan's (2014, 2000) adoption of the Minimalist Program (Chomsky 1995, 1993) in what he terms a constraint-free approach. This approach contrasts with what he calls constraint-based approaches or code-switching specific approaches which, in his view, depend on supplementary mechanisms, such as reference to specific languages or code-switching constraints, to account for language mixing (MacSwan 2014:230, 2010, 2000). MacSwan (2014:76), for example, considers a language index unnecessary to account for mixed speech since grammars are formally blind to the languages they generate. His constraint-free approach is summarized below (MacSwan 1999:146):

Nothing constrains code-switching apart from the requirements of mixed grammars.

The same operations that build monolingual phrase structure are taken to also apply to code-switches; thus, language specific properties are avoided at the outset since these enter late in the numeration. Under this approach, it is possible to code-switch at any juncture provided that morpho-phonological dependencies are not disrupted due to the fact that the morpho-phonologies of different languages have different properties (MacSwan 1999 as cited by Stabler and MacSwan 2014:257; MacSwan 2014)[14]. One challenge is that "code-switching" under this approach is not clearly defined. The term seems to refer to any type of alternation including both words and

---

[14] In principle, this approach coincides with the Poplack's (1980) free morpheme constraint (MacSwan 2005).

multiword constructions with mixed etymologies, despite copious data suggesting that these represent different linguistic phenomena (Poplack and Meechan 1998b)[15].

As a model couched within the generativist framework, it aims to determine not only what code-switches are permissible but also what structures cannot be generated (MacSwan 2005:2). Thus, there is a reliance on informant elicitation, judgment, and introspection regarding often decontextualized or invented examples (Stabler and MacSwan 2014; Cantone and MacSwan 2009; MacSwan 2010, 2005). What is not discernable in this approach is whether structures deemed grammatical constitute productive patterns or if such structures are evaluated by an appropriate sample of speakers (e.g. proficient bilinguals for whom code-switching is a discourse mode) rather than by a select speaker or select few. While much has been learned about language through these approaches, unless experimental procedures properly control for speakers' internalized normativity in relation to community-based mixing strategies, it is possible that these findings might neither reflect actual use nor the underlying grammars constraining use of mixed discourse (Poplack and Torres Cacoullos 2015), in particular with regard to intrasentential code-switching,[16] signaling the need for complementary approaches. Speakers, for example, may reject the very structures they use, based on internalized norms rather than grammaticality (Milian Hita 2014:fn18; Muysken 2000:29; Labov 1996). As noted in French-English data (Poplack 1987), for example, social stigma interacts with linguistic factors to shape code-switching behavior, resulting in minimal intrasentential code-switching amongst those speakers. This, in turn, may lead speakers

---

[15] Both the following items are discussed as code-switching: *parquear* 'to park,' and ¿*Funciona* **the computer** *de tu hermano en la oficina*? 'Does your brother's computer in the office work?' (MacSwan 2005:7,10).
[16] The Functional Head Constraint (Belazi et al. 1994), for example, which was based on judgment tasks, has not found support in production data (Moro 2015). This proposal predicted that switching could not occur between functional heads such as AUX/MOD, NEG, COMP, QUANT/NOM and DET and their complements (Belazi et al. 1994); however, some of these very sites emerge as favored code-switching sites (e.g. Lipski 2008:55).

to judge those types of code-switching which are permissible in the language pair as ungrammatical [17].

These models differ in whether or not they treat all manifestations of language mixing as the same and whether they then attempt to account for them under one model. Furthermore, these disparate approaches differ in terms of the aspect of speaker competence that is considered central. The minimalist approach, or constraint-free approach, for example, aims to ascertain the internal system of rules that govern code-switching (MacSwan 2014, 2005). A third difference among these approaches is consideration of the role of extralinguistic factors on constraining language. Both the MLF model and Muysken's typology give some attention to extralinguistic factors, but these are not part of generativist models of language. Therefore, the goals of each approach may differ significantly, and, consequently, the type of data that is considered appropriate for analysis as well as the methodological tools also vary.

Thus, some researchers, even when testing approaches such as MacSwan's (2014, 1999), advocate for the use of production data, even if this is limited to reported data (Milian Hita 2014). Others urge for the use of strategies for mitigating confounding variables, such as the social stigma associated with code-switching, when using judgment tasks (González Vilbazo et al. 2013) and for the application of quantitative reasoning rather than analysis of single examples (Poplack 2015). Alternatively, triangulation of methodologies has been suggested in the literature (Pérez Leroux et al. 2014:288)[18]. No framework has a privileged window on the study of bilingualism and language contact, and theoretical pluralism is always to be preferred to theoretical monism. There is much to be learned about bilingualism, language-mixing, etc., drawing on the complementary

---

[17] In addition, when it comes to bilingual data, differing levels of bilingual ability are correlated with different patterns of mixing (Poplack 2004) which may consequently affect intuitions and judgments of bilingual data. The same can be said for the age of acquisition of the second language (Gingrás 1974 as cited in Lipski 1985:18). To complicate matters more, intrasentential code-switching remains primarily an oral mode of communication which further confounds the use of written judgment tasks.

[18] In fact, some researchers consider experimental data together with production data (Liceras 2013).

strengths and weaknesses of different approaches. What the variationist method (Labov 1972) can contribute is the quantitative analysis of naturalistic or vernacular data. The method was developed to ascertain and account for the variable use of structures by speakers in a given speech community and knowledge of the grammatical constraints governing variable use is considered the central aspect of speaker competence (Weinreich, Labov and Herzog 1968:188). Thus, the approach that is pursued in the ensuing chapters is one that privileges spontaneous speech and the community-based motivations for language mixing. The kinds of questions this thesis poses are ones which the variationist framework is particularly apt to address. The most recent findings and contributions of this approach are discussed in the next section.

2.3.2 Variationist approaches to code-switching and borrowing

Two constraints on code-switching have been proposed from variationist approaches to language mixing: the equivalence constraint and the free morpheme constraint (Poplack 1980). A later corollary to these two constraints is the Nonce Borrowing Hypothesis (Sankoff et al. 1990; Poplack et al. 1988). As will be shown, together they account for both code-switching and borrowing. A detailed discussion is provided below, including the evolution of the Nonce Borrowing Hypothesis.

The free morpheme constraint and the equivalence constraint are summarized in (I) and (II):

I. Free Morpheme Constraint (hereafter, the FMC)

Codes may be switched after any constituent in discourse, provided that constituent is not a bound morpheme. This constraint holds true for all linguistic levels but the phonological (Poplack 1980:585-586)[19].

---

[19] This means that if a non-native lexical item is rendered in one syllable with native phonology and in another syllable in non-native phonology, this is considered transference of phonological patterns and not a "switch" between bound morphemes. For example, in the production of *excuse*, where there is aspiration of /s/ before a voiceless consonant in the first syllable and the second syllable is produced with English pronunciation, this represents transference of the aspiration tendency of /s/ in Caribbean Spanish and not a switch (Poplack 1980:586). The issue of phonology is further clarified: "…a switch may not occur between a bound morpheme and a lexical form unless the latter has been phonologically integrated into the language of the bound morpheme" (Sankoff and Poplack 1981:5).

Included under this constraint are idiomatic expressions, such as cross my fingers [sic] and hope to die and Si Dios quiere y la virgen ('God and the Virgin willing') which are considered here to behave like bound morphemes in that they show a strong tendency to be uttered monolingually (Poplack 1980:585).

II. Equivalence Constraint (hereafter, EC)

Code-switches will tend to occur at points in discourse where juxtaposition of L1 and L2 elements does not violate the syntactic rule of either language, i.e., at points around which the surface structures of the two languages map onto each other. According to this simple constraint, a switch is inhibited from occurring within a constituent generated by a rule from one language which isn't shared by the other" (Poplack 1980:586).

The notion of code-switching and equivalence between grammars is one that appeared early in the literature (Pfaff 1979; Lipski 1978; Lehtinen 1966 as cited in Muysken 2000:11) but was clearly named and stated as a constraint within a model of code-switching by Poplack (1980) and later formalized in Poplack and Sankoff (1981) and Sankoff and Mainville (1986).

Both constraints, however, met with much criticism and apparent counterevidence. One of the early critiques of the EC, for example, was that it was specific to Spanish-English data or other typologically similar language pairs and might not apply to typologically distinct languages or agglutinative languages (Gardner-Chloros 2009:96; Halmari 1997:76-77; Myers-Scotton 1993:31). Nonetheless, the EC has been upheld in a number of corpora, both with typologically similar (Poplack 1980; Pfaff 1979) and typologically distinct language pairs (Budzhak Jones 1998c, 1997; Eze 1997; Poplack et al. 1989), as has the FMC (Budzhak Jones 1998c; Sankoff et al. 1990; Poplack et al. 1988; Berk-Seligson 1986; Pfaff 1979). Thus, the constraints are not specific to Spanish-English or typologically similar language pairs. Furthermore, from their inception the EC and FMC were stated as tendencies; they were never intended to be categorical. In addition, they are not expected to be "uniformly pertinent" to all bilingual communities (Poplack et al. 1989:390) since smooth intrasentential switching may be limited by typological constraints or community

norms. Poplack (1993:276) contends that because code-switching is a discourse mode, it cannot be predicted when a code-switch (or a borrowing) will occur at the local level.

Other critiques of the EC and the FMC were presented in the form of presumed counterevidence taken from judgment data (e.g. MacSwan 2010:286-287; Toribio 2001:206-207). In (8)[20], for example, a code-switch is presumed to be permissible in both (8a) and (8b) since word order is the same for Spanish and English, but (8a) is judged ungrammatical. Yet, as discussed above, there is much disagreement on the validity of judgment tasks for ascertaining constraints on code-switching.

8. a. *__The students had__ *visto la película*.
    b. *Los estudiantes han* __seen the movie__.
    'The students had seen the movie.'

Counterevidence from production data was also presented but in the form of isolated single-word examples[21] as shown in (9) where *parents* is considered to be a code-switch in violation of the EC.

9. __Parents__ *te* __depend__ *hona ε*
    'It depends on the parents.' (Romaine 1986:40-42)

In this example, English and Punjabi do not share the same prepositional word order; therefore, a "code-switch" should not occur at the boundary between *parents* and *te* since the languages are not equivalent at this juncture. This example illustrates the need for understanding the difference between a code-switch and a borrowing and the relevance of this difference to modeling constraints on code-switching. Punjabi, has postpositions; therefore, the English noun *parent* is occupying a Punjabi syntactic slot and not an English one, which suggests that it is integrated into Punjabi. Bearing in mind that code-switching is not simply an alternation but rather a switch between grammars of the respective languages, *parents* might well be a borrowing, and the status

---

[20] Originally cited in Belazi et al. (1994). Glossing is from cited source.
[21] See also Myers-Scotton (1993), Eliasson (1991), Bokamba (1987), Bentahila and Davies (1991), among others.

of such items must first be ascertained before it is decided that they constitute relevant data for assessing code-switching constraints.

Likewise, single-word examples offered up as violations of the FMC were of little value in refuting the FMC. Eliasson (1991), for example, cites an English-origin word adapted into Maori: *helptia* 'to be helped', as counterevidence to the FMC which prohibits a switch between a lexical stem and a bound morpheme unless the lexical stem has been phonologically integrated into the language of the bound morpheme[22]. As in the case of *parents* in the Punjabi-English example, *helptia* appears to be integrated into Maori. A quantitative comparative analysis would be helpful in this case to show whether items such as *helptia* pattern with Maori and are borrowings or if they pattern with English and are code-switches; in which case, they would be relevant (in the aggregate) to testing code-switching constraints. Isolated decontextualized examples are generally not the most suitable source of evidence for testing constraints. Rather, the most frequent forms found in data should constitute the main body of evidence and that infrequent forms should be taken as possible "fluke phenomena, performance errors, and the like" (Muysken 2000:29), a position shared by the variationist approach (Tagliamonte 2012; Poplack and Meechan 1998; Poplack 1993; Labov 1972:95). This is not to say that the counterexamples are to be dismissed, but rather that when they are insufficient, they may require further exploration. In other words, "...the single-word examples offered up repeatedly in the literature as counterexamples to one or the other constraint are simply irrelevant without systematic corpus-based assessments of whether they are, at least in the aggregate, borrowings or switches" (Sankoff et al. 1990:97). In addition, the FMC was intended to also encapsulate fragments longer than a word (Budzhak Jones 1998c; Sankoff et al. 1990), such as switching within idiomatic expressions (Poplack 1981:175). According

---

[22] See Eze (1998:38-50) on non-native single words in some African languages; Halmari (1997:76) on Finnish-English; Eliasson (1991) on Maori-English and also a summary of other language pairs; and Romaine (1989:137-138) on various language pairs.

to the FMC, it is possible to switch full sentences as well as any constituent within the sentence provided that the constituent consists of at least one free morpheme (Poplack 1981:175). This is further clarified by Sankoff et al. (1990:97):

> *There is another, less important, aspect to the free morpheme constraint as originally formulated, namely that switches should not occur across a morpheme boundary where one of the morphemes is at the end of a clearly monolingual multiword fragment in one language and is bound to another at the beginning of a clearly monolingual multiword fragment in the other language. Were many such switches to occur, given the multiword context of the free morpheme in its original language, there would be less justification in regarding it as a borrowing, and the free morpheme constraint would be falsified.*

The counterevidence based on single words was perhaps inadvertently fostered by the now well-known example *eat-iendo* 'eating,' which was initially considered to be a violation of the FMC due to the so-called "switch" between English (phonologically unintegrated) *eat* and Spanish bound morphology *–iendo*, '–ing' (Poplack 1980:586). Consequently, each time a lexical item with etymology from one language and morphology from another was identified in a corpus or considered acceptable in a judgment task, it was considered a violation of the FMC regardless of phonological integration which was an essential criterion of the FMC. This is in part because the FMC and EC were considered to apply to all alternations, including single and multiword alternations. Recall that at the time these constraints were proposed, there were unclear distinctions between code-switching and borrowing, and code-switching most often referred to any alternation to a non-native fragment that was not both linguistically and socially integrated into the recipient language, an issue that is addressed in later work (cf. Poplack et al. 1988). Some of the confusion regarding the FMC and EC, then, derives not from the proposal itself, but rather from the working definitions of code-switching and borrowing current at the time and which therefore underlie the original proposal. Furthermore, the constraints initially attempted to account for both code-switching and borrowing under a "single model" (Poplack 1980:585), long since abandoned

by Poplack and associates given that the early definitions of borrowing and code-switching were later shown to be inadequate. Borrowing was shown to be a morphosyntactic process where social and phonological integration are only optionally present (Poplack et al. 1988), and to be produced from mechanism distinct from that which produces code-switches (Sankoff et al. 1990; Poplack et al. 1988). Subsequently, it became evident that borrowing and code-switching could not both be addressed by the FMC and EC. These more recent developments in combination with the original proposals permit an adequate and empirical account of bilingual data.

Finally, the FMC and EC were stated as tendencies (Poplack 1980; see also Muysken 2000)[23]. If exceptions to constraints on monolingual speech are not themselves categorical, then it should not be surprising that exceptions may also be found in bilingual speech. Poplack (1980:fn6) herself noted some cases which appeared to be exceptions: a switch is permissible at the subject/verb juncture if the subject is a full NP but not if it is a pronoun. In addition, in (10), below, there is an alternation (considered a "switch" at the time) found at the NP-ADJ boundary which is not an equivalence site for Spanish and English. This was considered to be a violation of the EC at the time[24].

10. *Las palabras* **heavy-duty**, *bien grandes se me han olvidado* (40/85)
    'The heavy-duty words, big ones, I've forgotten.'

However, as the distinction between code-switching and borrowing began to be empirically defined more clearly, the data that was considered valid (or not valid) for testing the EC and FMC also began to be more clearly defined. If code-switching and borrowing are different processes, then borrowings are not good candidates for testing the EC since they are not governed by it.

---

[23] Though they did account well for the Spanish-English data (Poplack 1980; Pfaff 1979; Lipski 1978).
[24] Since then studies have shown that (10) is best analyzed as a borrowing given the level of integration. The adjective is postposed to the noun, which is consistent with Spanish adjectival placement.

In the following section the refinement of how the FMC and EC apply to bilingual data and the concomitant development of the NBH is discussed. Given the misunderstandings in the literature since the EC and FMC were first proposed, some of which were noted above, it is important to clarify their development if the variationist approach to language mixing is to be the basis of analysis in this thesis.

2.3.2.1 The evolution of the nonce borrowing hypothesis

As indicated above, the FMC and EC relied on assumptions regarding borrowing and code-switching that were current at a time when studying bilingual speech was still in a relatively early stage (Poplack 2012:644). Code-switching, for example, had only just begun to be recognized as systematic and rule-governed (Lipski 1985; Gumperz 1982; 1978; Pfaff 1979; Timm 1975; Hasselmo 1970). Thus, in Poplack's early work (1980:583-584; Poplack 1981:170), code-switches and borrowing were alluded to as distinct phenomena but, in keeping with the literature of the time, the differentiation was based largely on phonological criteria (Poplack 1980:583) or criteria of social integration (Poplack 1987:56; Poplack 1980:598), and single items were not separated from multiword fragments if they were found to lack social integration (Poplack 1980:602)[25]. Single nouns, for example, are referred to as a form of switching, i.e., "noun-switching" (Poplack 1981:171; 1980:602). As the research continued to evolve, momentary or nonce borrowings were acknowledged but also kept separate from "borrowing on the community level," i.e., diffuse or established loanwords (Poplack 1987:72), though the relationship between the FMC and nonce borrowing began to emerge since it was found that the FMC, which prohibits mixed morphologies and mixed phonologies within a word, could be "circumvented" through momentary borrowing

---

[25] "Spontaneous switches of **words**, sentences and larger units at a turn boundary, not involving any change in interlocutors, were also considered to be code-switches if they exhibited…" lack of integration into Spanish. (Poplack 1980:585); (emphasis my own).

(1987:69). Thus, it became apparent that phonological integration was not necessarily a concomitant process to morphological and/or syntactic integration.

In fact, Poplack et al. (1988) found that there were a large number of items that did not fit the traditional definition of borrowings in terms of both social integration or phonological integration. Methodological criteria were established in order to determine, relative to the corpus from which data was drawn, if such items were widespread in the community, recurrent, used only once (nonce), or idiosyncratic. These were then compared to unambiguous borrowings, i.e., socially established loanwords that were linguistically integrated, and it was found that they were morphosyntactically integrated in the same fashion as the established loanwords, which by definition are both socially and phonologically integrated (Poplack 2012; Poplack et al. 1988). Thus, in order to ascertain the status of ambiguous items, both code-switching and borrowing must first be defined. The ambiguous items can then be measured against these clear-cut definitions. Furthermore, these *a priori* definitions of borrowings and code-switches ought to appeal to generally agreed upon definitions in the literature rather than being ad hoc formulations.

Sankoff et al. (1990) further developed this work by assessing single English-origin items in a typologically distinct language pair (Tamil-English) and were able to show that single lexical items, used even once, showed linguistic integration at the outset which is consistent with a borrowing analysis. These facts are elaborated in the nonce loan hypothesis "…, which basically states no more than that borrowing, whether nonce or established, is a phenomenon of language mixture distinct from code-switching and is operationally distinguishable as such, at least at the aggregate level, adds to the predictions of the free morpheme constraint in its most important context, the occurrence of single words from one language in a sentence or multiword fragment of the other language" (Sankoff et al. 1990:97). Since then, numerous studies have shown that nonce borrowings differ from established loanwords only in diffusion (Torres Cacoullos and Aaron 2003;

Poplack and Meechan 1998; Poplack and Meechan 1995:200; Poplack 1993:256; Poplack et al. 1989:392) and phonological integration (Sankoff et al. 1990:74; Poplack and Sankoff 1988:1179; Poplack et al. 1988:72).

The criterion of phonological integration, central to the FMC and formerly crucial to differentiating manifestations of bilingual speech, was eventually rejected as a reliable diagnostic for differentiating code-switching and borrowing and as a requisite characteristic of borrowing (Poplack et al. 1988; Poplack 1985). This was because in the French-English data, fluent bilingual speakers engaged in applying a distinctive stress pattern to English nonce borrowings, a feature which was not applied in monolingual French or English discourse (Poplack 1987:70). Second, and perhaps more importantly, phonological integration of borrowings was found to be gradual and to occur as a function of social integration and diffusion; therefore, the phonology of borrowings, specifically nonce borrowings, is highly variable in synchronic data and consequently unreliable as a diagnostic (Poplack et al. 1988:72). Third, phonological transference is not unusual in bilingual communities; as such, the pronunciation of non-native items with full or partial native language phonological features complicates the use of phonological integration as a diagnostic (Budzhak Jones 1998c:51-52; Poplack 1987:72; Poplack 1980:586). Finally, the use of phonological integration as a diagnostic rests on the assumption that bilinguals have completely separate phonologies, which may not be the case (McClure 2001; Poplack et al. 1988; Poplack 1987:72; Poplack 1980:586).

One of biggest contributions of the nonce-borrowing hypothesis, and the quantitative studies supporting it, is the finding that the accepted definitions of borrowing were based on a faulty premise, i.e., that all levels of linguistic integration proceed in the same fashion and that social and linguistic integration were closely intertwined and even required mechanisms in the borrowing process. Syntactic, morphological, and phonological integration do not occur at the

same rate; the former two are abrupt while the latter is gradual (Poplack and Dion 2012; Poplack et al. 1988). Accordingly, morphosyntactic integration is independent of social integration; it is only phonological integration that has a close relationship with social integration. Another major contribution from the studies that gave rise to the NBH is the evidence pointing to borrowing and code-switching as distinct processes along with the methodological tools to distinguish them.

As a result of this distinction being made, the NBH also presents a strong challenge to the notion of word-internal switching, if "switching" represents the presence of two separate grammars and the alternation between them. When mixed morphologies are present within one word such as *helptia* or *eatiendo,* and such items are analyzed quantitatively, it is found that the such items tend to be integrated into the recipient language, retaining only the etymology of the donor language but not its grammar. Therefore, these are not examples of word-internal "switching." Thus, the predictions of the FMC, EC, and NBH, all empirically motivated, together accounted for the vast majority of surface manifestations of language mixing.

The current study avails itself of the major contributions of the NBH as discussed above: (i) borrowing is a morphosyntactic process (it may or may not involve phonological and social integration); (ii) borrowing and code-switching are distinct mechanisms; and (iii) these mechanisms can be distinguished in the aggregate by using the tools of the variationist comparative method.

2.4 Verbal Borrowing and Bilingual Compound Verbs

Most of the studies discussed above focus on nouns, in part because these tend to be the most borrowed category (Muysken 2000; Matras 2007; Poplack et al. 1988). However, verbs still constitute a considerable portion of the data in some corpora (Meakins and O'Shanessy 2012; Poplack et al. 1988; Budzhak Jones 1998; Eze 1998, 1997; Nishimura 1995; Pfaff 1979). In the current study, for example, English-origin verbs make up 14% of the English-origin single items. In

view of this fact, coupled with reports in the literature that Spanish speakers in Belize frequently employ bilingual compound verbs (BCVs) (Hagerty 1996), special attention is paid to the English-origin verbal items in this study.

Spanish-English data has been reported to present three main patterns of verbal borrowing[26]: (i) integration into the –*ar* class of verbs (Sánchez 1995[27]; Pfaff 1979), (ii) bare verbs[28] (Silva Corvalán 1999; Pfaff 1979:301), and (iii) bilingual compound verbs (e.g. *hacer retire*) (Fuller Medina 2015, 2013, 2010, 2005a,b; Balam et al. 2014; Wilson and Dumont 2014; Wilson 2013; Toribio et al. 2012; Hagerty 1996; Silva Corvalán 1994; Rubinstein 1979; Pfaff 1979; Reyes 1976). The preferred strategy employed tends to differ by community, though bare verbs seem to be the most infrequent pattern. The strategy of bilingual compound verbs in Spanish-English data has recently begun to garner increased attention despite having been noted early in the literature (Hagerty 1996 and Rubinstein 1979 in Belizean Spanish; Pfaff 1979 and Reyes 1976 in Spanish in the U.S.). It is reported as a minor pattern in Spanish in the U.S. (Jenkins 2003; Silva Corvalán 1994; Pfaff 1979) while it is claimed to be frequent in Belize (Hagerty 1996).

The first full-length treatment of Spanish-English BCVs was by Jenkins (2003) who provided a descriptive study of production data from Southwest Spanish. This was followed by a quantitative empirical study of BCVs in Belizean Spanish (Fuller Medina 2005a,b) addressing the structure of BCVs, frequency, and aspectual constraints, with later theoretical analyses examining the internal structure of BCVs (Fuller Medina 2013, 2010). Since then, other quantitative studies have ensued which address questions raised in these earlier studies regarding cognitive load in relation to the use of BCVs (Wilson and Dumont 2014) and frequency constraints (Balam et al.

---

[26] Pfaff provides a more detailed typology which is not replicated here since many of the tokens appear too low in number to constitute a regular pattern and also because her typology, while exhaustive, contains a mix of both code-switches and borrowings.

[27] Data consists of a corpus of written text.

[28] "*…caminaba…y luego* ***transfer***," "…I used to walk…and then transfer" (Pfaff 1979:301)

2014; Wilson and Dumont 2014; Wilson 2013). Aspectual constraints have also been revisited (Balam et al. 2014).

Because BCVs have been documented in numerous language pairs and have a long history of being analyzed in these language pairs[29], various proposals have aimed to describe their use and structure (e.g. Muysken 2000; Ritchie and Bhatia 1996; Romaine 1986; Moravcsik 1975), to account for the apparent novelty of the structure (Wilson 2013; Gardner Chloros 2009; Edwards and Gardner Chloros 2007), to ascertain if they present counterevidence for proposed constraints on code-switching (e.g. Romaine 1995,1986), and to determine why bilinguals use these structures (Jenkins 2003; Myers-Scotton 2002; Edwards and Gardner-Chloros 2007)[30]. These are discussed in detail below.

With regard to the question of why bilinguals use BCVs, it has been suggested that it is an attempt to fill lexical gaps (Jenkins 2003; Reyes 1976) and that BCVs may have the function of vocabulary extension (Muysken 2001:185). And, in fact, most of the token types produced by Belizean Spanish speakers were associated with technology (e.g. *hacer* download) suggesting that novel or previously unknown contexts can give rise to borrowing via structures such as BCVs (Fuller Medina 2005). Nevertheless, the creation of new terms in this fashion is a regular process found in monolingual language use. Thus, such uses in bilingual speech need not be taken as compensatory.[31] In addition, speakers have been found to use both BCVs and monolingual simplex verbal equivalents in their speech, suggesting that they have both forms available to them (Wilson 2013; Hagerty 1996:136-137). As such, the data does not support the lexical gap hypothesis.

---

[29] For an overview see Muysken (2000:184-220), also Wohlgemuth 2009 on an inventory of language pairs. For Japanese-English (Nishimura 1995); Igbo-English (Eze 1997); Mayan-Spanish (Suárez Molina 1996).
[30] An exception to this is González-Vilbazo and López (2012, 2011) who use bilingual data comprised of BCVs to test theoretical constructs; their primary goal is to test the hypothesis that little *v* is the head of a vP phase.
[31] *hacer clic* 'to click,' for example, is found in monolingual Spanish.

With regard to the cognitive load hypothesis, Jenkins (2003) notes repair phenomena in the vicinity of some BCVs which he analyzes as a compensatory strategy due to the presumed heavy cognitive load which is created by trying to recall the monolingual Spanish verb. Wilson and Dumont (2014:5), however, found no evidence of this. They examined disfluencies in the intonation units (IUs) containing BCVs and compared these to other IUs of mixed discourse and monolingual Spanish and found no significant difference between the three datasets. Thus, disfluencies preceding BCVs cannot be taken as indications of a heavy cognitive load.

Related to this is the consideration that BCVs, or light verb strategies, are a more "cumbersome" borrowing strategy for bilingual speakers (Wichmann and Wohlgemuth 2008). Based on a cross-linguistic inventory of loan verbs, they propose a loan verb integration hierarchy (LVIH) as seen below in (11).

11. LIGHT VERB STRATEGY < INDIRECT INSERTION < DIRECT INSERTION < PARADIGM INSERTION

In this hierarchy, verbs borrowed via strategies further to the right of the hierarchy are considered to be most integrated and require fewer mechanisms for integration while those borrowed via strategies to the left are less integrated, requiring more mechanisms. The light verb strategy, which includes BCVs, produces verbs that are least integrated according to these authors and require "a whole extra constituent" (i.e., the light verb) to integrate the non-native verb. This remains to be empirically tested as the authors note that the level of bilingualism in the community may play a role in which strategy is chosen. Indeed, if the light verb strategy is the most cumbersome, then the question arises as to why this would be chosen as the primary strategy of verbal borrowing in some communities[32].

---

[32] The counterargument can be made that BCVs are, in fact, the simplest strategy since the process of adaptation requires navigating not only the morphological adaptation of each individual English verb but also the morpho-phonological changes necessary with borrowings (Hagerty 1996; Silva Corvalán 2001:288). By using a BCV, all these

Related to the idea of BCVs emerging to fill lexical gaps is the proposal that BCVs are used only with low frequency words but not high frequency ones such as *walk* (Reyes 1976), possibly because low frequency words might be harder to recall, as might also be the case with new words (Jenkins 2003:198). Three studies tested this hypothesis by examining the frequency of the English-origin item in the BCV. The first study found that English-origin verbs, defined as high frequency according to their Spanish equivalents (e.g. *eat 'comer'*), occurred at very low rates (4.5%) in BCVs in Belizean Spanish (Fuller Medina 2005a,b), thus supporting Reyes' (1976) claim. However, two later studies, which operationalized frequency following Fuller Medina (2005), reported a much higher percentage of high frequency verbs in BCVs, though they found a similar tendency of BCVs occurring primarily with low frequency verbs (Wilson 2013; Balam et al. 2014)[33].

It is apparent that a deficiency hypothesis or monolingual bias continues to prevail in the literature (Wilson 2013; Gardner Chloros 1995) as most of the discussion around why bilinguals use BCVs, such as lexical gaps and cognitive load, appear rooted in this position. The idea that bilinguals have a heavier cognitive load is a widespread one, often being exploited to account for

mechanisms can be said to be simplified since the English stem requires no changes and the inflectional endings need only be known for *hacer.* The issue of simple versus complex (or "cumbersome") is not an easy one to address and not one that the variationist approach wrestles with since this approach is less concerned with why speakers mix languages.

[33] Frequency was not established in any of the studies relative to the corpus being analyzed, i.e., not according to the varieties under investigation but rather according to lexical frequencies of the Spanish equivalent as found in the Corpus del Español (Davies 2002). Reyes' (1976) hypothesis was originally with regard to lexical frequency in Chicano Spanish. Wilson operationalized frequency by using *walk*, cited by Reyes (1976) as a high frequency verb that would not occur in a BCV in Chicano Spanish, as a measure of high frequency so that any verbs that had the same frequency as the Spanish equivalent *caminar* in the Davies corpus were considered high frequency. Following Fuller Medina (2005b), Balam et al. (2014:249) operationalized high frequency words as those that had more than 2000 occurrences in all conjugated forms in the Corpus del Español (Davies 2002). The frequency criteria used are problematic for a number of reasons, particularly from a variationist perspective, and since the former study did not use the same criteria as the latter ones, the results are difficult to interpret even if the same tendency was found in all three studies. Word frequency would be best defined according to the varieties and communities under study (Poplack and Dion 2012) and not by a corpus that contains various genres and varieties of Spanish as in the case of the Corpus del Español (Davies 2002). This issue is further complicated by the bilingual nature of the data. Why should English-origin verbs be more easily accessed? And if frequency is related to lexical access, are the English equivalents to the Spanish verbs more frequent than the Spanish verbs or simply more available if speakers are English-dominant? The question of frequency is one that surfaces as a means to try to explain why bilinguals use BCVs and is not of interest here. Furthermore, given the challenges outlined here in operationalizing frequency, this question is not directly addressed in this thesis.

bilingual data in spite of limited empirical support. The data with regard to BCV use is a case in point. It is more likely that speakers use BCVs simply because they can, as part of their individual and community repertoire (Poplack 1987) and not necessarily as an attempt to address the presumed cognitive load associated with being bilingual.

Moving away from proposals on why BCVs are used, aspectual constraints and the question of BCVs as code-switches or borrowings are now discussed. Some languages use a *do*-verb for more dynamic verbal borrowings and a second verb for stative, or less dynamic, verbal borrowings via BCVs (Muysken 2000; Romaine 1989, 1986). In the Spanish-English literature, few of the studies present distributional data, however, the BCVs that have been reported in the have tended to be with dynamic verbs (see also Jenkins 2003:202-203). Silva-Corvalán (1994) further suggests a stative/dynamic split interacting with tense where the perfective/imperfective distinction disappears with stative verbs as they are marked with imperfect morphology while non-statives maintain the distinction and appear in BCVs in the preterite and [*estar* + participle] constructions in the imperfect[34]. Spanish does not appear to have a second verb for statives as in other languages, so it was unclear if this meant that statives could not be borrowed via BCVs.

The hypothesis that BCVs might be aspectually constrained has been tested by using judgment tasks in combination with production data from semi-directed interviews where participants were asked to describe a picture or tell a story about what might be happening in the picture (Fuller Medina 2005a,b)[35]. BCVs in the production data were categorically with dynamic verbs while judgment tasks showed no significant difference between stative and dynamic verbs. A later study (Balam et al. 2014), which followed the same mixed method of using production data in conjunction with judgment tasks, found a similar discrepancy in their results. The production data

---

[34] The dataset was quite small, with only two tokens of BCVs reported.
[35] Most of the pictures were designed to depict a possible BCV as reported in the literature (see Fuller Medina 2005b).

showed a strong tendency to use BCVs with non-stative verbs (95%), but the judgment tasks showed no difference in acceptability of stative/non-stative verbs in BCVs. These two studies bring to light some important methodological considerations. It is highly questionable whether paper and pencil judgment tasks can capture the interplay of Spanish, English, and Kriol which is a natural feature of mixed speech in Belize (Fuller Medina 2005b)[36]. In addition, the aforementioned studies did not clearly define and operationalize lexical aspect. This issue is addressed in the analysis of verbs in Chapter 4

The question of most interest to this thesis, however, is the status of the English-origin item in BCVs. They have been termed borrowings (Fuller Medina 2015, 2013, 2005a,b; Wichmann and Wohlgemuth 2008; Torres Cacoullos and Aaron 2003; Toribio 2001:212; Eze 1998; Hagerty 1996; Sankoff et al. 1990), code-switches (Balam et al. 2014; González-Vilbazo and López 2011; Nishimura 1995), or, alternatively, BCVs are considered to be some novel bilingual construction, such as an emergent hybrid (Wilson 2013), manifestations of a third grammar (Romaine 1995, 1986, Gardner-Chloros 2009; Edwards and Gardner-Chloros 2007), a kind of special bilingual syntax (Muysken 2000:219) or products of creolization (Balam et al. 2014; Gardner Chloros 2009; Pfaff 1979).

González-Vilbazo and López (2011), for example, state clear assumptions in analyzing code-switching, i.e., no special mechanisms are needed, items from two languages contribute to building linguistic structure, mixed speech is just another example of the language faculty, etc. At the same time, however, they seem to use code-switching as a cover term for mixed speech, and BCVs are thus considered to be code-switches (Ibid: 835, 842). They suggest that word-internal

---

[36] In an attempt to address this issue, Fuller Medina (2005b) combined the paper and pencil tests with an audio recording of sentences read by a native speaker. Based on the inconsistency between the results of the judgment task and the production data, Fuller Medina concluded that the judgment tasks were not likely an appropriate tool for studying BCVs in Belizean Spanish. The fact that Balam et al. (2014) potentially had a similar result seems to lend support to this conclusion.

switching is empirically real and that borrowing and word-internal code-switching can be distinguished (Ibid: 840), though they do not make it clear how Spanish-German *utilizieran* ('to utilize'), considered to be word-internal switching, is different from Spanish-English *cliquear* ('to click') which they consider to be a borrowing. Likewise, Balam et al. (2014:256-257), citing Sebba's (1998) congruence approach, consider BCVs to be code-switches where *hacer* is used as a way to resolve the "switch incompatibility between the two languages" and that BCVs represent a switch site *created* by speakers[37]. Neither study provides independent measures of what constitutes a code-switch. Similarly, other authors consider BCVs to be borrowings (Fuller Medina 2005, 2005b; Toribio 2001; Hagerty 1996) but, again, do not provide a clear rationale for their decision. Hagerty (1996) includes them in his analysis of loanwords, and Toribio (2001:212, fn16), citing cross-linguistic evidence, considers BCVs to be "indicative of a general process of word-formation."

The idea that some third system is responsible for the creation of BCVs is also lacking strong empirical support. It has been claimed, for example, that BCVs are hybrid constructions found in various language pairs that do not have native counterpart constructions on which the BCVs can be based (Wilson 2013; Edwards and Gardner-Chloros 2007). Additionally, Muysken (2000:219) claims that BCVs must result from a special syntax since when these appear in language pairs involving Indic languages, the leftmost element is verbal, which is not permitted in monolingual verbal compounding in Indic varieties (Tamil, Punjabi, and Hindustani). However, Sankoff et al. (1990:77) argue that in the case of monolingual Tamil, the helping verb *PaNNu*[38] 'do' is a lexical transitive verb which takes an NP as an object, but in the mixed Tamil-English data it

---

[37] It is unclear what the authors mean by "switch incompatibility" as both the sites where BCVs appear and the boundary between *hacer* and its complement are sites where both grammars coincide or overlap, i.e., are sites of equivalence.

[38] Normal-aove **people** *vantu* **generalize** *paNNuvaa.*
      (adv) (emph)   (filler)   do (3p-pl-fut)
  'Normally people will generalize.' (287/M).
Transcription conventions for Tamil are retained from Sankoff et al. (1990:96).

acts as a pro verb and not a lexical verb. Thus, there may be differences in bilingual data in how the helping verb behaves and the type of complement it accepts. It is precisely in cases like this that the tools of the variationist method are helpful. When vernacular data is used as a benchmark, rather than grammars or standard varieties, the variability present in language as it is used. The of *PaNNU* 'do,' is a case in point [39]. In the same fashion, despite Wilson's (2013) claim to the contrary, light verb constructions are well-documented in monolingual Spanish (Bosque 2001) and may be seen as analogous to BCVs (Torres Cacoullos and Aaron 2003).

According to Muysken's typology, BCVs fall under both his categories of alternation and congruent lexicalization. He categorizes them as alternation if they show adjunction and as congruent lexicalization if they are formed from a language pair that shares [particle + verb] constructions (2000:185). Though, as noted, he does consider that a "special kind of bilingual syntax" underlies the formation of BCVs (Ibid: 219). Romaine (1989, 1986), who presents a systematic analysis of BCVs in Punjabi-English, within a Government and Binding approach, suggests that BCVs constitute a third grammar[40].

The majority of the studies discussed above have some common denominators. The decision to consider BCVs as code-switches, borrowings, or some other type of formation tends to be: (i) ad hoc as there are seldom independent criteria provided; (ii) based on single examples or sparse data rather than regular community-based patterns; (iii) based on vague definitions of code-switching and borrowing; or (iv) post-hoc rationalizations. Because the English-origin items in

---

[39] Sankoff et al. (1990) also found that Tamil, considered to be strictly OV also exhibits VO order, though in a limited fashion. Thus, the inherent variability in language is not always captured in formal grammars.

[40] While this is ultimately Romaine's (1989:139) conclusion, her analysis actually suggests a borrowing analysis (though she does not claim this), and she does lend credence to the borrowing analysis presented by Sankoff et al. (1990) as well as to DiSciullo et al. (1982). For example, the Punjabi gender agreement rule that is applied by analogy to English-origin nouns which are objects of BCVs suggests that there is no code-switch between the *do*-verb and its complement into English and, instead, that both the English verb and the nominal object of the BCV are treated according to the Punjabi system (Romaine 1989:140). Consequently, the evidence of integration into Punjabi suggests that they are both borrowings.

BCVs carry no affixed overt morphology, they are sometimes assumed to be bare and sometimes infinitives. A principled quantitative examination can provide much needed insight into this issue particularly because of its relevance to modeling language mixing.

At this point, it is worth providing some details of three variationist studies which include an analysis of BCVs, though they do not have this construction as the main focus of study. In these studies, all other language material was extracted and analyzed to identify regular patterns in the data. Additionally, clear definitions for code-switching and borrowing based on level of integration into the host language are provided. The first study analyzes Igbo-English language mixing (Eze 1998, 1997) where a subset of the bare English-origin verb forms was found to co-occur with an Igbo helping verb in a structure similar to that found in monolingual Igbo verb formations. As a result, it was concluded that these English-origin verbs, even though they were lacking Igbo morphology, were, in fact, integrated syntactically into Igbo and thus analyzed as borrowings as per the criteria of integration into the host language. In a second study on Tamil-English data, Sankoff et al. (1990) found that bare English-origin items that occurred in *PaNNu* constructions were syntactically integrated into Tamil via the pro-verb *PaNNu* and were being treated as native Tamil forms. In a third study, Torres Cacoullos and Aaron (2003) report that a number of bare nouns appearing in conjunction with *hacer* were, in fact, predicate nouns and thus verbal in nature. They behaved similarly to monolingual Spanish predication (but differently from English) and were thus integrated into Spanish and, therefore, not single-word code-switches.

2.5 Summary

A number of issues in the literature are lacking resolution. First, there is inconsistency in terminology, perhaps due in part to the second, more significant issue which is whether or not the surface manifestations of language mixing can be reliably identified. Related to this issue is the

status and function of BCVs as they are often considered code-switches, borrowings, or new bilingual innovations. A third concern relates to grammatical constraints on intrasentential code-switching. Nonetheless, as the preceding discussion has shown, a number variationist studies show code-switching and borrowing to be distinct, they can be reliably distinguished, and that they should be distinguished. One result from these studies, which aids greatly in this endeavour, is the finding that borrowing is an abrupt morphosyntactic process, independent of social and phonological integration. While the approaches discussed in this chapter may each offer a particular view of language, the variationist method (Labov 1972) offers tools to systematically analyze large datasets of naturalistic or vernacular speech. Surface manifestations of language mixing are identified by establishing *a priori* definitions of code-switching and borrowing to set the yardstick against which data whose status is unclear can be assessed. In addition, the communities' own vernaculars are used as benchmarks, and, thus, this method includes rather than eschews the variability found in natural speech. This thesis makes use of these methodological and conceptual tools to analyze all the surface manifestations of language mixing in the data and subsequently to reveal the community-specific patterns of language mixing employed by the speakers in the cohort analyzed.

CHAPTER 3 METHOD AND DATA

3.1 Introduction

The current chapter presents an overview of the study context, Belize, with brief descriptions of the history of Spanish, English, Kriol and indigenous languages in the country. Also discussed are features of the Belizean context that make it ideal for the study of language mixing. This is followed by a detailed description of the framework and application of the variationist method in this study, including the challenges of applying this method to language mixing. A full account of the data collection procedures is provided, as well as a description of the corpus and the subset of data analyzed for the current study. The final section describes coding procedures for the preliminary distributional analysis.

3.2 Belize

Belize is located on the Atlantic Coast of Central America, sharing borders with Mexico to the north and Guatemala to west and south of the country. The country is politically divided into six districts (Figure 1). The six administrative districts are comprised of the northernmost district of Corozal followed by Orange Walk and Belize in the central coastal area. To the west, bordering Guatemala, is the Cayo district, and Toledo and Stann Creek are in the south. In addition, several cayes[41] can be found along the coast.

---

[41] Cayes are islands off the coast of Belize. They are part of the territory of Belize and the northern cayes are politically part of the Belize District.

Figure 1: Political Map of Belize

According to the Statistical Institute of Belize (2013:6), the results of the 2010 national census indicate a total population of 324,453. Approximately half the population lives in urban centres and the other half in rural areas, as indicated in Table 1. The urban centres correspond roughly to the district capitals with the exception of Belmopan, which is the national capital[42]. The largest urban centre is the old capital, Belize City, followed by Orange Walk Town, San Ignacio/Santa Elena, and Corozal. It is important to note that the Statistical Institute of Belize

---

[42] Belmopan and San Ignacio/Santa Elena are both in the Cayo district.

(hereafter, SIB) considers all these communities to be urban centres despite the large disparity in population figures between them. Belize City, for example, is the largest urban centre and has more than five times the population of Corozal Town. Nonetheless, the Spanish-speaking communities such as Orange Walk and Corozal, for example, have similar populations; therefore, the SIB's definition of these centers as urban is maintained for the purposes of this thesis. Smaller communities that are not the district capitals are considered rural.

Table 1: Population by rural and urban distribution

Division	Population	
	N	%*
RURAL	145 832	45
URBAN	176 621	55
Belize City	57 169	17
San Ignacio/Santa Elena	17 878	6
Belmopan	13 939	4
Orange Walk Town	13 708	4
Corozal Town	10 287	3
Dangriga	9 593	3
Punta Gorda	5351	2

*percent of total household population 322,453

The main ethnic groups as indicated by the 2010 Census (SIB 2013), are Kriol, Mestizo/Spanish/Latino, Maya, Garinagu[43], Mennonite, East Asian, and South Asian, with Mestizos forming the largest ethnic group. The languages spoken in Belize correspond to a large extent to these groups as indicated in Table 2. The largest linguistic groups are speakers of English, Spanish, and Belize Kriol, an English-lexified Kriol[44]. The high number of persons claiming Kriol as

---

[43] *Garinagu* refers to the people and *Garifuna* to the language.
[44] I adopt the term Belize Kriol to refer to the English lexified Creole spoken in Belize, in keeping with the National Kriol Council http://www.nationalkriolcouncil.org/the_culture. (also Udz 2012 p.c.)

the language they currently use supports the view of Kriol as the lingua franca (Escure 2013a,b;

Udz 2012; Young 1995). While English is the official language of Belize and is the language of

government and the educational system, only 63% of the population report English as one of the

languages they speak. Furthermore, the census cautions that "the possibility exists that some

respondents who indicated they speak English might in fact have been referring to Creole" (SIB

2013:21)[45].

Table 2: Language spoken by persons over the age of four (Adapted from *Table 9* SIB 2013:21)[46]

Language	N	% total population
Spanish	165,296	56.6
Creole (Kriol)	130, 467	44.6
English	183, 903	62.9
Maya (Ketchí, Mopan, Yucatec)	30, 748	10.5
German	9, 364	3.2
Garifuna	8, 442	2.9
Chinese	2600	0.9
Other/not stated/applicable	4266	1.5
TOTAL	292,263	

When the largest linguistic groups are analyzed by district, as presented in Table 3 below, it

becomes evident that significant numbers of English, Kriol, and Spanish speakers are found in the

same geographic space, furnishing the conditions necessary for language contact and stable

bilingualism. Stable bilingualism can be said to have geographic, diachronic, and usage criteria.

First, it is characterized by the coexistence of the linguistic groups in question in the same

[45] Self-reports can be subject to the impact of social factors, such as prestige associated with the language (Romaine 1989:25).
[46] Totals do not add up to the country population totals as speakers could indicate more than one language which suggests bilingualism and multilingualism.

geographic space (Silva-Corvalán 1989; Lewis 1978, 1972:275). Second, this coexistence must take place over an extended period of time (Gómez Molina and Gómez Devis 2004:35; Lewis 1972:275). The languages in question must be used frequently and/or in multiple domains by the same speakers (Lewis 1972), and no one language is in the process of being replaced by another (Lewis 1972)[47]. All of these criteria are met in the context of Belize. As previously noted, Spanish-English-Kriol contact can be traced back to at least the 1850's, and these languages have coexisted within the same geographic space, with varying degrees of contact, over this time period. Finally, the census data, discussed in further detail below, indicates high numbers of speakers of these languages.

When census data on language is broken down by district, as indicated in Table 3, specific areas of the country are revealed to be Spanish-dominant. In the Orange Walk district, 86% of people indicate that they speak Spanish; this is followed by the Corozal district with 85% and the Cayo district with 72%. These are the regions of the country where Spanish predominates in contrast to the other three districts where less than 40% of people report that they speak Spanish. The Belize district is English/Kriol dominant and Stann Creek is Kriol dominant. Not shown here is that 68% of the people living in Toledo report a Mayan language as the language they speak, and thus, it is a Maya dominant region. Given the high percentages of those that report Spanish and English (and Kriol in the case of the Cayo district) as a language they speak, and the fact that respondents clearly indicate that they speak more than one language there must be overlap between these groups and significant bi/trilingualism amongst speakers can be inferred from

---

[47] In addition to stable bilingualism, Lewis (1972:276-277) categorizes bilingualism into three additional categories: dynamic bilingualism (characterized by a "greater tendency towards assimilation"), transitional (languages share overlapping function), and vestigial (almost complete assimilation and bilingualism is symbolic). In other terms, Gómez Molina and Gómez Devis (2004:35) consider bilingualism to be 'intense' if it takes place over an extended period of time and 'extensive' if it is widespread and has a high number of speakers, though they are not specific about how long or how many speakers would reflect 'intensive' and 'extensive' bilingualism.

census data in these districts. It is these very regions of Belize that were targeted for the current study.

Table 3: Language spoken by persons over the age of four (percent of total district population). (Adapted from Table 9 SIB 2013:21)

	Percent of district population speaking language					
Language	Corozal	Orange Walk	Cayo	Belize	Stann Creek	Toledo
Spanish	84.7	85.6	71.5	34.1	39.3	28.2
English	54.4	62.2	66.7	72.5	52.0	47.9
Creole (Kriol)	18.9	16.8	39.9	63.6	67.4	47.2

The census data are presented here as a guide to understanding the linguistic landscape of Belize as it relates to the current study. It also provides independent criteria supporting the decision to target the northern and western districts of Belize for data collection. The census data does, in fact, reflect what is commonly held to be true in Belize with regard to language and ethnicity, i.e., that Spanish speakers traditionally reside in the northern and western regions of Belize. However, certain caveats should be noted. The SIB does not report data for bilingualism or multilingualism[48], for example, nor does it report which variety of Spanish people speak. Of the 56.6% of people in Belize who speak Spanish, it is not indicated in the census data how many speak one of the two Belizean varieties of Spanish or how many speak one of the other regional (more recent) varieties such as Guatemalan, Salvadorean, Honduran, or Mexican Spanish. Notwithstanding these limitations, it is possible to gain an overview of the study context and see clearly that the conditions for stable bilingualism are present.

---

[48] An issue also noted by Udz (2012:192). See also Romaine (1989:25-28) for a discussion of some problems with census data, in particular with regard to bilingualism.

3.2.1 Belizean English

English and English-lexified Creoles are spoken in various pockets in the Central American Region (Holm 1983), but Belize is the only Central American nation that has English as its official language. This came about due to the British establishing settlements, and later a colony, in what was to become Belize. British settlements were, in fact, the first documented settlements of the European invaders in the region now known as Belize; this, despite the fact that Spain had laid claim to that territory. These settlements were not authorized by the British Crown, however, as they consisted largely of British buccaneers[49] who would lay in wait to rob Spanish flotillas of their wood cargo (Shoman 1995:19). The buccaneers later decided to cut their own logwood and mahogany[50] rather than rob the flotillas and, as a result, established settlements and 1763 Spain granted permission for woodcutting (Ibid: 173).

Spain and Britain eventually entered into a conflict over territorial rights, which culminated in the Battle of St. George's Caye in 1798. This conflict has never been fully settled, leading to a contemporary long-standing claim by Guatemala to the territory encompassing Belize (Mayr 2014; Shoman 1995:78). Nonetheless, 1798 serves as the decisive moment in the history of the region indicating how and why Belize was eventually established as an English-speaking nation in a region dominated by the Spanish language. In 1862, the colony of British Honduras was established; the name was later changed to Belize and, in 1981, independence was obtained from Great Britain (Mayr 2014; Shoman 1995). English remains the official language and is the language of media, government, education, and a second language for many.

---

[49] Originally a term applied to French and British Europeans who settled in the first half of the 16th Century in Hispaniola and thought to mean "user of bucan," a wooden frame used to smoke meat (Mayr 2014:72-73). It is a term that became synonymous with pirate, though Mayr (2014:72-73) distinguishes them from privateers and pirates. See also Latimer (2009) who defines buccaneers as privateers who were licensed by the British to attack the Spanish.
[50] Logwood and mahogany are types of wood.

3.2.2 Belize Kriol

Belize Kriol is an English-lexified Creole considered to be the lingua franca in Belize (Escure 2013a,b; Udz 2012; Decker 2005; Le Page and Tabouret-Keller 1985). It coexists with English in what can be described as a diglossic relationship (Ferguson 1959), where the language of prestige is used in formal settings while the language of "low prestige" is used in non-formal settings. Thus, despite the growing appreciation for Belize Kriol and the concomitant lessening of social stigma, Kriol is still largely reserved for informal contexts while English is used in formal and official contexts. While Kriol has covert prestige it remains highly stigmatized and considered as "bad English" by many (Udz 2012; Hoebens 2000; Young 1995; Woods 1991; Le-Page and Tabouret-Keller 1985). At the same time, attitudes towards Kriol have changed considerably, particularly among younger people, no doubt as a result of the work of the National Kriol Council over the past 20 years (Udz 2012). Furthermore, Kriol has been said to be gaining ground linguistically (Escure 2013b) and has even been noted to be a contributing factor in language loss in some Garinagu communities (Ravindranath 2009; Joseph 2003; Bonner 2001) which reflects the covert prestige it holds as well as status as a lingua franca and marker of national identity (Udz 2012; LePage and Tabouret Keller 1985). In fact, Kriol appears to be moving toward establishing itself as a marker of Belizean identity rather than solely the language of those speakers who are ethnically Creole (Udz 2012:201).

The main lexifier language for Belize Kriol is English. West African Languages which are considered to have contributed grammatical and lexical features to Belize Kriol include Akan, Bantu (Escure 2013b), Yoruba, Twi, and Igbo (Young 1995), as well as the indigenous Miskito language of the region, though the Miskito influence may be limited to lexical items (Young 1995). Because slaves were often brought through Jamaica, and Belize had an administrative relationship with Jamaica, it is also believed that there is a close relationship between Jamaican Creole and Belize

Kriol. Young (1995:34-37), for example, cites an overlap of about 2000 lexical items between Belize Kriol and those listed in the Dictionary of Jamaican English (Cassidy and LePage 1967)[51].

As noted in Table 2, Belize Kriol ranks third among the most commonly spoken languages in Belize, and, as such, is one of the main languages that is in contact with Spanish. Some speakers in the current study claim Belize Kriol as one of the languages they speak, and some even claim it as their dominant language. This is also evidenced in the data, as will be seen in Chapter 6, where some of the code-switching data contain Kriol features or full clauses in Kriol. Nonetheless, I did not consider the Kriol data substantive enough, in comparison to the monolingual Spanish and English data, to carry out a separate quantitative analysis which could establish the underlying grammar of Belize Kriol. Thus, a Kriol comparison dataset is not employed at this time. Kriol features are noted where they appear in multiword fragments and where single items (e.g. *bway*, 'boy') in Spanish discourse are uncontroversially Kriol. In addition, when speakers use single items in Spanish discourse that are etymologically English, it is next to impossible to determine if they are being taken from Belizean English or Belize Kriol by surface inspection alone. Consequently, they are referred to as *English-origin items*[52]. A full comparative analysis, along with the development of the necessary diagnostics, to determine if the etymologically English items are constrained in any way by Kriol grammar, and therefore Kriol rather than English, is left for future research. Assessing whether or not English-origin material is integrated in Spanish is still possible, however, given that both English and Kriol differ from Spanish in terms of diagnostics such as inflectional morphology on verbs and gender agreement on nouns.

---

[51] Details of the grammar of Belize Kriol can be found in Escure (2013a,b), Decker (2005), Hoebens (2000), Greene (1999), and Young (1995, 1973).

[52] This also raises the question as to the discreteness of English and English-lexified Creoles, an issue which remains unresolved in the literature, as well as the question of whether or not speakers have separate lexicons. These issues are beyond the scope of this thesis.

3.2.3 Spanish in Belize

Spain had established a presence in the region now known as Central America as early as the 16th century but never created settlements in Belize; thus, the presence of the Spanish language in Belize is not due to Spanish colonization as is the case for most of Central America. Spanish in Belize has historically come from two main sources, Mexico to the north and Guatemala to the west[53]. Cardona Ramírez (2010:27,46) identifies a third Spanish-speaking zone in the south, corresponding roughly to the Toledo district where "encontramos, durante el trabajo de campo, una cantidad considerable de hispanohablantes"[54]. Very little is known of Spanish in this southern region, and further research might reveal it to be a third variety with a similar historical presence in Belize to the two varieties under study here. Shoman (1995:88) for example, cites migrations in the late 1800s into southern Belize from other Central American countries. This suggests that there may be an additional community of stable bilinguals that has yet to be documented, thus highlighting the need for further research on Spanish in Belize.

Contemporary migrations from the neighboring Central American countries dating back to the 1980s have also taken place (Shoman 1995; Young 1995: Palacio 1990), which has led to an influx of new varieties. But, as with the Spanish in southern Belize, however, little or nothing has been documented regarding the impact, if any, of these new varieties on Belizean Spanish[55]. In the current study, these contemporary input varieties are not considered in the definition of Belizean

---

[53] These are considered the main input varieties, though it is not categorical. Some participants in the west, for example, claim that parents or grandparents came from Mexico to work in the Chicle industry (extraction of resin from Chicle trees). This is also confirmed by Pinelo (2015) who describes Mexican revolutionaries and counterrevolutionaries migrating to Belize, though information is not provided on how large these groups may have been.

[54] "During the fieldwork we found a considerable number of Spanish speakers"

[55] One can imagine that effects such language revitalization, dialect leveling, or a reduced frequency of language mixing might be possible outcomes, but these have yet to be studied. At this time, it cannot be confirmed whether or not this influx has had an effect on Belizean varieties of Spanish since Spanish in Belize remains largely undescribed. The current corpus and this thesis starts to fill this gap and provide a basis for future research. In addition, it remains an empirical question whether or not the populations that migrated to Belize were sizeable enough to influence local varieties.

Spanish[56]. Therefore, for the purposes of this thesis, the definition of Belizean Spanish is based on: (i) the historical record where input varieties can be traced vis-à-vis major migration patterns, as detailed below, and (ii) on Hagerty's (1979) phonological analysis which identifies two major varieties of Spanish in Belize. Belizean Spanish is defined as one of two varieties that resulted from two specific historical events in the 19th century. The first is the large-scale migration of the Mestizo and Mayan people fleeing the Caste War in Yucatan in 1850 who then settled in the northern region of the country (Mayr 2014; 219, 225-228; Shoman 1995: 86). This population brought with them the input variety for what is now the Spanish spoken in the north: Northern Belizean Spanish. The second historical event that shaped the place of Spanish in Belize is the migration from Guatemala in the 1860s, also of Spanish-speaking and Maya-speaking people (Shoman 1995: 87-88; Hagerty 1979:22). The migration from Guatemala stemmed to some extent from political unrest, but economic factors also played a major role as the Chicle Industry[57] attracted workers from both Guatemala and Mexico (Shoman 1995:87; Hagerty 1979:22). As early as 1839, workers from Guatemala were documented in mahogany camps in Belize (Shoman 1995:87). They brought with them the input variety or varieties for the Spanish which is now spoken in the west of Belize: Western Belizean Spanish[58].

It is harder to date and detail migration patterns from Guatemala, but migration from Mexico is better documented. The Caste War was a specific and significant event which led to large-scale migration in a very short period of time; in addition, data from the 1861 census clearly indicates that a significant portion of the population of Belize was from Mexico. According to this census, 38% of the total population of Belize at that time were born in Mexico while 9% were born

---

[56] A detailed description of the criteria for speakers of the targeted varieties of Belizean Spanish is outlined in *Data Collection*.

[57] *Chicle* refers to the resin that is extracted from the Chicle tree which was used in the manufacture of chewing gum. Workers in this industry were known as *chicleros*.

[58] There is some evidence that migration into western Belize continued well into the first decade of the 20th century, though probably not in the same numbers (Pinelo 2015).

in Central America, though it is not clear to what extent Guatemalans accounts for this 9% of Central Americans (Dobson 1973 and Barker n.d. as cited in Palacio 1990:12). Thus, detailed information on founding populations from Guatemala is not as readily available.

While the documentation on founding populations and their corresponding input varieties provide an essential rationale for considering Belizean Spanish to be composed of two main varieties, a second reason for defining Belizean Spanish as two distinct varieties can be found at the level of phonology (Cardona Ramírez 2010; Hagerty 1979). Hagerty (1979) considers Belize to be on an isogloss between Yucatecan Spanish and Guatemalan Spanish with one of the major phonological distinctions between the two varieties being the production of rhotics. Spanish has two rhotics, a tap /ɾ/ and trill /r/ which contrast in intervocalic position (Navarro Tomás 1918) (*pero~perro*, [peɾo]~[pero] but~dog). In non-contrastive contexts, the trill occurs word-initially, in prenuclear position, and word-internally following the consonants *n, l, s*. It may also occur when in coda position and followed by a consonant (Hualde 2013; Harris 1969). The tap occurs intervocalically and following consonants other than *n,l,s* but does not occur word-initially (Navarro Tomás 1918). The tap varies with the trill in coda position when followed by a consonant.

This tap /ɾ/-trill /r/ contrast is found in the west, while in the north a tap /ɾ/-retroflex approximant /ɻ/ contrast is prevalent[59]. This means that an English sounding 'r', or what has been called a retroflex approximant, is found in Northern Belizean Spanish in canonical positions where a trill would be found in mainstream varieties of Spanish and Western Belizean Spanish (Cardona Ramírez 2010; Hagerty 1979: 79-81). As a consequence, *perro* (dog) would be produced [pero] in the west but [peɻo] in the north. The retroflex approximant is also found in both prenuclear and

---

[59] I aim to capture the general contrast between the two regions with regard to rhotic production. For a more fine-grained analysis see Hagerty (1979:79-82) as well as Cardona Ramírez (2010:41-45) who present multiple variants within the general contrast indicated here and also discuss various processes of neutralization, assibilation, and so forth.

postnuclear position, and, as a result it is found word-initially, *ropa* [ɹopa] 'clothes,' as well as word-medially, *enredado* [enɹedado] 'tangled,' *perla* [peɹla] 'pearl'. This phonological difference is well-known amongst Spanish speakers in Belize and is one of the ways that northern speakers are identified in Belize. It is socially stigmatized regionally and to a lesser extent locally.

Another variant that appears to be geographically constrained and that may constitute a second distinguishing feature of these varieties, is the word final realization of the alveolar nasal [n]. In the west, the alveolar nasal tends to become velar, *pan* 'bread' [paŋ], which mirrors Guatemalan Spanish (Lipski 1994:184), the input variety for Spanish spoken in the west. In contrast, in the north, there seems to be some variability, with /n/ being realized word-finally mostly as [n] and minimally as [ŋ] (Cardona Ramírez 2010:40), consistent with Yucatecan Spanish (Michnowicz 2008; Lipski 2008:85) which is the input variety for Northern Belizean Spanish[60]. In contrast to rhotics, this variant would appear to be below the level of consciousness, as is the case in other varieties of Spanish (Lipski 2012:7).

Differences at the level of morphosyntax have yet to be systematically documented, however, one major difference is the use of verbal *voseo*. While Quilis (1990) reports an absence of *voseo* in Belize, Hagerty (1979: 13,102) notes its ample use in the west and its absence in the north, a finding confirmed by data from the current study. Full *voseo* (or pronominal *voseo*)[61] was not found in the current corpus, though anecdotal reports from participants in the west suggest that the variation between verbal and pronominal *voseo* may be present in the region.

---

[60] According to Cardona Ramírez (2010), Quilis did not report a velar nasal in his survey of Spanish in Belize conducted in the '80s. Consequently, Cardona Ramírez suggests that the velarization of word-final [n] may be a new pattern resulting from the major migrations in the '80s from Central America where this pattern is well-documented. However, the velar variant was documented earlier by Hagerty (1979:86), though its use was reported to be sporadic at the time. Further research is required to determine what the norms are for varieties of Spanish in Belize and any effects that reintroduction of this pattern may have had on the two varieties of Spanish spoken in Belize.

[61] *Voseo* refers to the use of the pronoun *vos,* 'you' to address an interlocutor. Corresponding verbal forms are also generally used (*vos sabés* versus *tú sabes,* 'you know'). When the verb forms are used without the pronoun *vos*, it is known as verbal voseo. The use of the verb forms in conjunction with the pronoun *vos* is known as full voseo (Fontanella de Weinberg 1999).

Other features of Spanish in Belize include a number of lexical items that can be attributed to English (in **boldface**) and Mayan languages (in CAPS).   Kriol features, such as the future temporal marker *a* are also evidenced in the discourse of some Spanish speakers. Examples drawn from the current corpus are presented below.

Kriol

12. | Para | mí, | mi | punto | de | vista, | John | *a* | top | the | polls |
     |------|-----|-----|-------|-----|--------|------|------|------|------|-------|
     | for | me | my | point | of | view | John | will | top | the | polls |

'for me, from my point of view, John. will top the polls again.'

(BSEK/ 45NJ/45NJ_43NJ_3/01:23:03/)

Mayan[62].

13. | Y | le | dan | su | mazorca, | su | PIBINAL[63.] |
    |---|-----|----------|------|----------|------|---------|
    | And<small>CL</small> | give 3.<small>PL</small> | | their | corn | their | PIBINAL |

'And they give them their corn, their PIBINAL.' {context: food served at the end of prayers}

(BSEK/54N/53N_54N/00:06:57.240)

English

14. | Pero | yo | lo | quería | **trayar** |
    |------|-----|-----|--------|--------|
    | but | I | it | want1.<small>SG.IMP</small> | try.<small>INF</small> |

'But I wanted to try it.'

(BSEK/58N/58N/11:28.0)

In addition, other features that have been documented in other Spanish-English contact situations are also found in Belizean Spanish. These are detailed below with the caveat that they have not been analyzed quantitatively, and therefore, it is not known at this stage if they are representative

---

[62] Quilis (1990) also reports *Chiich, chichi* (grandmother); *xix* (crumbs, a little bit) and *Tu'up* (youngest child) which is also expressed with the Spanish diminutive *–ita*: *tu'upita* These did not appear in the current corpus, but they are terms familiar to me as lexical items used in Belize.

[63] Pibinal: corn roasted in an underground pit oven. *Píib* – underground, *nal* -corn (Suárez Molina 1996) http://www.mayas.uady.mx/articulos/voces.html Accessed December 12, 2015.

of the communities' language use, if they are idiosyncrasies or anomalies, or even if they can be attributed to cross-linguistic influence as is claimed in some cases.

Third person auxiliary form used in contexts for first person auxiliaries

15. Pero   nunca, yo   nunca <u>ha</u>   <u>visto</u> nada
     But   never I   never have3.SG   seen nothing
     'But never, I have never seen anything.'

(BSEK/58N/58N/00:15:51.561)

This use has been documented in Spanish in the U.S. (Torres Cacoullos and Travis 2013; Bills and Vigil 2008; Lipski 2008) as well as in Belize by Hagerty (1996:137) who attributes it to simplification towards the English system. However, Lipski (2008:96) lists this feature as one typical of "rustic" or rural Spanish dialects and not limited to Spanish in contact with English.

Non-Standard Possessive Pronoun Use

Duplicated possessives

Duplicated possessives have been attributed to retentions from Old Spanish and are well documented in Mexican Spanish (Company Company 1991) and in other Latin American varieties (RAE 2010:349). In these constructions, a possessive pronoun (*mi* 'my') appears in the [Possessed N *de* Possessor N] possessive construction. As a result, there are two references to the possessor, the possessive pronoun, and the PP [*de* N], hence the term duplicated possessives. According to prescriptive norms, a definite determiner rather than a possessive pronoun would be required in possessive constructions such as the one exemplified in 45 below[64].

---

[64] Quilis (1990:144) notes this usage in Belize with personal pronouns: *su casa de él*, 'his house'.

16. fue en <u>su casa</u> de mi prima
    went3SG.PRET to POSS house of my cousin

'She went to my cousin's house.'

[indef DET + possessive pronoun + possessed NP] possessive constructions

The construction illustrated in (17), which marks possession with both the indefinite determiner *un* and possessive pronoun *mi*, has also been attributed to retention from Old Spanish (Lipski 2008: 159, 184) which permitted such constructions (RAE 2010:346). Pato (2002), however, argues for convergence between the Spanish construction and pre-existing parallel constructions in Mayan languages in Guatemala. These examples were far more frequent in interviews with western speakers.

17. el que estaba era <u>un mi</u> tío.
    One that be 3SG.LOC be3SG INDEF.DET POSS uncle

'The one that was there was my uncle.'

Another feature of possessive constructions is the use of possessive pronouns for inalienable possession where mainstream varieties of Spanish might use the definite article (Hagerty 1996). Because it mirrors English, this use has been attributed to English; however, Bullock and Toribio (2006) suggest that this use is not new to Spanish and therefore may not be due to contact with English.

18. tenía de todo color en <u>mi pelo</u>
    have 1SG.IMP of all color in POSS hair

'my hair was dyed all kinds of colours.'

A feature also found in Spanish in Belize is the use of *uno/una* in places where it may be deemed redundant, as is the case with demonstratives such as *este uno* which appears to be a kind of calquing of the English construction *this one* (Hagerty 1996:137). In (19) below, for example, *la otra* corresponds to English *other one* and *una* is not necessary in the Spanish construction.

19. Shit!   ¿cómo se   llama <u>la</u>   <u>otra</u>   <u>una</u>?
    Shit   how   CL   name the   other   one

   'Shit what was the other one's name?'

   (BSEK/ 45NJ/45NJ_43NJ_3/01:05:19)

Both gender and number agreement appear to be applied with these forms, thus, the paradigm includes: *otra una, otro uno, otras unas, otros unos* (Ibid). Whether or not this construction is due to contact with English cannot be confirmed at this time.

The two varieties of Spanish in Belize will be referred to as Northern Belizean Spanish (NBS) and Western Belizean Spanish (WBS)[65]. Historically the cayes have the same input as NBS spoken on the mainland and, for this reason, they are subsumed under the variety of NBS[66]. The designations north and west are broad and do not necessarily correspond to the political divisions in the country as illustrated in Figure 1 above**.** Likewise, the political division of the country does not correspond to the linguistic one. Nonetheless, it is commonly understood in Belize that the northern districts are Corozal and Orange Walk. For the purposes of this thesis, northern Belize corresponds roughly to Corozal, Orange Walk, as well as the cayes, and western Belize to the

---

[65] The Cayo district is sometimes considered to be a central zone or central-western zone (Cardona Ramírez 2010:45,46), but the Cayo district which corresponds to the zone for WBS is considered in Belize to be the western region of the country (e.g. Shoman 1995:139) and was originally called the Western District (Shoman 1995: 87); thus, the term Western Belizean Spanish is retained. This also has the advantage of distinguishing this region and variety from Cardona Ramírez's (2010:46) "Zone 0," which corresponds to the central coastal area and is not Spanish-speaking.

[66] Participants often report that Spanish is spoken differently in different communities. Some participants, for example, suggest that the Spanish spoken in the Cayes is different from mainland NBS, which may be possible due to the geographic isolation of the Cayes; however, until these differences can be empirically verified, for the time being the Spanish from the Cayes is subsumed under NBS and the two main distinctions of NBS and WBS are maintained.

western parts of the Cayo District. These regions then, are where we find NBS and WBS respectively.

3.2.4 Indigenous languages in Belize

As seen in Table 2, above, 10.5 % of the population claim one of three Mayan languages as on that they speak. These three languages are Yucatec Maya, spoken predominantly in the north and Ketchí and Mopan Maya spoken in the South. As noted above, Mayan words can be found in Belizean Spanish and at least three speakers from the north reported passive bilingualism in Yucatec Maya. According to Wilk and Chapin (1990) Yucatec Maya has had more contact with other linguistic groups in Belize as compared to Mopan and Ketchí and census data would appear to support this as 55% of Yucatec Maya speakers reside in the Spanish-dominant districts of Corozal and Orange Walk where (SIB 2013: Table L1.1). In comparison 77% of Ketchí speakers and 46% of Mopan speakers reside in the southern district of Toledo where 68% of the population report a Mayan language as one that they speak (SIB 2013: Table L1. 1; Table 2). Wilk and Chapin (1990) further suggest that Yucatec-Maya is being displaced by Spanish and English and according to the *UNESCO Atlas of the World's Languages in Danger* (Mosely 2010), it is a language that is "definitely endangered". "Ketchí" is considered "vulnerable" and Mopan is "severely endangered" though both these linguistic groups have more than double the number of speakers of Yucatec-Maya. The role of Kriol in language displacement of indigenous languages is not known nor is the extent of Spanish-Maya bilingualism. As noted earlier in the description of Belizean varieties of Spanish, Cardona Ramírez (2010) identified a third Spanish-speaking zone in southern Belize and little, if any, documentation exists regarding these communities of speakers, therefore, the level of contact with the Mayan languages and the role of Spanish in language shift or loss is also unknown.

The Garifuna language is also an indigenous language in the region. It is an Arawakan language spoken in Belize, Honduras, Guatemala, and Nicaragua and is considered to be "vulnerable" (Mosely 2010; Ravindranath 2009). Spanish-Garifuna bilingualism is more common in the other Central American countries while in Belize Garifuna-Kriol or Garifuna-English bilingualism is more likely. Of all the indigenous languages in Belize, Garifuna speakers are the smallest group and the vulnerable status of the language has been documented to some extent and is considered by some to be a "dying language" (Ravindranath 2009; Escure 2004; Bonner 2001). Belize Kriol is considered to be the language that is displacing Garifuna (Escure 2004; Joseph 2003; Bonner 2001) in contrast to the situation of Yucatec Maya. As with the Mayan languages, however, little is known as to the role of Spanish, in language shift or the extent of bilingualism or multilingualism among Garifuna speakers.

3.3 Sampling Methods of the Variationist Approach

The variationist method "places the locus of language in some sort of social order (the speech community) rather than the individual" (Figueroa 1994:70). This is "because the variable structures contained in language are determined by social functions, [therefore] idiolects do not provide the basis for self-contained or internally consistent grammars" (Weinreich, Labov and Herzog 1968:188). Thus, knowledge of these structures and the variable use of these structures constitute part of the linguistic competence of members of the speech community (Ibid). Once a community is identified, locating (and recruiting) participants who possess this linguistic competence becomes essential for the collection of appropriate data (Poplack and Meechan 1998:128; Poplack 1993: 262-263). Within the variationist approach, appropriate data is comprised of the spontaneous, vernacular speech of native speakers of the varieties in question from well-defined communities. In this way, data is assured from speakers who possess knowledge and

competence of both the grammatical and community norms of language use. The notion that the community norms play a role in the grammar of their respective communities is particularly important in bilingual communities where some modes of discourse, such as code-switching, may be highly stigmatized, and consequently may be eschewed even where the language pair permits it (Poplack 1993).

The current study utilizes the snowball technique (Milroy and Gordon: 2003:32) for recruiting and locating participants. This technique utilizes participants' social networks to recruit more participants. The researchers then attempt to continually tap into the participants' networks for the duration of fieldwork. This technique has the benefit of tapping into the speakers' immediate networks and thus enables data to be collected from a relatively cohesive group of participants within the speech community[67].

3.3.1 Accessing the vernacular

As noted above, appropriate data consists of the vernacular speech of native speakers of the varieties in question. In its simplest definition, vernacular speech is casual unmonitored speech (Labov 1984). Two additional components are crucial to this definition. First, the vernacular refers to the variety learned in preadolescence (assumed to remain stable once acquired), and second, the rules that govern the variability found in the vernacular have been observed to be more regular than more "formal" or "superposed" styles learned later in life" (Labov 1984:29); it is therefore considered the most systematic data and, thus, the most appropriate data for analysis[68].

---

[67] This may also have the unwanted result of a biased sample. In an effort to avoid biases in sampling, for example, random sampling may be employed (Tagliamonte 2006; Schilling-Estes 2013), and while this technique has been used successfully (Poplack et al. 1988; Poplack 1989), without the resources necessary for the compilation of a mega corpus, it can often lead to unmanageable corpora (Schilling-Estes 2013:33; see also Mallinson et al. 2014).

[68] In keeping with this assumption, pre-adolescent speakers were not included in the speaker sample though it should be noted that not all researchers agree with the assumption of language stability in post adolescence (see Schilling-Estes 2013).

For some researchers the vernacular is seen as an "abstract object" analogous to the standard, and thus elusive, making it difficult for the researcher to know when they have captured it (Becker 2013:96; Milroy and Gordon 2003:50). Along these same lines, others caution against the common assumption in received definitions of the vernacular that speakers have one true vernacular (Schilling Estes 2007:173; Eckert 2003). Speakers can style-shift, for example, between the vernacular and less vernacular forms even in casual speech, and thus, there exists a "multiplicity of vernaculars" (Poplack 1982:2-3). In other words, there are no single-style speakers (Labov 1984:30; see also Schilling Estes 2008, 2007). Therefore, it is perhaps most useful to think of vernacular data as data that closely approximates that variety learned in preadolescence, that reflects the internally consistent grammar of the community's language use, and that is used in casual speech. This type of data can be obtained by employing the tried-and-tested tools of variationist sociolinguistics.

Underlying this approach to data collection is the attention to speech model. In this model, casualness varies along a continuum of the amount of attention paid to speech (Labov 1984)[69]. This means that speakers often shift to more formal modes when they know they are being observed or when interacting in formal contexts. This leaves researchers with the conundrum whereby "the aim of linguistic research in the community must be to find out how people talk when they are not being systematically observed; yet we can only obtain these data by systematic observation" (Labov 1972:209). This phenomenon is known as the *observer's paradox* (Labov 1972). As a consequence, efforts to access the vernacular involve mitigating the observer's paradox by finding ways to draw participants' attention *away* from language and language use and facilitate the use of the vernacular (Schilling Estes 2013: 93, 96, 98). Traditionally, two main approaches are used. The first relates to having in-group members of the community collect data,

---

[69] See Schilling Estes (2013) and Meyerhoff et al. (2012) for further discussions of the attention to speech models.

and the second is for (preferably) in-group members to employ the sociolinguistic interview (Labov 1972) for data collection.

The interviewer's status in the community is considered important since speakers presumably pay more attention to speech, and consequently use more formal speech, with outsiders than with other community members or perceived insiders (Poplack 1981). As a result, some researchers insist on in-group members for entering the community and conducting interviews (Tagliamonte 2006; Torres Cacoullos and Aaron 2003; Poplack 1993:[70,71]). Poplack (1981) has shown that qualitative and quantitative differences arise, for example, when code-switching data was collected by an outsider versus when it was done by an insider. She asserts that this is because code-switching is an in-group mode of speaking, one that could not be accessed by an outsider. The insider/outsider dichotomy, however, is multifaceted[72], and some variationists consider the in-group membership requirement as extreme (Meyerhoff et al. 2012:127). Where a researcher is not an in-group member, various tools such as repeat visits to the community (Blake 1998; Moyer 1992), ethnographic methods (Eckert 1989), employing community fieldworkers (Poplack and Tagliamonte 1991; Poplack et al. 1988; Poplack 1987:66), or using outsider status as an advantage by adopting the role of student or learner (Hazen 2000) can be used to access data[73]. This thesis, in keeping with traditional variationist methods (Tagliamonte

---

[70] For a recent example of criteria for defining insiders or in-group members for the purposes of data collection and bilingual corpus compilation, see Torres Cacoullos and Travis (2014) and Travis and Torres Cacoullos (2013).

[71] How the researcher enters the community can also potentially affect which participants they gain access to and consequently the kind of data that may be collected (see Schilling Estes 2013; Tagliamonte 2006:22).

[72] It is perhaps most useful to think of the insider/outsider dichotomy as being on a continuum representing membership within the community, with the idealized in-group member being at one end of the continuum and the 'outsider' being at the other end. In this way, insiders and outsiders are always positioned relative to the speech community, leading to different levels of membership in a community. Ideally, the interviewer will be positioned closest to the insider-end of the continuum. The insider/outsider dichotomy is problematized far more in sociological and feminist research. For a detailed description from these perspectives see Gooden and Hackett (2012) on entering their own Caribbean communities. For recent treatments in sociolinguistics, see Levon (2013) and for general discussions, see Schilling Estes (2013).

[73] It is generally accepted that interviewers who share the same characteristics as their interviewees will not only enter the community more easily but also capture the most natural speech; however, it is not a guarantee (e.g. Rickford and

2006; Poplack 1993) privileges the status of the insider for data collection, and holds that the value of the role of insider cannot be underestimated. This is not only because an insider is most likely to share the same vernacular, but because their knowledge of the community permits them to formulate questions that are relevant to the community and even likely to elicit narratives (Torres Cacoullos and Travis 2013:7).

The second approach to drawing attention away from speech, and consequently mitigating the observer's paradox, is the sociolinguistic interview (Labov 1984). This is the standard tool utilized in variationist studies (Schilling-Estes 2013; Tagliamonte 2012, 2006; Poplack 1993; Labov 1984)[74]. This 'interview' initially consisted of four components designed to elicit a range of styles; it included: a conversation with the interviewer, reading a list of minimal pairs, words in isolation, and a short narrative (Meyerhoff et al. 2012:129; Labov 1984:32-33). Typically, the sociolinguistic interview now refers to the conversation component alone, which utilizes a series of modules designed to elicit natural speech by building from more general questions to more personal questions, often ending with questions about language (Tagliamonte 2012; 2006). The general questions allow participants to get used to the interview so that they are comfortable answering questions later in the interview that relate to more personal experiences. The topic, then, also plays a role in attention to speech since speakers likely pay least attention to speech when engrossed in the telling of personal experiences (Tagliamonte 2012, 2006; Labov 1984, 1972). This technique is intended to mitigate the observer's paradox.

The series of modules allow for the interviewer to introduce topics that will likely elicit narratives and avoid "yes/no" questions. Likewise, participants can talk as much or as little as they

McNair Knox 1994) and should be seen as a strong tendency rather than a given. Any researcher must always negotiate the insider/outsider dynamic and its impact on data collection (e.g. Friesner (2009) on being an outsider and potential impact on data).

[74] See Labov (1984) and references therein for more information of the development of the sociolinguistic interview.

want about a given topic and are free to introduce topics themselves (Tagliamonte 2006; Poplack 1993:261). The range of topics includes childhood games, community practices, and so forth. Consequently, questions can easily be tailored to the study context (Schilling Estes 2013:94). It is intended to be a conversation that focuses attention on the participants' stories so that attention is drawn away from their language use, thus facilitating the use of natural or vernacular speech. Other strategies of countering the observer's paradox are to conduct group interviews or to leave a recorder running (with permission) while two or more people talk, without necessarily guiding the conversation through the modules of the interview.

Since the sociolinguistic interview was first conceived, it has successfully been used countless times; however, a number of critiques have arisen, one of the most common being the 'danger of death' question where participants are asked if they were ever in a situation where they thought they were going to die. This question worked well in Labov's early work in New York, but many researchers have since found it less useful (Schilling Estes 2013; Meyerhoff et al. 2012; Milroy and Milroy 1978), and in the current study this question was rarely used[75]. A more important but less frequent criticism concerns the nature of the data (Schilling Estes 2008; Eckert 2003). It is important to remember, however, that one of the reasons the Sociolinguistic Interview continues to be used is that it permits the collection of large amounts of data in a short period of time (Schilling Estes 2007). It has been shown in a number of studies to be suited for the collection of naturalistic data, particularly if researchers take into consideration both that speech is shaped by "a range of contextual factors" and that the vernacular is an abstraction that is multidimensional (i.e., there is a range of casual styles), in which case, the utility of this tool is not diminished (Schilling Estes 2007:174).

---

[75] Instead with older participants in coastal areas, for example, I could ask about their experiences during tropical storms and hurricane threats.

3.4 Data Collection[76]

In this section I describe in detail how I apply the methodological tools above to the Belizean context, i.e., how I circumscribed the speech communities, located and recruited participants, and collected data via sociolinguistics interviews. Subsequently, the corpus is described with details provided as to the subset of interviews that are analyzed for the current study.

3.4.1 Locating speakers

According to the census data described above, Spanish is spoken mainly in the north and west of Belize. Data was collected in both these regions, following Hagerty (1979) and in keeping with an earlier study of Belizean Spanish (Fuller Medina 2005b). The rationale for collecting data in both regions of Belize is twofold. First, it aims to document Spanish in Belize since only a handful of studies have researched this variety of Spanish (Fuller Medina 2015, 2013, 2010, 2005a,b; Balam et al. 2014; Cardona Ramírez 2010; Hagerty 1996, 1979; Quilis 1990). While the focus is on language mixing among a select cohort of speakers, some characteristics of both local varieties are outlined, as noted above, and also provided are some indications as to the monolingual systems in the Spanish spoken in Belize. In Chapter 4, for example, in assessing the integration of English-origin nouns, mixed data is compared to monolingual Spanish data and it becomes evident that speakers employ gender and number agreement rules in monolingual Spanish just as speakers do in other varieties of Spanish. Likewise, in Chapter 5, the monolingual Spanish data offers insight into the clitic system in these varieties of Spanish setting the stage for a full quantitative analysis such as has been carried out by Silva Corvalán (1994) on Los Angeles

---

[76] Ethics approval was obtained from the University of Ottawa REB (File number: 06-13-37), as well as in Belize from the Institute of Social and Cultural Research (ISCR H/2/19)

Spanish. In order to build a complete picture of Spanish in Belize, it is particularly important to include both varieties in systematic analyses.

The second reason is that these two regions have been previously identified as distinct dialect regions, based on a detailed phonological study carried out by Hagerty (1979) (see also Cardona Ramírez 2010:77). Therefore, in order to understand patterns of language mixing in Belize data, both varieties must be analyzed, particularly since it is known that community norms may constrain language mixing, and each region could potentially have a distinct pattern of language mixing (Poplack 1993; Poplack et al. 1988; Poplack 1987). In Fuller Medina (2005b), for example, data on bilingual compound verbs (e.g. *hacer retire* 'to retire') was collected from both Spanish-speaking regions rather than just from one community, for this very reason. Bilingual Compound Verbs (BCVs) had not been previously studied in a systematic quantitative manner in the Spanish-English literature, consequently, the extent of their use was not known[78]. Data from Belizean varieties of Spanish showed that BCVs were used both in the north and in the west with no difference in the number of tokens produced by speakers in each region (15/34), though a smaller number of northern speakers produced the same number of tokens as a larger group of western speakers, suggesting a more pervasive use in the north, a result that could not have been found without casting a wide net (i.e., in both regions) (Fuller Medina 2005b: 90-91). It is also a result that can now be tested in this thesis with a larger dataset.

The communities in the two regions targeted for this study are recognized as predominantly Spanish-speaking. These are some of the same communities that had been identified in an earlier study (Fuller Medina 2005b) and that census data have shown to have high numbers of Spanish speakers. Speakers from these communities who speak the local varieties of

---

[77] Very little has been documented on other areas of grammar as is the case with language mixing in general.
[78] In the U.S., for example, it had been reported to be highly restricted to Southwest Spanish (Jenkins 2003).

Spanish as well as English and/or Kriol were targeted. As noted earlier, however, Belize has had significant immigration from neighboring Central American countries in the 20th century, representing a second major stage in the history of Spanish in Belize. These more recent varieties of Spanish have not been in contact with English and Kriol to the same extent as the earlier varieties from the north and west as described above (i.e., approximately 40 years or less versus 150 years of contact)[79]. Consequently, for the purposes of the current study, these speakers (first, second, or third generation) are not considered to be speakers of varieties of Belizean Spanish even if they may now be Belizean nationals and speak Spanish[80].

Who then is considered a Belizean Spanish-speaker? Following previous recruitment criteria (Fuller Medina 2005b), Belizean Spanish-speakers are defined as persons who currently use Spanish in their everyday lives, who grew up in Belize, whose parents/guardians also grew up or were born[81] in Belize (either in the north or the west), and who acquired Spanish at home. Any person who met these criteria but had also worked or studied abroad in a Spanish-speaking country was excluded when I further refined the corpus[82]. Participants for the study also had to fulfill the important criteria of using at least two of the languages in question regularly, i.e., Spanish and English, or Spanish and Kriol, as indicated by responses regarding language use in the language module of the interview. This information was also garnered from participant observation.

---

[79] While these populations may be sizeable enough to warrant study in terms of dialect contact, dialect change and convergence, they do not fall within the purview of this investigation and as such are not pursued further.

[80] This is in no way a comment on their belonging (Hall 1995:4), nationality, or identity.

[81] If a participant or parent was born in Mexico or Guatemala simply due to access to medical care, these participants were not excluded, as it is not uncommon, particularly in border communities, to access medical care in Guatemala or Mexico. If parents or participants were born in non-neighboring countries (El Salvador, Honduras, etc.), then these participants were excluded from the corpus.

[82] I chose to exclude these speakers on the basis of extended immersion in a monolingual variety of Spanish and the potential pressure to shift to the host variety. Some participants spoke openly about the linguistic discrimination they faced based on their local variety. The data itself also supported this decision when interviews were later reviewed. Their speech was highly monolingual, though they often claimed to engage in the language mixing typical of their community, suggesting non-vernacular usage in the interview. In addition, salient markers of Belizean Spanish, such as the retroflex approximant associated with NBS, tended to be absent for NBS speakers in this category, and finally, these speakers tended to indicate that they had actively worked on 'improving' their Spanish.

In terms of participant recruitment and entering communities, I was able to make use of my secondary networks to initially recruit participants, but most recruitment was done through the support of the Houses of Culture (detailed below). I was fortunate to benefit from these kinds of support offered to me as a result of being a local researcher. I also share a pan-Belizean identity with participants, as well as the shared cultural identity that comes along with being a native speaker of Spanish, English, and Kriol. In addition, language mixing is a part of my own vernacular speech. I had personal connections to each community because I had spent time as a child or teenager in each community and because I either had family ties in the community or was connected through friends from the community. Thus, participants were able to associate me not only with the research but with community members and the local community centres.

The Houses of Culture in each community played a central role in participant recruitment which was partially supplemented by contacts from my secondary networks. The Houses of Culture (hereafter HOC) are local community centers whose programs are centered on cultural activities and are generally well known in the community. In the west, the HOC plays a particularly central role in the community and acts as a meeting place and drop-in for community members of all ages. This HOC often provided space for interviews, introduced me to potential participants, and/or allowed me to participate in activities, which facilitated my meeting potential participants. Thus, each HOC became a kind of home base from which I could operate and any association with a community organization did not appear to impose limitations on data collection such as adding a more formal dimension to interviews when conducted at the HOC[83]. Data collection in the west coincided with community events hosted by the HOC, therefore, I attended community events and

---

[83] See Childs et al. (2001:167) for an example of how partnering with a small community's town council facilitated their research.

helped out with events where possible[84]. In the northern communities I was also able to participate in events, though my time there did not coincide with as many as in the west. As in the west, the HOCs in northern communities are very active and also aided greatly in participant recruitment and provided space for conducting a number of interviews.

I originally created recruitment flyers to be posted in the community at places like the HOCs, but, in fact, these were distributed instead at local festivals and directly to potential participants when they visited the HOC. As a result, many participants in the current study were those people who frequented the HOC regularly. This initial group of participants then put me in touch with other potential participants in their network. Consequently, most participants were recruited through "snowballing" (Milroy and Gordon 2003:32) which often had its starting point at one of the HOCs[85].

Recruitment through community centers also allowed for ethnographic observation, in particular in the west. Because the HOC in this region was highly active and had numerous people of various age groups coming in and out on a daily basis, and because I happened to be there for some community events, I was able to become further acquainted with language use in the community. I was also able to meet a number of people from the community in one central location and in a space that they frequented often and presumably felt comfortable. This often permitted me to have brief conversations with potential participants, note interactions at the center, and then on a

---

[84] I mention this participation because it ultimately enhanced the data collection process as it gave me an even better sense of the speech community and helped me adjust questions in the interviews. While this kind of participation is often touted as a "strategy" for entering a community, I am already part of the larger community and already had personal connections in the specific communities. Participation was genuine and personally motivated. Any positive exposure it could have given me is simply a fortunate coincidence. At the same time, it should be noted that if the efforts and interest were not genuine, this would have been easily detected and would have no doubt worked against me, particularly as an "insider."

[85] There was one northern community where speakers of the local variety under the age of 60 could not be found. This raises cause for concern, particularly since this community had historically been an isolated one and the lack of speakers of the local variety may signal a shift to English and/or Creole at the expense of Spanish. In addition, it raises the question of how a variety of Spanish might be lost in a region dominated by Spanish.

separate occasion interview them. In addition, being hosted by a family for two weeks in this same community gave me direct access to observing everyday language use and also led to a high level of interaction with any individuals living in the home or who were frequent visitors to the home. This meant that if any of these individuals participated in the study, they would have already become accustomed to interacting with me and seeing me with the recorder. This afforded me the option on a few occasions of simply leaving the recorder running (with permission) while two or three participants chatted.

In the fall of 2013, over a two-month period, I carried out sociolinguistic interviews with participants from a total of eight communities. I carried out additional interviews in two of these communities again in the summer of 2014. As a result of both field visits, a corpus of interviews with 51 speakers was compiled and is further described below. A large number of people volunteered to participate, particularly in the west, and I did not have "no-shows." I also adopted a policy of interviewing everyone who made themselves available to participate and completed interviews, albeit often shorter ones, even if they did not fit the criteria (e.g. studied abroad)[86].

Interviews generally lasted one hour, but in one case the interview was approximately half an hour[87], and five interviews ran between 1.5- to-3 hours. Length of the interview was always determined by the participants' time constraints. Four interviews were with groups of two or more[88]. Topics introduced for the interviews were generally about community life, childhood games and rules, and folktales and supernatural beings[89] but had to be tailored somewhat for different

---

[86] Despite screening criteria, it often did not become apparent until the interview was underway that speakers did not fit the criteria discussed above. Nonetheless, the "interview everyone" policy had positive outcomes even if some individual interviews were excluded from the corpus, since this resulted in getting to know more community members and ultimately recruiting more participants. In addition, these interviews now comprise a comparative dataset which can be analyzed for any number of features related to the acquisition of community norms, individual differences, and socially indexed variables.

[87] Despite the brevity of the interview, it nonetheless yielded 113 tokens of other-language material.

[88] I considered these joint interviews, and they are indicated with a "J".

[89] See also Hagerty (1979) on the use of folktales and accounts of supernatural beings.

communities. Questions about folktales and supernatural beings do not work well for eliciting narratives in larger urban centres, for example, just as questions about hurricanes are not the most relevant in non-coastal communities. I was also able to modify some more sensitive questions, such as those regarding level of education. Since I knew that many smaller communities never had a high school, with middle-aged and older speakers, for example, I could ask if there used to be a primary school or high school when they were growing up rather than asking directly about their level of education. When asked in this way, participants usually included in their response their own experience with school and level of education.

A language module was also included and came towards the end of the interview. In this component participants were asked to rate their language ability in Spanish, English, and Kriol and to indicate their preferred or dominant language, if they had one. They were also asked to indicate language use for each language, i.e., peer groups, school/work, and home (immediate or extended family as well). At the end of the interview, after debriefing, each participant in the study, a symbolic remuneration of $15BZD was offered. This was intended to to be a token of gratuity to thank participants for their time and to also cover any costs (e.g. travel)[90] incurred by participating in the study.

In keeping with the ethics requirements for this study, and in order to maintain anonymity and confidentiality, each participant was assigned an alphanumeric code in the corpus. This consisted of the interview number followed by an N or W to indicate the region they are from. The code 13W indicates that this was interview 13 and the participant is from the west and a speaker of WBS. The code *13WJ/11_12_13J/10:55:6* indicates a sample from participant 13 found in the audiofile *11_12_13J* (a group interview with participants 11 and 12) at the 10:55:06 minute mark.

---

[90] For example, three participants chose to travel to the community center in town from villages that had been amalgamated with the main town. Remuneration was more than enough to cover their travel cost. See Schreier (2013:27-28) for a critical discussion of remunerating participants in sociolinguistic research.

Consequently, the format for the identifier associated with each token is: participant#/audiofile/timestamp. All interviews that were conducted in 2014 additionally have the number '14' as part of the participant alphanumeric code. All interviews carried out in 2013/2014 are compiled in the Belizean Spanish, English, and Kriol Corpus (BSEK). Where names of people were used during the interview, they were replaced with initials or pseudonyms in any examples used here. Additionally, following Rickford (1986), I have also chosen not to name the specific communities and to refer only to the region[91].   In the following sections, the makeup of the corpus is described.

3.4.2 Nature of the corpus

Table 4-Table 6 present a summary profile of the participants interviewed for this project and who met the criteria of a Belizean Spanish-speaker for the purposes of this study. Most of the speakers in the west were younger speakers accounting for 55% of participants in this region, while in the north the reverse was true with 52% of speakers falling into the "older" category. In Table 5, most participants are grouped into three education categories. *Primary* includes participants that had some primary schooling or those who completed primary school. The same applies to *High School*, it is the maximum level of education completed or partially completed. The last category includes participants currently enrolled in *Sixth Form or Higher*[92], who had completed sixth form or who completed or partially completed university studies.

---

[91] These measures go beyond ethics requirements but are taken in order to account for the ways in which confidentiality and anonymity requirements differ between the Belizean and Canadian contexts as well as the consequences these might have for protecting the identity of the participants. This thesis must be deposited in Belize as a requirement of ethics approval from the Institute of Social and Cultural Research, and preliminary findings of this project have already been presented in Belize (Fuller Medina 2015, 2014); therefore, in trying to keep the best interests of the participants in mind, as I understand them, I chose to err on the side of caution. See Childs et al. (2011) on the challenges of navigating confidentiality and anonymity in small communities. Much of the same issues apply to Belize.
[92] Sixth form refers to two years of college that follow the four-year secondary program.

Table 4: Participants by AGE and REGION[93]

AGE	REGION					
	WEST		NORTH		TOTAL	
	N	%	N	%	N	%
Young (15-35)	16	55	7	29	23	45
Middle (36- 55)	5	17	4	20	9	18
Older (56- 95)	8	27	11	52	19	37
TOTAL	29	57	22	43	51	100

Table 5: Participants by EDUCATION and REGION

EDUCATION	REGION					
	WEST		NORTH		TOTAL	
	N	%	N	%	N	%
Primary	8	27	11	50	19	37
High School	8	27	6	27	14	27
Sixth Form or higher	13	45	5	23	18	35
TOTAL	29	57	22	43	51	100

Table 6: Participants by SEX/GENDER and REGION

Sex/Gender	REGION					
	WEST		NORTH		TOTAL	
	N	%	N	%	N	%
MALE	11	38	8	36	19	37
FEMALE	18	62	14	64	32	63
TOTAL	29	57	22	43	51	100

---

[93] These age groupings cover the range of ages in the corpus and are not intended to reflect any particular stage of language acquisition, development, or change, though they correspond to the age categories of young, middle, and older speakers in Poplack (1989); see also Tagliamonte (2013:47) on age grading and the range that is considered 'middle years.'

3.4.2.1 Nature of the data

The above-described interviews provided both highly monolingual data, from primarily rural older speakers and, more importantly, for the purposes of the current investigation, rich data of mixed speech. The selected interviews constitute a rich repository of vernacular usage as instantiated by personal narratives, code-switching and borrowing. Other features of casual speech included minimal response, *mhmm*, exemplified in (20) below where both the interviewer (indicated by "I") and the participant employ this feature[94] and, for northern speakers, the retroflex approximant appeared consistently in their speech[95].

20. Y que   de veras really mean   it, no? {I: **mhmm**}   **mhmm**, así
    'And really mean it, no? {I: mhmm}   mhmm, that's how I am.'

                                                              (BSEK/08W/08W_03/00:18:38/)

The presence of personal narratives, jokes, and laughter are also considered features of informal speech (Poplack and Tagliamonte (2001:76,81). This is illustrated in (21) below along with the rich language-mixing found in the data analyzed for the current study. Both single and multiword English-origin fragments are used in addition to Kriol in this account of a visit to a "fortune teller." English is in **boldface**, Kriol in *italics*, and interviewer speech is indicated with an "I" and enclosed in { } brackets.

---

[94] Other minimal responses include those indicating "yes", *ahaahn*, which is used more by Creole speakers, and which corresponds to Spanish *ajá*. Mhmm can mean "yes" or can just be an affirmation to include the speaker and punctuate discourse. These types of minimal responses are typically used in casual speech in Belize. While I indicate an example of my own use, this is one place where shared vernacular features was not an advantage and had to be monitored on my part so as to not overlap with the participants' speech, which is typically what would happen in spontaneous interactions.
[95] Recall that this variable is characteristic of Northern Belizean Spanish and one that is often manipulated by speakers who had worked and/or studied aboard.

Fortune teller:

21. **So, that**, ese parte si puedo decir que es cierto pero aparte de eso, {I: y fuiste aqui?} *Ahaahn*, pero después, ehm, no sabía si hacer **trust or** no, porque me hizo **freak-out** después. Porque entonces dije yo: **if that could be true, imagine** si me dice algo **negative** {laughter} lo que sea **good** me gusta escucharlo pero ya cuando sea algo **negative,** "*oh, no I no beleev ina dat*" {loud laughter} (BSEK/58N/04N14_58_N/01:00:48.285)[96]

'So that, that part I can say was true but apart from that {I: and you went here?} Yes, but then after, ehm, I wasn't sure if to trust it or not because it freaked me out afterwards. Because, then I said: if that could be true, imagine if she tells me something negative {laughter} whatever is good I like to hear it, but then when it's something negative, "Oh no I don't believe that". {loud laughter}'

Finally, as a result of the ethnographic observation discussed above, I was able to note other features of casual speech which I am familiar with given knowledge of the varieties spoken in Belize. I was able to verify that many of the same features, as described above, were also used by speakers when not being recorded [97].

3.4.2.2 Transcription procedures

Interviews were transcribed using the linguistic annotator tool ELAN (Sloetjes, H., & Wittenburg, P. 2008). Where partial extraction was completed, this was done using ELAN as well. In keeping with the variationist principle of studying language as it is used, interviews were transcribed verbatim with no adjustment of non-standard features to prescribed forms. In this way, features such as the possessive constructions described above could be captured, as well as all the language-mixing data[98]. In addition, false starts, incomplete utterances, and so forth, often

---

[96] Transcriber notes are indicated in brackets { } and Interviewer speech is preceded by "I".
[97] See Blake (1997) for similar assessment of data where recorded speech is compared with speech observed when participants where not in an interview being recorded.
[98] Distinguishing between Kriol and English (and sometimes even Spanish) was not always clear-cut. Some features that were more lexical could not be confirmed as English or Kriol and therefore were transcribed as English when ambiguous. For example, monosyllabic items that had some phonological overlap: English *the*, Belizean English *de/the*, and Kriol *di* were sometimes difficult to distinguish, though this variability is not one that is analyzed here. Additionally, mono-syllabic responses such as "no" are homophonous between English, Spanish and Kriol.

thought to be associated with production or recall difficulty for bilinguals, were also captured and noted, as will be seen in Chapter 6: Code-switching.

3.4.2.3 Interview selection

In order to identify the highest language mixers in the corpus, a "language-mixing rate" was calculated for each interview. All non-Spanish items were extracted and counted for the first 20 minutes of the body of each interview[99]. All non-Spanish items were counted regardless of single or multiword status. In (22) and (23), for example, both the single item *transfer* and the multiword fragment *(if we don't…)* were each counted as one item. In (22), there are two single English-origin items (*transfer* and *account*), and in (23), there is one multiword fragment (*if we don't … crash*).

22. Entraba allá y lo hacía **transfer** en mi **account** yo hacía el cheque y se lo daba a Greg
    'It would go in there and I would transfer it, I would do the cheque and I would give it to Greg'
    (BSEK/45NJ/45NJ_43NJ_3/01:06:01/)

23. Sabes que Susana, **if we don't take over** *dem* **bills, the bills, the credit cards and the, and the** *lee* **loans that they had, they were going to crash,** iba a hacer crash.

    'You know what Susana, if we don't take over those bills, the bills, the credit cards and the, and the little loans that they had, they were going to crash, they were going to crash.'
    (BSEK/45NJ/45NJ_43NJ_3/01:03:35/)

Extractions were done after the first twenty minutes of the interview, with the exception of one interview, which was less than 30 minutes long, in which case extraction was done from the whole interview. Speakers were considered low language mixers if the number of alternations or non-Spanish items (irrespective of fragment length) fell within the range of 0-25. Medium language mixers were those speakers that used between 26 and 55 non-Spanish items in their speech, and high language mixers were those that utilized over 56 non-Spanish items in the 20-minute

---

[99] The body of the interview is considered here to be the segment between the first 15 minutes of the interview and the language component at the end of the interview. Typically, this is about 40 minutes long, with the exception of interviews that were shorter or longer than an hour.

segment. The resulting categories give a measure relative to the other speakers in the same corpus and make it possible to select interviews which potentially contain the richest language mixing data. This information is summarized below according to region and the age groups. The young speakers are between 15-35 years of age, middle-aged speakers between 36-55, and older speakers between the ages of 56-95.

Table 7: Level of language mixing by region and age

	LOW mixers 0- 25			Total LOW	MEDIUM mixers 26-55			Total MEDIUM	HIGH mixers 56+			Total HIGH	Total
AGE ⇒	Y	M	O		Y	M	O		Y	M	O		
REGION ⇓													
North			**7**	7	**4**	3	3	10	**4**	1	0	5	22
West	5	2	**6**	13	**5**	2	2	9	**7**		0	7	29
Total	5	2	13	**20**	9	4	5	**19**	11	1	0	**12**	51

Key: Y=young, M=middle age, O=old

The low language mixers are predominantly from the west (13/20) and tend to be older (6/13), though close to half of the low language mixers in the west are young (5/13). On the other hand, medium and high language mixers are evenly split between regions and are predominantly young. There are no older high language mixers. In the north, only older speakers are low language mixers, and most of the young people are medium to high language mixers. In the west, only young speakers are high language mixers. That language mixing should appear to correlate with age is not surprising: older speakers would have had less access to English, especially in the smaller, more rural communities, as is the case in the west, simply because there was less access to formal education. In addition, historically the west was more isolated than the north (LePage and Tabouret-Keller 1985:64; Hagerty 1979), as noted previously. The language mixing rates are not

discussed here further since they are a general measure of language mixing used to select interviews for analysis in the current study. Normalized borrowing and code-switching rates for the participants selected for this study are discussed in Chapter 6 where code-switching and community patterns are discussed.

The interviews from the twelve high language mixers were selected for analysis given the focus of the current study on language mixing. It is in these interviews that the ideal data would be found. These speakers form a relatively homogenous group in terms of age, though there is one speaker who is in their early 50s and another who is thirty years old. Most had completed high school with the exception of the two youngest speakers who were currently enrolled in high school. The breakdown by region and gender is presented in Table 8 below [100]. Five of the twelve speakers are form the north while seven are from the west though the number of hours of recorded time is approximately equal[101].

Table 8: Participants by region and gender

	REGION					
	WEST		NORTH		TOTAL	
SEX/GENDER △	N	%	N	%	N	%
MALE	4	67	2	33	6	50
FEMALE	3	50	3	50	6	50
TOTAL	7	58	5	42	12	100

---

[100] Traditionally, efforts are made to analyze corpora that are balanced for sex/gender; therefore, this break-down of participants is presented. Nonetheless, at this time I do not assess correlations between language use and sex/gender. Despite the received wisdom that women use more standard features or are leaders of linguistic change, sex/gender is not necessarily directly linked to linguistic behavior (Eckert 1989). It has been shown, for example, that correlations between sex/gender often mask other factors such as mobility (Poplack et al. 1988:96) or that bilingual ability rather than sex/gender is a stronger predictor of code-switching (Poplack 1980:612). Instead, I look directly at social factors of region and bilingual ability. See also (Fuller Medina and Roy 2010) for a critique of the use of biological sex as the operationalization of gender in quantitative analyses.

[101] Three participants from the west were recorded together in one session.

3.4.2.4 Linguistic profile of participants

A linguistic profile for each participant was obtained via triangulation of information from self-reports, the data itself, and the ethnographic observation previously discussed in Chapter 3 (Travis and Torres Cacoullos 2013:5). It was found that participant self-reports corresponded to observations of their language use and the recorded data. One minor exception was with 45NJ who reported not caring much for Kriol and not using it, yet their interview was one that provided possibly the most Kriol features (e.g. the plural marker *dem* and the future temporal marker *a*). Reports of restricted Kriol use in the west tended to coincide with the ethnographic observations in the community.

Following standard procedures in variationist research on bilingual data, self-reports were used to obtain information on participants' proficiency and language use (Travis and Torres Cacoullos 2013; Poplack, et al. 2006; Poplack et al 1988: fn3, Poplack 1989, 1980:592-593)[102]. A language component was included in the interview rather than utilizing standardized tests to determine bilingual ability[103]. These data were then used to operationalize bilingual ability for the sample analyzed. Participants were asked how often they speak Spanish, English, and Kriol, and with whom, to rate the three languages (passable, well, excellent, not well), and to indicate which is their dominant language or the language they use most of the time. This information was then used to classify language use as *restricted* and *non-restricted*. If they reported using Spanish, Kriol, or English at home with immediate family, as well as with extended family, with peers, and/or

---

[102] For a detailed treatment of self-reports see Poplack et al. (2006:191) who utilize self-reports to calculate a French proficiency index based on "1) age and means of acquisition, (2) frequency of use, (3) degree of proficiency, and (4) language used with selected interlocutors". This score is used to place speakers in groups ranging from low to high proficiency. While the current study does not calculate a Spanish proficiency score, it follows the above-mentioned variationist studies and, more specifically, self-reports include points 2, 3, and 4, from Poplack et al (2006). As more speakers are added to the analysis in the future, a more fine-tuned method of making use of self-reports such as a language index used by Poplack et al (2006) may have to be implemented.

school/work, then this was considered non-restricted use. If they used one of the languages in only one of these social contexts, then language use was considered to be restricted.

Furthermore, these categories of language use were translated into categories of bilingualism: low-bilinguals and bilingual or trilingual. All speakers included in the corpus met the criteria of bilingual, however, as in many bilingual corpora, they have varying degrees of bilingualism, as supported by self-reports[104].   In order to capture this, if a participant reported non-restricted use of two languages and also reported both as their dominant language, then they were regarded as bilingual. If they reported non-restricted use of all three languages and claimed no one language as their dominant one, then they were classified as trilingual. No additional assessment was made as to their level of bi/trilingualism. On the other hand, if speakers reported restricted use in two or more languages, and also named only one language as their dominant language then these speakers were coded as "low-bilinguals" regardless of the language they reported as their dominant one, since their self-reports seem to reflect a tendency towards monolingualism. This means that those speakers who reported English dominance together with restricted use of Spanish and Kriol are considered low-bilinguals for the current purposes. Likewise, those speakers who reported Spanish dominance but restricted use of English and Kriol are also considered low-bilinguals. It is important you note that the categories summarized represent a way of operationalizing the major differences between the two groups that emerged in the sample analyzed; "low-bilingual," bilingual, and trilingual are not absolute terms but rather reflect the differences within the sample as reported by participants[105]. As will be seen in discussion of community patterns in Chapter 6, whether speakers are Spanish or English-dominant appears to

---

[104] Poplack (1980) also found that amongst the bilingual speakers in the Puerto-Rican study, there were varying degrees of bilingualism according to self-reports.

[105] For a detailed treatment of self-reports from Spanish-English bilinguals and the correlations between these and language mixing behaviours see Poplack (1980).

be less important than whether or not they use two or more languages in non-restricted ways and also identify more than one language as their dominant one. Rather than assign the label English-dominant or Spanish-dominant, I opt to call them "low-bilinguals". These data are summarized below in Table 9.

Table 9: Language use, language dominance and bilingualism

REGION	SPKR	DOM LANG.	LANGUAGE USE			BILINGUALISM
			SPANISH	ENGLISH	Kriol	
N	45J	B	N	N	R	Bilingual
N	58	K	N	N	N	Trilingual
N	05N14	K	R	N	N	Bilingual
N	62N	E	R	N	R	low bilingual
N	60NJ	E	R	N	R	low bilingual
W	29W	B	N	N	R	Bilingual
W	35W	B	N	N	N	Trilingual
W	08W	T	N	N	N	Trilingual
W	38W	S	N	R	R	low bilingual
W	11WJ	S	N	R	R	low bilingual
W	12WJ	S	N	R	R	low bilingual
W	13WJ	S	N	R	R	low bilingual

Key:
Dominant Language: B=both languages, T= all three; K=Kriol; E=English; S=Spanish. (those labelled N under language use)
Language use: N=non-restricted (used at home, immediate family, with peers, and/or school/work), R=restricted (used only at home with immediate family, or with peers, or only at school/work)

The NBS speakers generally report to be Kriol or English-dominant while the WBS speakers report more Spanish dominance. Claiming Kriol as a dominant language, as some participants do, can potentially be taken as a reflection of covert prestige and shifting attitudes towards Belize Kriol. Though not shown here, all northerners, except one, rate their Spanish as "pasable" (passable) while westerners rated their Spanish more highly as "*bien*" (good) or even "*excelente*" (excellent). For WBS speakers for whom English is restricted, English is used mostly in

school or outside the community, and restricted use of Kriol is mostly attributed to low ability in this language or even an uncertainty as to whether or not their language use could be considered Kriol. As Table 9 shows, there is a high correspondence between language dominance and language use, though it is not categorical. Those speakers who claimed to speak two or three languages at a comparable level also reported using them in non-restricted ways.

3.5 Theoretical Framework

The current study employs the framework of comparative variationist sociolinguistics (Poplack 2000; Poplack and Tagliamonte 2001)

3.5.1 Variation theory and language contact

One of the main premises of the Variationist Method is that language varies systematically, rather than randomly. As a consequence, this systematic variation can be modeled quantitatively (Young & Bayley 1996). Central to this approach is the *linguistic variable* (Labov 1982) which is made up of a class of variants that may have the same grammatical function or referential meaning (Tagliamonte 2006:9; Poplack 2000:88). A speaker then has a choice in which form will be utilized in discourse; the variationist approach attempts to uncover the patterns in the distribution and conditioning of the competing variants in discourse by way of large-scale quantitative analysis.

Rather than analyzing only contexts where the variant under study surfaces, the principle of accountability (Labov 1972) dictates that analysts must additionally consider those contexts where a variant could have occurred but did not. In other words, variants must be contextualized in terms of the larger structured grammatical subsystems to which they belong. For example, in order to study the inflected future in Spanish (e.g. *viajaré*, 'I will travel'), it is necessary to consider not only tokens of this variant, but also competing variants in the same grammatical subsystem, i.e., the

periphrastic future (*voy a viajar*, 'I am going to travel'), and present tense forms with future temporal reference (*viajo mañana*, 'I travel tomorrow'). Likewise, all contexts where the variants in question cannot or do not alternate must be excluded, e.g. habitual non-future meaning (*viajo sóla siempre* 'I always travel alone'). This constitutes the *variable context*. All tokens that fall within the variable context would then be analyzed quantitatively to ascertain which factors constrain the variable under study[106]. Language contact phenomena present a challenge to the application of this method since code-switches and borrowings are not prototypical variables as described above. This is discussed further in the subsequent sections.

3.5.2 Language contact phenomena and the challenge of circumscribing the variable context

In the case of language mixing, adhering to the principle of accountability means that one would not only have to account for all cases of actual code-switches, but also all instances of where a code-switch could have occurred but did not (Poplack 1993; Meyerhoff et al. 2012:124). Since in typologically similar language combinations, such as Spanish-English-Kriol, there are many permissible switch points, this would prove an enormous task (but see Sankoff and Poplack, 1981). Furthermore, a trilingual speaker in the current study can chose between Spanish, English, Kriol, or code-switched discourse, or some combination thereof, to express something, and there are no reliable ways to predict which discourse mode will be chosen. In instances where CS is chosen, there is currently no way to predict whether speakers will actually switch at a given permissible switch site. In fact, since code-switching is "first and foremost a discourse device, … a code-switch

---

[106] A multivariate analysis is generally carried out which considers the effects of multiple independent factors simultaneously. This type of analysis produces probabilities or factor weights which are a measure of the favouring or disfavouring effect of a particular factor and also indicates which factors are statistically significant. This type of analysis can generally be carried out with prototypical variables where multiple linguistic and extralinguistic factors are thought to constrain variability. The current data did not lend itself to a multivariate analysis. Nonetheless, the underlying grammar can still be accessed by examining regular patterns of use and the application of language specific grammatical rules.

is no more predictable at the *local* level than, say, a curse or a joke" (Poplack 1993:276). Borrowing presents a similar dilemma since "any content word is fair game" (Ibid: 277).

However, this does not preclude the application of the principles of the variationist method. The principle of accountability can still be adhered to, and the variability that is the object of study can still be circumscribed. One of the earliest applications of the variationist principle of accountability to bilingual data can be found in Pfaff (1979:295) where she considered all other language material, whether single or multiword fragments. She also reported on all the relevant data and not only on select individual examples. Likewise, Poplack (1980), in her study on Puerto Rican Spanish, also extracted (and reported on) all instances of other language material in the data, and in this way adhered to the Principle of Accountability and the object of study was still clearly defined as was the context for the variable linguistic phenomena under study (Poplack et al. 1988). Thus, in studying bilingual data, circumscribing the object of study is a fundamental premise. This is the methodology employed in the current study.

All non-Spanish fragments were extracted without defining the status of any of this material *a priori*. In addition, other-language material is quantitatively analyzed so as to reveal major patterns in language use rather than analyzing single items, which may be perceptually salient but not reflective of regular patterns. Regular patterns refer to a "series of parallel occurrences (established according to structural and/or functional criteria) occurring at a non-negligible rate in a corpus of language use" (Poplack and Meechan 1998:262). This approach enables these types of patterns in the natural speech of the community to be identified and distinguished from isolated or idiosyncratic occurrences that may be found in the speech of bilingual speakers but that are not representative of productive or recurrent patterns (Poplack 1993:263). In addition, by analyzing the regular patterns of the monolingual vernaculars of speakers, it then becomes possible to determine if the other-language material shows patterns that are parallel or dissimilar those in monolingual

discourse. Prescriptive norms would be of little use in identifying the system membership of the items in mixed speech that are in question. In studying possessives, for example, any English-origin nouns that enter into the non-standard possessive constructions outlined above might be deemed to *not* pattern with Spanish if prescriptive norms are used as the benchmark rather than the patterned use of such possessive constructions in the varieties used by the speakers.

3.5.3 The application of the comparative method to language mixing

The comparative method (Poplack 2000; Tagliamonte and Poplack 2001) as applied to language contact provides useful methodological tools for analyzing possible outcomes of language contact such as language mixing and contact-induced change. In the comparative method, tokens from different subsets of data from the varieties under study are analyzed separately while keeping factors constant across groups for the same linguistic variable (Meyerhoff 2009). This gives information as to how the same variable is constrained in different datasets under study. The results from these independent analyses are then compared to determine where the subsets of data may pattern similarly or in disparate ways. In the case of language mixing, a working hypothesis is that if the conditioning of items from language A (LA) embedded in discourse of language B (LB) pattern with comparable items in the LB system (and is different from the LA system), then this is taken as evidence of integration and LA items may be considered borrowings. Likewise, if LA items pattern with the LA system and this patterning is different from the LB system, then these may be considered code-switches since two grammars are present in the discourse (Poplack and Meechan 1998: 130). Thus, the comparative method is useful in determining if non-native material in a recipient language are borrowings (integrated) or code-switches (unintegrated) and therefore, produced by separate underlying mechanisms.

In order to determine if the English-origin fragments are integrated into Spanish or if they retain English grammar, the comparative method takes advantage of conflict sites between languages (Poplack and Meechan 1998). These are areas where languages differ in terms of rates and/or conditioning. By the same token, areas where both languages are homologous in their grammars (e.g. word order) are sites of equivalence. Conflict sites serve as the axis for the quantitative comparison of the patterning of those single items whose status as a code-switch or borrowing may be ambiguous (Poplack and Dion 2012; Torres Cacoullos and Aaron 2003; Poplack and Meechan 1998a). This is because it is at these sites that it can be seen which grammar is operating on the item in question. Spanish, for example, requires number agreement between determiners, modifiers, and the nouns they modify, while English only does so in limited contexts (third person possessives, indefinites, and demonstratives). Example (24) below shows number agreement between the determiner *los* ('the') and *midwives*, which reflects a Spanish agreement rule indicating that *midwives* is integrated into Spanish[107]. However, a single example does not indicate if the application of the number agreement rule constitutes a regular pattern or an anomaly. In addition, it is not known if, in the aggregate, such English-origin nouns show the same patterning of number agreement as monolingual Spanish nouns or if they behave in ways that are more akin to English nouns in English discourse. If the former is true, then they can be considered borrowings, but if the latter is true, then they can be considered code-switches. In order to determine the underlying grammar constraining English-origin nouns, for each conflict site, all eligible tokens are subjected to quantitative analysis to reveal their patterning. Eligible nouns in comparative datasets for Spanish and English are also analyzed so that the conditioning of nouns

---

[107] Spanish also requires gender agreement which in this example appears to be lacking if we assume based on real-world experience that midwives are generally women. However, as will be discussed in Chapter 5, the masculine determiner is used with reference to superordinate concepts in Spanish. This highlights the need to examine data via quantitative analysis to uncover the patterns of use. Gender agreement or apparent lack thereof is discussed in detail in Chapter 5.

for each language is revealed. When the mixed data is compared to these datasets it becomes possible to see if it patterns with Spanish or English. Accordingly, at the coding stage, no decision is made as to the status of the item. In the case of number agreement, tokens are coded fro the presence of plural markers on the determiner on nouns in each dataset and only then are they subjected to quantitative analysis to determine if the English-origin items parallel the Spanish pattern of agreement.

24. Llaman   a   los   **midwives**   en   las   casas
    call3PL.PRES   to   the   midwives   to   the   homes

    'They call the midwives to the homes.'

(BSEK/05N14/05N14/ 00:02:26)

Given that it has been shown that borrowing is largely a morphosyntactic process, which is only optionally accompanied by phonological and social integration (Poplack et al. 1988), and that phonology is an unreliable diagnostic, only indices of morphosyntactic integration are employed to assess the status of single English-origin items in Spanish discourse.

Comparison datasets are created by extracting corresponding categories from monolingual Spanish stretches of discourse as illustrated in (25) and monolingual stretches of English shown in (27). By creating and using the comparison datasets from the same speakers, the inherent variability in each variety is taken into account so that tokens are not analyzed for language membership according to prescribed norms but rather according to the vernacular norms to which speakers are actually adhering (Poplack and Meechan 1998:130). In the following examples, Spanish nouns are underlined and English-origin nouns are in bold.

Spanish nouns in Spanish discourse

25.
cuando	ya	resultan	esos	<u>conflictos</u>	y	todo,	ya
when	already	result3PL.PRES	those	conflicts	and	everything	then

no	es	nada	good
no	is	nothing	good

'Then when those conflicts come about and everything, nothing good comes of it.'

(BSEK/08W/08W_03/00:17:56)

English-origin nouns in Spanish discourse

26.
En	el	mismo	**school,**	en	el	palapa.
In	the	same	school	in	the	palapa

'At the same school, in the palapa.'[108]

(BSEK/11WJ/11WJ_12WJ_13WJ/00:07:12)

English nouns in English discourse

27. They can win in any **seat** in Belize City right now. (BSEK/45NJ/45NJ_43NJ_3/01:54:12)

As will be seen in Chapter 4, a comparison dataset for verbs that have object clitics associated with them is also extracted in a similar fashion.

3.6 Methodology

This thesis aims to determine patterns of use among Belizean Spanish speakers who engage in language mixing and to determine what the surface manifestations of this language mixing might be. Where intrasentential code-switching is identified, data is examined to assess the grammaticality of those code-switches as per predictions of the equivalence constraint and free morpheme constraint (Poplack 1980). In order to respond to these questions, single English-origin single items must be separated from multiword fragments, and within the subset of single words, lexical and functional categories must also be separated and clearly identified so that comparative

---

[108] Tatch roof structure like a gazebo.

analysis for corresponding categories can be carried out across the three datasets (Spanish, English, Kriol).

3.6.1 Extraction and coding of data

All items that were in a language other than Spanish were extracted from 12 interviews.

Number of Words

In order to determine the overall community pattern of mixing, i.e., borrowing and/or code-switching, all data extracted were first coded according to whether they were single word or multiword items as illustrated in (28-29) All English origin items are indicated in **boldface.**

Single-word

28. | Llaman | a | los | **midwives** | en | las | casas |
    |--------|---|-----|----------|-----|-----|-------|
    | call3PL.PRES | to | the | midwives | to | the | homes |

    'They call the midwives to the homes.'

    (BSEK/05N14/05N14/ 00:02:26)

Multiword

29. | Tenían | cabañas, | estaban | utilizados | **by** | **visitors** |
    |--------|----------|---------|-----------|--------|----------|
    | have3PL.IMP | cabañas | be3PL.IMP | use.PTCP | **by** | **visitors** |

so	**we**	**camped**	**out**
so	**we**	**camped**	**out**

    'They had cabañas, they were used by visitors so we camped out.'

    (BSEK/58N/58N/07:51.4)

3.6.2 Distribution of data

Single words were found to make up the majority of the data as indicated in Table 10 below. All non-Spanish fragments were also coded for language. Single items, such as *midwives* in (28) above, were coded as English. While Spanish is also in contact with Kriol and some speakers

claim Kriol as their dominant language, it is difficult to determine if a lexical item is Kriol or English when it is in isolation, given the overlap in lexicon between English and Kriol. A finer-grade distinction was not made between English and Kriol for single lexical items in isolation with the exception of items such as *gyal* 'girl,' and *bway* 'boy' which are Kriol (these appear in italics below)[109]. Other items were coded as Mayan (in uppercase), as in (31), or 'other.' This last category included words from other languages such as *karaoke*, which were infrequent. The vast majority of the single non-Spanish items were English. Mayan words were limited mostly to names of food, plants, or supernatural beings, and Kriol to the items previously named.

Table 10: Distribution of single and multiword items

	North	West	TOTAL
# of words	N %	N %	N %
Multiword fragments	428 37	206 26	**634** *32*
Single items	730 63	591 73	**1321** 68
TOTAL	**1158**	**797**	1955

30. y   se   muere   el   de-l   otro   lado   *bway*,   murieron
    and   CL   die   one   from-DET   other   side   boy   died

    como   cuatro
    like   four

    'And then the one from the other side dies, boy, like four of them died.'

    (BSEK/45NJ/45NJ_43NJ_3/01:08:16)

---

[109] In example (34) further below, *microwave* might be the same word used in Kriol and English and are considered etymologically English. As noted earlier, Spanish differs sufficiently from both English and Creole that the conflict sites chosen are still useful for assessing the integration of such single items.

31. Y le dan su mazorca, su PIBINAL[110]
    and CL give their corn their PIBINAL

    'And they give them their corn, their PIBINAL.' {food served at the end of prayers}
                                          (BSEK/54N/53N_54N/00:06:57.240)

The vast majority of multiword fragments were English as were the single items. An English-origin multiword fragment is exemplified in (32) while a Kriol-origin multiword fragment can be seen in (33). Tokens were considered Kriol based on morphosyntactic features. This means that the fragment was considered Kriol if it contained identifiable and uncontroversial Kriol lexical and functional features as per documented features of Belize Kriol (Escure 2013; Herrera et al 2007; Decker 2005) as well as my own knowledge of the language. In (33), for example, we find *bally* 'guy/man,' the completive marker *don* and no copula consistent with Belize Kriol. This can be contrasted with (32) which is internally consistent with English as evidenced by the English copula *is*, and the English indefinite determiner marker *a* versus the Kriol indefinite determiner *wah*.[111]

32. **This** **is** **a big** **thing** que pasó acá
    This is a big thing that happened here

    'This is a big thing that happened here.'
                                          (BSEK/58N/04N_58_N/00:52:20)

33. *No,* *he* *don* *in,* *da* *bally* *don* *in,*
    No he COMPL in that guy COMPL in

    guaranteed!
    guaranteed

    'No he is in already, that guy is in already, Guaranteed!'
                         (BSEK/45NJ/45NJ_43NJ_3/01:52:53) {local elections}

---

[110] Pibinal: corn roasted in an underground pit oven. Píib – underground, nal -corn (Suárez Molina 1996) http://www.mayas.uady.mx/articulos/voces.html Accessed December 12, 2015.

[111] The notion of discreteness between Creoles and their lexifier languages remains unresolved in the field (Patrick 2004; Winford 1988; Bickerton 1973) and given that I did not consider prosody and phonology; I was conservative with the coding of Kriol fragments. Unless items in the fragment could be confirmed as Kriol, they were coded as English. Though it was rare to find fragments which were ambiguous.

Because the Kriol fragments identified in this fashion were not substantial enough for quantitative analysis, they are not analyzed separately and are simply indicated in italics in any ensuing examples.

Category of single-word items

Each single item was coded for functional and lexical categories (see 34-42). The last category, 'other,' includes infrequent items such as kinship terms, interjections, and numbers. Non-Spanish prepositions, conjunctions, pronouns, complementizers, and determiners or quantifiers were not found in the corpus[112]. Discourse markers are listed separately as function may be more pragmatic since they carry little lexical meaning. Discourse markers (DM) were further coded for type of discourse marker (e.g. *like*, *then*).

Lexical categories

Nouns

34. El   **microwave**   cayó,   like,   sólo   cayó   así,
    DET   microwave   fell3SG.PRET   like   alone   fall3SG.PRET   like.that

    'The microwave fell, like, it just fell like that.'

                                                                    (BSEK/29W/29W/11:00.1)

Proper nouns

35. El   **Split**   es   el   lugar   más   bonito   para   nadar
    the   Split   is   the   place   more   pretty   to   swim.INF

    'The Split is the nicest place to swim.'

                                                                    (BSEK/08W/08W_02/8:12.6)

Adjectives

36. Y   ese   día   llevaba   una   blusa   **strapless**

---

[112] This is in line with most borrowability hierarchies which place functional or closed class items lower on the scale of borrowability (Matras 2009; Winford 2003; Muysken 2000; Romaine 1995). It is also consistent with corpora of spontaneous speech (Poplack et al. 1988; Poplack 1980) which generally show nouns as being the most borrowed category.

and	that	day	wear1SG.IMP	a	blouse	strapless

'And that day I was wearing a strapless blouse.'

(BSEK/11WJ/11WJ_12W.J_13WJ/01:39.6)

Verbs

37. Pero	yo	lo	quería	**trayar**
but	I	CL	want1SG.IMP	try.INF

'But I wanted to try it.'

(BSEK/58N/58N/11:28.0)

Bilingual compound verbs

38. Hacen	**shake**	el	pan	three times.
do3PL.PRES	shake	the	pan	three times

'they shake the pan three times.' {explaining game called "kick the pan"}

(BSEK/29W/29W/06:51.)

Other

Kinship terms or forms of address

39. le	digo:	"pero	¿cómo	**gran**?
CL	say1SG.PRES	but	how	gran

'I say to her: but how gran?'

(BSEK/58N/58N22:47.1)

Interjections

40. **Shit**!	¿cómo	se	llama	la	otra	una?
shit	how	CL	name	the	other	one

'Shit! What was the other one's name?'

(BSEK/ 45NJ/45NJ_43NJ_3/01:05:19)

Discourse markers

There were no single functional categories such as complementizers, pronouns, and so forth in the data. Single items which have a discourse pragmatic function are subsumed under discourse markers and are exemplified in (41)-(42).

41. Me   gusta   bastante   Bon Jovi,   **right**, y   un   sueño mío
    CL   like   a.lot   Bon Jovi   right   and   a   dream mine

    era   ir   a   ver   a   Bon Jovi
    was   go.INF to   see.INF   DOM   Bon Jovi

 'I like Bon Jovi a lot, right, and a dream of mine was to go see Bon Jovi.'
                                                    (BSEK/08W/08W_03/23:23.0)

42. yo   me   sentí   **like** {laughter}
    I   CL   feel1SG.PRET   like
    'I felt like'
                                    (BSEK/11W.J/11W.J_12W.J_13W.J/10:47.0)

The distribution of single non-Spanish items in the data is shown in Table 11. The largest categories, and the ones addressed in this thesis are highlighted. Nouns make up the largest category, followed by discourse markers and verbs. These three categories, along with multiword fragments, are analyzed in the thesis. Nouns and verbs are assessed for system membership, discourse markers are assessed for their potential function as code-switch triggers, and multiword fragments are analyzed first for their status in terms of type of switch (intrasentential or intersentential, for example) and then, once intrasentential switches are isolated, these are evaluated against the equivalence constraint and the free morpheme constraint. Other types of code-switching such as intersentential switching are not relevant for assessing syntactic constraints because the alternation between languages takes place between independent clauses where a syntactic boundary does not need to be navigated by the speaker.

Table 11: Distribution of lexical categories of non-Spanish single items

PART OF SPEECH	N	%
Nouns	732	55
Discourse markers	221	17
Verbs	191	14
Adjectives	90	7
Adverbs	40	3
Numbers	21	2
Kinship terms	21	2
Interjections	5	0.4
TOTAL	**1321**	**100**

Given that individual Mayan examples are infrequent and comparative Mayan data is unavailable, these are not analyzed in the current study. Only English-origin material is analyzed in the ensuing chapters. The datasets in the following chapters are further refined by the exclusions of tokens that are difficult to code because they were ambiguous in nature due to poor sound quality or incomplete clauses as a result of interruptions. The coding procedures for the analysis of verbal items, nominal items, and multiword fragments are discussed individually in the respective chapters.

CHAPTER 4 VERBAL BORROWING

4.1 Introduction

This chapter focussed on the analysis of single English-origin verbs in Spanish discourse as these account for 14% of the current data. This chapter responds to two main research questions: (i) Are the English items in bilingual compound verbs (henceforth BCVs) borrowings, code-switches or do they represent some other phenomena (e.g. emergent hybrids, creolization)? (ii) Are there aspectual restrictions on the English-origin verbs that can occur in BCVs and what does this tell us about the grammaticalization of *hacer*? In order to respond to these questions, this chapter proposes some methodological tools for circumscribing BCVs so that the object of study is clearly defined. Similarly, aspect is operationalized (following Walker 2010a) in a manner that draws on independent theoretically motivated criteria with a view of offering a potentially transparent and replicable analysis of aspect.

The disparate characterizations of BCVs will be addressed by making use of conflict sites (Poplack and Meechan 1998a,b), or areas of grammar where Spanish and English differ, in a quantitative analysis of 182 BCVs[113]. These conflict sites are morphological marking on verbs and variable placement of object pronouns. If in the aggregate English-origin items are shown to retain features of English grammar, such as verbal morphology, then they can be deemed single-word code-switches. If, however, they show systematic indices of integration into Spanish, via morphology and clitic placement, then they can be considered borrowings. In terms of aspectual restrictions on BCVs, previous studies have presented conflicting results, claiming, on the one hand, that no restrictions govern English-origin verb types in BCVs (Balam et al. 2014), while on

---

[113] As noted earlier, Kriol data was not deemed sufficient for inclusion in a quantitative comparative analysis. Belize Kriol verbs are reported to generally be bare forms that variably appear with preverbal markers (invariant in forms) to mark tense, mood and aspect (TMA) (Decker 2005). Such markers were never observed to co-occur with BCVs or in place of *hacer* and, as the data will show, there was no variability observed between English-origin forms unmarked for TMA and forms overtly marked for TMA via *hacer*. Thus, the comparison of BCVs to monolingual English is justified at this time.

the other, that since stative verbs are rarely found in BCVs, this suggests an aspectual restriction on the English-origin verbs (Fuller Medina 2013, 2005a,b). In view of conflicting findings, the hypothesis that the class of English-origin verbs in BCVs is aspectually restricted is tested by first operationalizing lexical aspect and then examining the distribution of stative and dynamic English-origin verbs.

I begin by providing some descriptive details about BCVs by examining their frequency and the type/token ratio of English-origin verbs, as well as the distribution of BCVs across the verbal paradigm. The object of study is then clearly defined, i.e., how BCVs can be distinguished from other bilingual constructions with *hacer*. In order to assess whether or not they are code-switches, the first step is to define what is being analyzed. The general coding for verbs and their distribution in the data is subsequently presented followed by analyses of the following conflict sites: (i) verbal morphological and (ii) object pronoun expression. Finally, the English-origin item in the BCV is analyzed with regard to lexical aspect.

4.2 Bilingual Compound Verbs (BCVs): Elements of the Compound

There is general consensus that the "*do–verb*" in a BCV functions as a kind of "helping verb"[114] rather than as a main verb (Wilson and Dumont 2014; Wilson 2013; Fuller Medina 2005a,b; Gardner Chloros 2009; Edwards and Gardner Chloros 2007; Muysken 2000) and has been analyzed as a light verb (Rao and Den Dikken 2014; Fuller Medina 2013, 2010; Gonzalez-Vilbazo and López 2012, 2011; Nakajima 2008). The English complement of the *do*-verb is not a direct object but predicates jointly with the *do*-verb as a constituent (Muysken 2000; Fuller Medina 2013, 2010; Edwards and Gardner-Chloros 2007; Romaine 1995; Moravscik 1975). In (43), for

---

[114] Muysken (2000) uses the term 'helping verb" to avoid confusion with the syntactic category AUX.

example, the BCV here takes the direct object *el dinero*, the English item *bury* is not an object of *hacer*.

43. Hacemos **bury** el dinero
    DO.1.PL bury the money

    'We bury the money.'
                                    (BSEK/29W/29W/00:08:02)

While the *do*-verb in these types of constructions has been called an auxiliary, auxiliaries occupy a different position in the syntax, and in BCV constructions *hacer* is better analyzed as a light verb (Rao and Den Dikken 2014; Fuller Medina 2013, 2010; Gonzalez and Lopez 2012, 2011). The term auxiliary is perhaps used as a synonym for light verb given that both auxiliaries and light verbs can be described as bleached of their main lexical meaning, and therefore are more functional than their lexical counterparts. Nevertheless, it is worthwhile exemplifying BCVs in auxiliary + non-finite type constructions, as in (44) below, to illustrate that *hacer* does not occupy the auxiliary position in the syntax. It co-occurs with auxiliaries rather than being in complementary distribution. The auxiliaries are in CAPS and the relevant constituent is underlined.

44. Como el inglés es el language que se HA
    Since the English is the language that CL have3SG.PRES

    hecho set
    do.PTCP set

    'And since English is the language that has been set'
                                    (BSEK/08W/08W_03/00:30:51)

45. y no lo ESTAMOS haciendo entertain porque
    and NEG CL be1PL.PRES do.PTCP entertain because

    no queremos tampoco hacer congest
    NEG want1PL.PRES either do.INF congest

    'and we are not considering them because we don't want to congest things either.' {context: applications}.
                                    (BSEK/45NJ/45NJ_43NJ_3/01:36:10)

This provides some support for the theoretical analyses of BCVs as light verb constructions (Rao and Den Dikken 2014; Fuller Medina 2013, 2010; González-Vilbazo and López 2012, 2011; Nakajima 2008) and for the cautionary approach taken by some researchers such as Muysken (2000) who refrain from using the term auxiliary for the *do*-verb in BCVs. In the early literature (Pfaff 1979; Reyes 1976) Spanish-English BCVs were thought to be based on Spanish causative constructions but the above analysis does not support this given that BCVs as illustrated in (45) have only one agent. However, in a causative construction, we would expect to see the thematic roles of causer or agent assigned to arguments of *hacer* in addition to the thematic role of patient or causee associated with the event that is being caused (see Wilson 2013; Fuller Medina 2005a; Jenkins 2003). Again, this supports the analysis of BCVs as light verb predicates given that BCVs also have only one agent or causer rather than two, despite the apparent presence of two verbs[115].

Light verbs tend to be form-identical to their main verb counterparts but are dependent on another event to predicate (Butt 2010, 2003). In (46), *hacer* forms part of a complex predicate whereas in an example such as *hice un pastel* 'I made a cake,' it is a main verb. As complex predicates, the two elements predicate jointly "by mapping their combined predicational content to a single monoclausal syntactic domain." (Butt 2010: 56). In (46) and (47) below, for example, *hacer* and the nominal complements (*un llamado* and *travesura*), in the underlined constituents, predicate jointly. The light verb is the same but the predication differs depending on the complement. Thus, the light verb contributes little content.

46. El presidente <u>hizo un llamado</u> a la población civil.
    'The president made a call to the civil population.'

<div align="right">(Bosque 2001:27)</div>

---

[115] Pfaff (1979) and Reyes (1976) suggest that Spanish-English BCVs may be modeled on causatives and González Vilbazo and López (2012) argue that Spanish-German BCVs may have a causative reading. I have argued against a causative interpretation of *hacer* in Spanish-English BCVs (Fuller Medina 2013, 2010, 2005b) as have Jenkins (2003) and Wilson (2013:125-126).

47. <u>Hacemos travesura</u> cuando andamos así en la calle
    'We do mischievous things when we are like walking around in the streets'
<div align="right">(BSEK/29W/29W/00:09:04/)</div>

Light verbs may also have simplex verbal counterparts such as in English - *take a shower* means the same as *to shower*. This is illustrated below for Spanish with *rezar/hacer rezos* 'to pray/to do prayers'. Additional light verb/simplex verb equivalents are illustrated further below in (50).

48. y ella rezaba
    'and she used to pray'
<div align="right">(BSEK/58N/58N/00:23:34.485)</div>

49. y se hacían los rezos
    'and they used to do prayers'
<div align="right">(BSEK/44N/44N24:54:500)</div>

50. a. Fumigar = hacer fumigaciones 'fumigate'
    b. Prometer = hacer promesa 'promise'
    c. Travesear=hacer travesura 'do mischief/be mischievous'

If we compare (43) above to (51) below, it can be seen that the predication is dependent on the complement of *hacer,* in this case the English-origin item, just as in the monolingual examples. This means that the complement of the light verb is contributing the argument structure of the predicate (Rao and Den Dikken 2014; Fuller Medina 2013, 2010; Wilson 2013; González-Vilbazo and López 2012, 2011).

51. Teníamos que   hacer **compete**,   ahm   hacer **create** un   booth, y
    have.to1PL.IMP   do.INF compete   ahm   do.INF create a   booth and

    hacer   **compete**   con   otros organizaciones
    do.INF   compete   with   other organizations

    'We had to compete, ahm, create a booth, and compete, with other organizations'
<div align="right">(BSEK/58N/58N/00:04:59)</div>

4.3. English-origin Verbal Items

As noted in Chapter 3, all etymologically English-origin verbal forms were extracted and coded for morphological marking without regard to the status of the item as a code-switch or borrowing. Once all tokens were coded, then they were subjected to quantitative analysis in order to reveal whether they pattern with the Spanish or English system of verbal morphology. These coded forms constituted 14% of the single English-origin items. English-origin verbs in Spanish discourse were then coded for verbs appearing in constructions with *hacer* as BCVs as in (52); Spanish morphological affixation as in (53), which will be referred to as synthetic integration; for zero marked verbs (54) and past participles (55). In addition, some Spanish verbs were used with English meanings and these were coded as calques as in (56) (*realicé*).

Bilingual Compound Verb

52. Graba videos, los <u>hace</u> **edit** y todo,
    record3SG.PRES videos CL do3SG.PRES edit and everything

    es bien buena.
    be3SG.PRES really good

    'She records videos, she edits them and everything, she's really good.'
    (BSEK/08W/08W_03/12:27.3)

Synthetic (adaptation to *–ar* class)

53. pero yo lo quería **trayar**
    but I CL want1SG.IMP try.INF

    'But I wanted to try it.'
    (BSEK/58N/58N/11:28.0)

Bare

54. … diseñar cosas, fotos, **edit**, todo
    design.INF things pictures edit.INF everything
    '…design things, pictures, edit, everything.'
    (BSEK/08W/08W_2/00:06:53)

Participle

55. tenemos   bastantes-,   diferentes   artistas **signed** al   label
     have1PL.PRES   a.lot   different   artists   PTCP   to-DET   label

'We have a lot of, different artists signed to the label'

(BSEK/08W/08W/ 04:10.0)

Calques

56. **Realicé**[116]   que   estoy   miseando   big   money
     realize1SG.PRET   that   is1SG.PRES   miss.PTCP   big   money

'I realized that I am missing out on big money'

(BSEK/45NJ/45NJ/45NJ_43NJ_3/ 13:57.3)

Table 12: Distribution of English-origin verbs in Spanish discourse

	N	%
Periphrastic (BCV)	167	91
Synthetic	7	4
Calque	5	3
Bare	3	1
Participle	2	1
Total	**184**	**100**

As Table 12 shows, 91% of verbs appear in a BCV construction and only 4% appear to be synthetically integrated. As the discussion further below will show, this result is quite different from what is reported for other Spanish-English data and other Romance languages in contact with English where English –origin verbs are integrated into one of the major verb classes (Poplack et al. 1988; Pfaff 1979). Verbs showing synthetic integration can be considered borrowings since they

---

[116] Calques or loan shifts (Haugen 1956) such as *realicé* were coded separately and not included in the category of –ar incorporated verbs since this is a Spanish word with an English meaning applied to it. (Spanish: *Realizar*= 'to fulfill, carry out'; English: *to realize*= 'to become aware'). These were infrequent but included in the distribution because they indicate a form of borrowing where the English meaning is borrowed but mapped on to a native Spanish word which has the same form as the English one. The difference with the type of borrowing of interest here is that it is not etymologically English and does not require morphosyntactic integration. See Hagerty (1996) for other examples in Belizean varieties of Spanish and Haugen (1956) for an early typology of types of borrowing.

show integration into Spanish via affixation of Spanish TMA morphology. The participles as in (55) can be considered single-word code-switches since they show English morphological marking and, thus, pattern differently from the synthetically integrated English-origin verbs. The status of the English-origin items in BCV constructions is not clear, however, since these items show no overt morphological affixation derived from either Spanish or English. BCVs could be code-switches given that the onset for the code-switch could be between *hacer* and the English item with a code-switch back into Spanish after the English-origin verb. Alternatively, these constructions might be borrowings, the majority of non-native items in bilingual data consists of single-words and these have tended to be borrowings (Poplack 2015; Poplack et al. 1988).

4.3.1 BCVs: Defining the object of study

In order to assess the nature of BCVs the "variable context" needs to be circumscribed given that *hacer* can occur as a lexical verb (57), functional verb (58) or in idiomatic expressions as in (59) (Solé 1966).

57. Maria   hizo   un   pastel
      María   make3SG.PRET   a   cake

      'Mary made a cake.'

58. a. El   presidente   hizo   un   llamado.
        the   president   make3SG.PRET   a   call

      'The president made a call'

      b. María   hizo   comer   el   pastel   a   Juan.
        María   make3SG.PRET   eat.INF   the   cake   to   Juan

       'Mary made Juan eat the cake.'

59. María   se   hizo   la   tonta.
María   CL   do3SG.PRET   the   fool

'Mary played the fool'

The principle of accountability dictates, that only eligible variants should enter into an analysis of any given linguistic variable, therefore, only bona fide BCVs should be included the analysis. But what constitutes a bona fide BCV? The literature does not provide an explicit method for demarcating and extracting BCVs. The structure of BCVs is sometimes considered to be *hacer* with a nominal complement (Sliva Corvalán 1994:47), but more often assumed to be [*Hacer* + infinitive] without a rationale as to why the English item is considered an infinitive. Extraction is simply described as extraction of all instances of [*Hacer* + English infinitive] (Balam et al. 2014), [*Hacer* + V$_{ENG}$] (Wilson and Dumont 2014), or [*Hacer* + V] and [*Hacer* + N], while forms modified by a determiner are generally excluded from analysis (Wilson 2013; Fuller Medina 2005a,b). Wilson (2013) provides somewhat more detail indicating that all instances of *hacer*, regardless of the language of the complement, were extracted and only those with an English bare infinitive as an adjunct were retained. In a similar fashion to Fuller Medina (2005), any English lexical items in BCVs that were modified by a determiner were excluded from analysis (e.g. *hacemos un get-together).*

Two main assumptions underlie the study of BCVs, at least in the Spanish-English literature, i.e., (i) BCVs are [*Hacer* + English infinitive] structures and (ii) the identification of the category of the English element is a straightforward procedure that can be made by examination of the surface form of the English-origin lexical item in isolation (see Jenkins 2003 for some exceptions to these assumptions). Both these assumptions call for further examination. In the

BCVs in (43) and (51) above for example, the English item (*compete*, *bury*) is clearly verbal[117] and can therefore be included in the analysis of BCVs. Many verbs and nouns in English, however, are form-identical as in (60) below, and since they show neither affixed verbal morphology nor nominal marking, they could be verbs or nouns. If the complement of *hacer* is a noun, then this could mean that *hacer* is a lexical verb with the noun as its object versus light verb *hacer* with a noun as its complement which serves to specify an event rather than refer to an entity.

60.

Y	me	dice,	ehm,	te	lo	hago	**promise**	que,
and	CL	say	ehm	CL	CL	do1SG.PRES	promise	that

que	si	tú	te	lo	pintas	yo	me
that	if	you	CL	CL	dye2SG.PRES	I	CL

lo pinto.
CL dye1SG.PRES

'And he tells me, ehm, I promise you that, that if you dye yours {hair} I will dye mine'

(BSEK/13WJ/11WJ_12WJ_13WJ/00:03:46)

It is necessary to distinguish BCVs from other Spanish-English constructions that might involve the lexical or main verb *hacer* which has an English-origin noun as an object. English-origin lexical items which are objects of main verb *hacer* would be better analyzed with nouns since such constructions would not be compound predicates. In other words, the surface structure of *hacer* + English-origin noun and *hacer* + English-origin verb may appear to be similar even if the function is different (Fuller Medina 2010; Jenkins 2003). Nouns that do not predicate jointly with a light verb do not fall within the "envelope of variation" for verbs. In addition, because verbs from monolingual Spanish and monolingual English will be used as benchmarks for determining the status of the English-origin item in BCVs, it is important to ensure that English-origin verbal items are clearly

---

[117] Contrast those examples with the nominal forms which are different: *burial*, *competition*, and *creation*.

identified to ensure that the same category is being compared across datasets in the comparative analysis.

To distinguish between these cases where the verb and noun are form-identical, three criteria are used for assessing to the data: (i) direct object taking properties of the construction (ii) referentiality of the English-origin item and (iii) and argument structure of the English-origin item. These criteria are discourse-based, rather than reliant on syntactic tests of grammaticality[118]. Consequently, those examples as in (61) that take a direct object (*los tacos*) were retained as bona fide BCVs given that *taste* is predicating together with *hacer* as evidenced by the direct object.

61. los judge   hacían   **taste**   los   tacos
    the judge   do3PL.IMP   taste   the   tacos
    'The judges tasted the tacos'

                                                          (BSEK/62N/62N/18:26.7)

At this point we still do not know if the English item *taste* is a borrowing or a code-switch but the above examples indicate that BCVs behave as constituents and that the English item (*taste)* is not a direct object, and therefore, not nominal and most likely verbal. Those items that had no direct object as in (62) below still remain ambiguous.

62. después   que   van   a   hacer   **protest,**   que
    after   that   go3PL.PRES   to   do.INF   protest   that

    van   hasta   el   aeropuerto
    go3PL.PRES   up.to   the   airport

    'And then that they are going to protest, that they are going all the way to the airport'
                                                          (BSEK/ 45N/45NJ_43NJ_3/ 35:13.8)

---

[118] See also Jenkins (2003:199-200)

These English-origin items either could not take a direct object, since they are intransitive, or simply did not appear with one[119]. The nouns in these examples were therefore assessed to determine if they were referential and if they have argument structure. Nouns are considered to be referential when they refer to "an object as an object, with continuous identity over time" (Dubois 1980:208 as cited in Torres Cacoullos and Aaron 2003:306). In addition, if the complement of a light/non-light verb is considered to have its own argument structure then it will select for the light counterpart (Wittenberg and Piñago 2011: 395). This means that a noun like *cake* or *orange* would potentially occur with the main verb *hacer* whereas a noun like *shower,* which has argument structure, would potentially occur with the light verb *hacer*[120] (Ibid)*.* If the English item, such as *protest* in (62) was found to be non-referential and also have argument structure then it was also considered verbal and these tokens were included in the analysis of BCVs. None of the tokens showed evidence of referentiality or tracking, i.e., the English-origin lexical item in the BCV was not co-referentially indexed with the same item in the ambient discourse. A summary of the criteria for categorizing the English –origin item in the BCV is presented below[121].

---

[119] Such examples could also take oblique arguments such as a PP complement, in which case they would still be considered verbal.

[120] Admittedly, these are weaker diagnostics but they still depend on the actual nature of the data rather than hypothetical possibilities of whether the constituent can take modification, can occur in particular structures etc.

[121] Object clitics could potentially be used for diagnosing the verbal or nominal status of the English-origin item, however, there seems to be some variability with regard to *le* and *lo* for dative/accusative marking (exemplified in (i), (ii) and (ii)), therefore, this option was not used. Their placement with respect to the complex predicate can still be used as a diagnostic for determining the system membership of the English-origin item since this placement relies on the finiteness of *hacer* regardless of its complement.

(i) Él trabajaba allí … pero LE hicieron **transfer** allí en Corozal, 'He used to work there…. but the transferred him to Corozal' (BSEK/35W/35W/00:22:18/)

(ii) LO hacía **transfer** en mi account, {money} (BSEK/45NJ/45NJ_43NJ_3/01:06:01)

(iii) soy so sure about it que sólo LE hago **blurt-out**. 'but it's something that I am so sure about it that I just blurt it out' (BSEK/08W/08W_2/00:14:46/)

4.3.1.1 Exclusion Criteria

(i) Morphological marking determiner or plural –*s*[122]
(ii) Nominal form (retirement versus retire)

4.3.1.2 Inclusion criteria

    (i)   The English item is clearly verbal *retire* versus *retirement*

When the nominal and verbal item have the same surface form e.g. *promise*, include as a BCV if it:

    (ii)   has a direct object which indicates that it is verbal,
    (iii)   if there is no direct object, then it must not be a referential or a "tracking" noun,
    (iv)   and must have argument structure (compare *cake* to *shower*)

It can be argued that some English-origin items that are marked overtly as nominal (e.g. by determiners) may also be light verb constructions and should be categorized as BCVs. In fact, cross-linguistically *do*-verb constructions have verbal, nominal, and adjectival complements (Muysken 2000 184-220; Romaine 1989). However, I opt to exclude such examples, on the basis of overt nominal morphology, as those constructions with non-verbal complements may exhibit different behavior from BCVs where the complement of *hacer* can be shown to be verbal. In addition, the criteria described above captures 90% of the data when applied to tokens of *hacer* with an English-origin lexical item. BCVs are predominantly of the [Hacer + V] type and, consequently, in keeping with the goals of this thesis to identify and analyze major patterns in the data, it is this subset that is analyzed. I leave the discussion of nominalized BCVs, which are infrequent, for future research.[123,124]

---

[122] Plural –*s* could be used as a diagnostic as there was no pervasive pattern of s-deletion evident on impressionistic analysis and, as will be shown in Chapter 5 Nominal borrowing, speakers make consistent use of –*s* to mark plural on count nouns.

[123] There were twenty tokens of English-origin items either marked with plural –*s* or with a determiner. Some of these English items are objects of *hacer* as a main verb as in (i) and some are possibly light verb constructions as in (ii).
    (i)   Hicimos nuestro **sawdust carpet**. *(*BSEK/58N/58N/09:15.1)
            'We did our sawdust carpet'
    (ii)   Yo los pongo a ellos a hacer **chores** (BSEK/05N14/05N14 /00:08:39)

Assuming that BCVs are compound predicates, if *hacer* is followed by an English-origin multiword (MW) fragment, this construction could potentially be a BCV as well. However, MW fragments and single words have been kept separate in this analysis, so far obscuring any BCVs that may be in the MW subset of data. Consequently, MW fragments consisting of *Hacer* + MW English-origin fragments were examined for BCVs. In keeping with the criteria described above, any examples with nominal marking were excluded (e.g. *hacemos meat pies*) leaving 24 potential BCVs. Most of these consisted of phrasal verbs or particle verbs as in (63) (*hacer set-up*) and thus were considered verbal. The remaining examples such as *hacer project movies or performances* in (63) were subjected to the same tests as described above.

63. sólo   cuando tenemos   eventos,   then   venimos   para   acá
    just   when   have1PL.PRES events   then   come1PL.PRES   to   here

    para   hacer   **set-up screen** y   hacer   **project movies or**
    to   do.INF   set.up   screen   and   do.INF   project movies or

    performances
    performances

    'Just when we have events, then we come here to set up the screen and project movies or performances.'

    (BSEK/58N/04N_58N/00:05:21/)

The tests showed that in examples such as (63), *project,* the English-origin item immediately following *hacer,* is verbal since *hacer project* takes the direct object *movies or performances,*

<hr>

'I put them to do chores'

[124] I have essentially dealt with the "words", or the surface forms, which are the complements of *hacer*. This is in keeping with the discourse-based approach of the variationist method. The underlying structure, however, can be analyzed as a categorizer + root. Roots essentially become "words" when they are combined with adjectival, nominal, or verbal features and, thus, are interpreted as nouns or verbs when they are in the context of certain morphemes (Arad 2003; Marantz 1997; Harley and Noyer 1999). Some roots may be interpreted as two different categories based on different environments which accounts for the same surface forms functioning as noun or verb (to taste, a taste; to protest, a protest). This analysis both accounts for the ambiguity regarding the complements of *hacer* noted above, and the observed differences between Spanish light verb constructions and BCVs. In Spanish light verb constructions, *hacer* takes a nominal complement and in BCVs it takes a verbal one, for the most part. Elsewhere, I analyze BCVs as little vP structures where the complement of hacer is a categorized verbal root because, following Harley and Noyer (1999:2), the nearest c-commanding f-morphemes are little *v*, Aspect, and Tense (Fuller Medina 2013).

therefore, *project* forms part of the verbal complex. Because the object of the BCV is in English, and not in Spanish, on the surface it may appear as an alternation to a multiword fragment at the boundary between *hacer* and *project*. However, when subjected to the criteria described above, examples such as (63) show similar results as the BCVs described above. This suggests that the English-origin verbal item is part of the Spanish verbal predicate and not the ensuing English-origin multiword fragment. Consequently, if there is a code-switch, it is most likely taking place after *project* and *set up,* between the verb and the object. Such alternations have been reported in other bilingual data where the code-switch is considered to occur at the boundary between a borrowing and the remaining multiword fragment rather than at the boundary between the two languages. This is exemplified below for English-Spanish (Torres Cacoullos and Aaron 2003: 320) and Igbo-English (Eze 1997:168, 170). For example, Torres Cacoullos and Aaron 2003:320) suggest that *supervisor*, in (64), may be a nonce loan followed by a code-switch since *supervisor* is not marked by a determiner which is consistent with Spanish but not English grammar. Likewise, Eze suggests that the noun *gate,* in (65) below, patterns with the lone English–origin nouns borrowed into Igbo discourse and that "…nothing indicates that the … nouns are not also borrowings into Igbo… the switch point would be after the noun *gate*, for instance, and not before it as it appears on the surface." (Eze 1997:170)[125].

64. My youngest daughter, ella es ø supervisor at the State Police (190.10)
    'My youngest daughter, she is a supervisor at the State Police

65. I̩ kpochigo ø gate by this time? (5/020)
    You locked gate by this time
    "Have you locked the gate by this time?"

---

[125] See also Romaine (1986:43-44) on a similar analysis. Romaine argues that in the case of Punjabi-English BCVs, the English-origin nouns that appear with BCVs show object marking and gender marking consistent with Punjabi, thus, indicating that they are integrated into Punjabi and therefore complements of the BCVs and not part of the BCV complex. This suggests that single English-origin nouns that follow BCVs in the current data are also separate words subject to indices of integration which apply to nouns.

Therefore, by applying the above-described criteria in a consistent fashion it becomes possible to adhere to the principle of accountability by clearly defining the object of study. The criteria applied are not intended to determine if the English-origin item is a code-switch or borrowing but rather to determine if the English-origin item is part of a BCV and belongs to the verbal or nominal subset of data. Consequently, this goal involves capturing the full distribution of BCVs in Belizean Spanish and further assessing the nature of the BCVs and potential restrictions on their formation.

4.3.2 Distribution, frequency, and productivity of BCVs in Belizean Spanish

Hagerty (1996) noted that BCVs appeared to be frequent in Belizean Spanish though a quantitative distribution of the data was not provided. The current data is therefore examined to determine if BCVs are frequent in Belizean Spanish, and not simply perceptually salient[126]. In addition, BCVs are reported to partake in various structures cross-linguistically (Wilson 2013:129-130; Balam et al. 2014; Fuller Medina 2010, 2005; Muysken 2000) but full distributional data is generally not presented[127], therefore, it is difficult to know how BCVs pattern syntactically or with respect to any restrictions on their use. As a first step in understanding the role and function of BCVs, their distribution across in the verbal paradigm must be mapped out. By so doing, it becomes possible to talk about the properties of BCVs once their syntactic distribution has been observed. The distribution of BCVs also aids in determining how productive their use is (Wilson and Dumont 2014: 8). If BCVs function in similar ways to their simplex monolingual verbal counterparts, then we would expect them to be just as productive and to also be found in a variety of structures, moods, tenses, verb types, and clause structures. Potentially BCVs could be used frequently yet be limited in the

---

[126] Phenomena that are perceptually salient can sometimes appear as frequent because they are noticed more readily.
[127] Balam et al. (2014) do provide limited distributional data as do Wilson and Dumont (2014) and Wilson (2013).

lexical types that occur and restricted to certain contexts. Thus, it is beneficial to assess both their distribution as well as the type/token ratio to determine if they are both frequent and productive (Wilson and Dumont 2014; Fuller Medina 2005a,b; Bybee and Thompson 1997; Hagerty 1996). As will be seen in the discussion on potential aspectual constraints, this distribution will become useful for assessing the presumed grammaticalization of *hacer* in BCVs (Balam et al. 2014; Fuller Medina 2013, 2010; Wilson 2013).

It has been reported that BCVs can be transitive or intransitive (Balam et al. 2014; Fuller Medina 2013, 2010, 2005; Wilson 2013), that they can appear in various tenses and moods (Wilson 2013:130; Fuller Medina 2005:73-74), and in different clause structures (Balam et al. 2014). The data is therefore coded for these factors to establish the distribution of BCVs and the contexts where they occur. The type/token ratio was also determined by coding each English-origin item appearing in a BCV for lexical type and it was found that there were 133 lexical types. The results of this distributional analysis will show if BCVs represent a regular pattern and if they are productive or restricted. Each BCV was coded according to the categories outlined below. An example of each is also included for clarification.

Transitivity

Transitive

66. Hacen   **shake**   el   pan   three times.
    do3PL.PRES shake   the   pan   three times

    'they shake the pan three times.' {explaining game called "kick the pan"}
                                                                    (BSEK/29W/29W/06:51.)

Intransitive

67. yo salí   en   July, yo   hice   **graduate**   en
    I   leave1SG.PRET in   July   I   do1SG.PRET   graduate   in

    September
    September

    'I left in July, I graduated in September'

    (BSEK/05N14/05N14/00:12:06.858)

Tense

Present

68. como   allí   en   Corozal   lo   que   hacen   **mix**
    like   there   in   Corozal   CL   that   do3PL.PRES   mix

    es   el   Creole y   el   English
    be3SG.PRES   the   Creole and the   English

    'like, there in Corozal, what they mix is Creole and English'

    (BSEK/12WJ/11WJ_12WJ_13WJ/03:27.3)

Progressive

69. se escucha   como   que   están   haciendo   **shake**   el,
    CL hear3SG.PRES   like   that be3PL.PRES   do.PTCP   shake   the

    la puerta de   atrás
    the door   of   behind

     'you could hear like they were shaking the, the back door'

    (BSEK/12WJ/11WJ_12WJ_13WJ/50:03.6)

Preterit

70. Porque   nuestra familia es   grande, so   se   hizo
    because   our   family   be3SG.PRES   big   so   CL   do3SG.PRET

    spread-out
    spread.out

    'Because our family is big, so they spread out'

Imperfect

71.

los judge	hacían	**taste**	los	tacos
the judge	do3PL.IMP	taste	the	tacos

'The judges tasted the tacos'

Periphrastic future

72.

después	que	van	a	hacer	**protest,**	que	van
after	that	go3PL.PRES	to	do.INF	protest	that	go3PL.PRES

hasta	el	aeropuerto
to	the	airport

'And then that they are going to protest, that they are going all the way to the airport'

(BSEK/45NJ/45NJ_43NJ_3/ 35:13.8)

Mood

Subjunctive

73.

que	todo	los	estudiantes	se	hagan	**gather**
that	all	the	students	CL	do3PL.PRES	gather

porque	ella	ahorita	va	descubrir	quién	es	ese,	esa
because	she	now	go3SG.PRES	discover.INF	who	is	that	that

traviesa		o	travieso
mischievous.one.FEM		or	mischievous.one.MASC

'That all the students should gather because she was going to find out who is the mischievous one.'

(BSEK/13WJ/11WJ_12WJ_13WJ/35:02.0)

Indicative

74.

yo	salí	en	July,	yo	hice	**graduate**	en

124

I	leave1SG.PRET	in	July	I	do1SG.PRET	graduate	in

September
September

'I left in July, I graduated in September'

<div align="right">BSEK/05N14/05N14/00:12:06.858)</div>

Clause type

Independent clause

75.
los	judge	hacían	**taste**	los	tacos
the	judge	do3PL.IMP	taste	the	tacos

'The judges tasted the tacos'

<div align="right">(BSEK/62N/62N/18:26.7)</div>

Subordinate clause

76.
que	todo	los	estudiantes	se	hagan
that	all	the	students	CL	do3PL.PRES

gather	porque	ella	ahorita	va	descubrir	quién
gather	because	she	now	go3SG.PRES	discover.INF	who

es	ese,	esa	traviesa	o	travieso
is3SG.PRES	that	that	mischievous.one.FEM	or	mischievous.one.MASC

'That all the students should gather because she was going to find out who is the mischievous one.'

<div align="right">(BSEK/13WJ/11WJ_12WJ_13WJ/35:02.0)</div>

Results of the distributional analysis, summarized in Table 13, show that BCVs do not appear to be limited to any particular tense, mood, or structure though the majority of examples occur in the simple tenses of which the present tense (39%) dominates. This is followed by infinitive forms where the BCV is an embedded verb (25%) as in (72) above. They occur mostly in the indicative mood (63%) and are found in independent clauses 68% of the time. BCVs can be both transitive

(57%) and intransitive (43%) as noted earlier. In summary, BCVs in the current data appear in different structures, tenses, and moods confirming previous findings for Spanish-English BCVs (Balam et al. 2014; Wilson and Dumont 2014)[128].

Table 13: Distribution of BCVs across tense, mood, and clause structures

	N	%
TENSE		
Present	70	39
Infinitive	46	25
Preterit	28	15
participle	20	11
Future	11	6
Imperfect	5	3
present perfect	2	1
	182	
MOOD		
Indicative	126	69
Subjunctive	56	31
	182	
CLAUSE TYPE		
Independent	124	68
Subordinate/relative clause	58	32
	182	
TRANSITIVITY		
Transitive	104	57
Intransitive	78	43
	182	

Bybee and Thompson (1997: 384) suggest both token frequency and the type frequency play a role in productivity but that it is type frequency that determines the degree of productivity. "Type frequency refers to the number of distinct lexical items that can be substituted in a given slot in a construction…. And the more lexical items that are heard in a certain position in a construction, the

---

[128] At this time, it is not known if this distribution of BCVs differs from the distribution of monolingual Spanish or English verbs. BCVs, for example, appear 68% of the time in independent clauses but it is not known if this reflects a general tendency for independent clauses to be dominant in discourse in these data or if this finding is specific to BCVs. I leave this for further research.

less likely it is that the construction will be associated with a particular lexical item…". (Ibid). Accordingly, the high type/token ratio as indicated by the 133 types found in 182 tokens, points to BCVs being productive. This type frequency shows that there are 133 distinct lexical items in the current data that can be substituted in the complement slot of *hacer* in the BCV construction. Thus, BCVs are not limited to a few lexical collocations. In addition, this lexical variability and concomitant low rate of recurrence of English-origin verbal items suggests that BCVs may be used with more types than what is captured by the current data as "high type frequency ensures that the construction will be used more frequently … making it more accessible for further use, possibly with new items." (Ibid). All speakers in the sample use BCVs and extensive variability in lexical type is evidenced even with this select group, which together with the distribution summarized in Table 13, further bolsters the finding that BCVs are productive in varieties of Belizean Spanish[129]. With regard to frequency of BCVs, English-origin verbs account for 14% of the single other-language items and 91% of these verbal items are BCVs, thus, Hagerty's (1996) observation that BCVs are frequent in Belizean Spanish is confirmed by means of quantitative analysis. Now that the distribution across of BCVs across the verbal paradigm has been established, the status of the English-origin item as a borrowing or a code-switch is assessed by analyzing conflict sites of verbal morphology and variable clitic placement.

---

[129] In their study on BCVs in New Mexican Spanish, Wilson and Dumont (2014) report that of 24 speakers selected, 12 speakers use BCVs. Since they analyze BCV use among a group of twelve speakers, as in the current study, it provides a point of comparison. Consonant with the findings above, they also report that BCVs are productive in their data though the overall rate and type/token ratio of 73/116 is lower.

4.4. Code-switching or Borrowing

4.4.1 Conflict site 1: Verbal morphology

The conflict site of morphological marking provides a means of determining if English items in Spanish discourse are integrated into Spanish or if they follow an English grammatical system. While both languages affix morphology to the verb, Spanish is considered to be morphologically richer than English as Spanish morphological endings encode person, number, and tense. Comparatively, English affixes tense morphology (*I loved the movie*) but the endings are generally invariant and person and number must be inferred from the subject[130] (Quirk et al. 1985:98), with the exception of the third person singular (*I love, he loves*). This difference is illustrated with a regular *–ar* verb in Spanish in Table 14 below.

Table 14: Verb paradigm for regular verbs in (Latin American Spanish) and English

Present tense		*Bailar*		*To dance*
Singular	yo	bail-o,	I	dance
	tú	bail–as	you	dance
	vos	bail-ás		
	él/ella	bail–a	he/she	dance-**s**
	usted	bail–a		
plural	nosotros/as	bail-amos	we	dance
	ellos/ellas	bail-an	they	dance
	ustedes	bail-an	you-pl	dance
preterite				
Singular	yo	bail-é	I	dance-d
	tú	bail–aste	you	dance-d
	el/ella	bail–ó	he/she	dance-d
	vos	bail-aste		
	usted	bail–ó		
plural	nosotros/as	bail-amos	we	dance-d
	ellos/ellas	bail-aron	they	dance-d
	ustedes	bail-aron	you-pl	dance-d

---

[130] In addition, there are a limited number of English irregular verbs which mark tense by a change in the root (*I eat* versus *I ate*) or number by a change in the root (*I am* versus *you are, he is*).

If single English-origin verbs in the data follow the English system, then we would expect

to see number and tense marked on third person singular forms of verbs in the present tense and –

ed on past tense forms. If the English-origin verbs follow the Spanish system then we would expect

to see person, number and tense expressed in the verbal morphological endings as illustrated

above for *bailar*. Unlike synthetically integrated verbs such as *trayar* which provide clues to system

membership via overt morphology affixed directly to the stem*,* the English-origin verbs in BCVs

show no morphological affixes. In example (77), below, if English-origin verbs were following an

English system we would expect to see 3SG *–s* on the verb *certify* but in fact none of these verbs,

regardless of person or tense, show any English inflectional morphology.

77. él   es   un   professional   make-up   artist   y   él
    he   be3SG.PRES   a   professional   make-up   artist   and   he

    hace   **certify**   a   personas   también.
    do3SG.PRES   certify   to   people   also

    'He is a professional make-up artist and he also certifies people.'

                                                              (BSEK/05N14/05N14 /00:25:59)

The English-origin verbs in BCVs also lack the Spanish morphological affixes, however, these

English-origin verbs are, in fact, marked for Spanish morphology, via *hacer,* and categorically so.

Required Spanish morphology is marked on *hacer* in *hace certify* rather than directly on the verb

certify via affixation. This is in keeping with similar findings reported in variationist studies for other

language pairs[131] as well as for Spanish-English BCVs (Balam et al 2014; Wilson and Dumont

2014; Fuller Medina 2013, 2005a,b; Pfaff 1979). In (77),for example, the first clause is unmixed

speech and shows agreement in number and person between the subject *él,*'he,' and the verb

---

[131] See Sankoff et al (1990) on Tamil-English; Torres Cacoullos and Aaron (2003: 321-323) on Spanish-English light verb constructions; and Eze on Igbo-English (1998:190,1997)

es,'is'. In the same fashion, agreement can be observed between the subject *él* 'he,' and *hace* in *hace certify* 'certifies' in the second clause. Recall that the predication depends on both the light verb and its complement and what is observed here is subject-verb agreement between the subject and the compound predicate, i.e., the BCV. The strategy of borrowing verbs via a *do*-verb is well recognized as a borrowing strategy (Wohlgemuth 2009; Wichmann and Wohlgemuth 2008). These English-origin verbs are not truly "bare" as they are marked for tense, person, number mood etc. and are found throughout the verbal paradigm as evidenced from the distribution presented earlier in examples 66-76 and Table 13 above. Cross-linguistically, the borrowed verbal item is frequently a form that is stripped down to its "bare lexical essentials" but this form is not sufficient for predication; light verb strategies are one way to mark the lexical stem as a verb or full predicate, to "explicitly mark its 'verbness'" (Matras 2007: 48-49). This is exactly what is observed in the current data. Furthermore, Spanish is a language where verb roots cannot surface without being marked morphologically; therefore, when a stem or root from English is integrated into Spanish in the current data it is integrated largely via *hacer*.

In addition, light verb constructions are well documented in monolingual Spanish (RAE 2009; Alonso Ramos 2004; Bosque 2001; Solé 1966) and claimed to be more frequent in Latin American varieties of Spanish (RAE 2009:57). Accordingly, the categorical pattern of TMA morphology being expressed on *hacer* is consistent with monolingual Spanish light verb constructions (with *hacer)* where only the light verb is inflected and not its complement. The main difference is that Spanish light verb constructions tend to take nominal complements (*hacer un llamado, hacer travesura, hacer rezos*). Nonetheless, as with BCVs, the light verb depends on the complement for predication, thus, where the complement is nominal it serves to specify the event. There is no observable evidence of the English system participating in BCVs, as the English-origin items are only etymologically English. This, coupled with the pattern of categorical morphological

130

marking on *hacer*, which is consistent with Spanish, suggests that the English-origin verb can be analyzed as being fully integrated into the Spanish system.

4.4.2 Conflict site 2: Clitic placement

Object pronouns present another conflict site that is used to diagnose the system membership of the English-origin verbs in BCV constructions. Object pronouns, or clitics have been observed to occur with BCVs and have been taken as evidence confirming the functional nature of *hacer* (Wilson 2013: 125; Fuller Medina 2005b: 80) as well as the unithood of the [*Hacer* + V] complex and its integration into Spanish grammar (Wilson and Dumont 2014:8; Fuller Medina 2005b:80). In this section, I analyze all BCVs that occur with clitics and find that BCVs behave in a similar fashion to monolingual Spanish verbs and that the English-origin item is integrated into Spanish. The conflict site between the two languages is exemplified first, including descriptive details of the pronominal systems of English and Spanish. This is followed by a rationale for using object pronouns as a diagnostic of system membership of the English-origin item in BCVs.

4.4.3 Pronominal systems of Spanish and English

The canonical position for English object pronouns is post-verbal as exemplified in (78)[132] ("but I am a human being and *we can do it*"). For clarity, the relevant verb and clitic are underlined in the examples discussed. In contrast, only subject pronouns occur in pre-verbal positions, or what Quirk et al. (1985: 337) call "subject territory". There is some variability in the position of indirect object

---

[132]I leave aside for the time being any variability in object pronoun placement that topicalization might introduce.

pronouns, i.e., they may appear immediately following the verb (*I gave him the book*) or after a full NP direct object (*I gave the book to him*) but they do not precede the verb (Ibid)[133].

78. …but I'm a human being and we can <u>do it</u> under the table.

(BSEK/45NJ/45NJ_43NJ_3/01:34:08)

Spanish, on the other hand, shows variability in the position of object pronouns, also known as pronominal clitics, depending on the finiteness of the verb[134]. When the verb is finite (*metieron*), *as in* (79), the clitic must be preverbal[135] also known as proclisis (*lo metieron*) (RAE 2009: 39). This is in stark contrast to English where object pronouns are post-verbal. When the verb that the object clitic is thematically associated with is non-finite in Spanish, the clitic may appear before the finite matrix verb as in (80) (<u>lo</u> *tenía que mantener*) or it may appear postverbally or enclitically attached to the verb that it is thematically associated with as in (81) (*empezar<u>lo</u>)* (RAE 2009:39; Fernández Soriano 1999). The former case is generally known as clitic climbing, though there is ongoing debate in the theoretical literature to account for the variability in clitic position[136]. What examples (80) and (81) show is that the clitic may appear before or after the verb, variable behaviour not observed in English.

---

[133] Furthermore, when the direct object pronoun rather than a full NP appears (*I gave it to him* vs *I gave the book to him*) the indirect object pronoun is postposed to the direct object, though some varieties of English are said to permit alternative ordering of the direct and indirect objects (*I gave it to him/I gave him it*) (Gast 2007; Comrie 2012). This type of alternation, known as the dative alternation generally occurs with full NPs (*I gave the book to him* vs. *I gave him the book*) (Jackendoff 1990; Larson 1988,1990)

[134] Variationist studies have found a number of linguistic factors affect the position of the clitic including topicality and information flow (Torres Cacoullos and Schwenter 2009) verb type, preceding material (Torres Cacoullos 1999; Davies 1995) as well as negation and animacy (Myhill 1988).

[135] Some contact varieties permit postverbal clitics with finite verbs (see Barnes et al. 2014 on Asturian Spanish who attribute non-standard enclisis with finite verbs to influence from Asturian Spanish where enclisis is the norm). Enclisis also occurs in mainstream Spanish when the verb is a command (*cómelo* 'eat it').

[136]. There are two main approaches that aim to account for clitic position. Under the movement approach, clitics are generated in argument position of the verb and subsequently move to adjoin to the finite verb (Kayne 1975). Clitics can also be analyzed as being base-generated as a constituent of a complex verb rather than in argument position of the verb (Borer 1984; Rivas 1977). In this latter case, the clitic does not move or "climb" from any previous position. Regardless of the analysis, what is observed in the data is variability in clitic position associated with finite and non-finite verbs. It is this patterning that is instructive in responding to the research questions.

79. <u>lo metieron</u>   en   su   carro
    CL put3PL.PRET   in   his   car

    'They put it in his car'

<div align="right">(BSEK/12WJ/11WJ_12WJ_13WJ/00:54:17)</div>

80. cuando   comienza,   ehm,   dar   las   frutas   <u>lo</u>   <u>tenía que</u>
    when   start3SG.PRES   ehm   give.INF the   fruits   CL   have.to1SG.IMP

    <u>mantener</u>   ¿no?   Hacerlo   spray,   después   hacer   harvest,
    maintain.INF   no   do.INF=CL   spray   after   do.INF   harvest

    es   trabajo
    is3SG.PRES   work

    'When it starts, ahm, to give fruit, I had to maintain it, no? Spray it, then harvest, it's work.'

<div align="right">(BSEK/38W/38W_1_2/00:18:59)</div>

81. todavía   no   hemos   hecho eso,   <u>quiero</u>   <u>empezarlo</u>
    still   no   have1PL.PRET   do.PTCP that   want1SG.PRES   begin.INF=CL

    este   año   que   viene,
    this   year   that   come3SG.PRES
    'We still haven't done that, I want to start it this coming year'

<div align="right">(BSEK/45NJ/45NJ_43NJ_3/01:47:51.350)</div>

Spanish has two sets of object pronouns or pronominal clitics. One set is accusative and the other dative, though the forms overlap with the exception of the third person forms and second person plural forms (Real Academia Española (RAE) 2009). These are summarized Table 15 below. English, on the other hand, uses the same set of pronouns (*me, you, him/her/it, us, you, them*) for both dative and accusative objects[137]. Spanish and English also differ in reflexive pronouns as indicated in Table 16 below[138].

---

[137] Additionally, when the indirect object follows the direct object (*I gave the book to him*) rather than the verb (*I gave him the book*) dative case is marked by 'to'.

[138] Spanish has various uses of the pronoun *se*, it may be aspectual, impersonal, reciprocal, passive or an unaccusativity marker (Bruhn de Garavito et al 2002). The various uses of the pronoun *se* are not discussed in detail as the combined data was infrequent (reflexives and impersonal uses were each less than 10% of the data N ≤ 5 for each).

Table 15: Direct and indirect object pronouns in English and Latin American Spanish

	Spanish		English	
DIRECT OBJECT	**Singular**	**Plural**	**Singular**	**Plural**
1st person	Me	nos	me	us
2nd person	te,	los, las	you	you
3rd person	lo, la	los, las	him, her, it	them
INDIRECT OBJECT				
1st person	Me	nos	(to) me	(to) us
2nd person	Te	les (se)[139]	(to) you	(to) you
3rd person	le, (se)	les (se)	(to) him, her, it	(to) them

Table 16: Reflexive pronouns in Latin American Spanish

	Spanish		English	
	Singular	**Plural**	**Singular**	**Plural**
1st person	Me	nos	myself	ourselves
2nd person	te (familiar) se (formal)	se (familiar/formal)	yourself	yourself
3rd person	Se	Se	his/her/itself	themselves

The differences between Spanish and English object pronouns lie in the actual forms of the pronouns (*te* vs *you; te* vs *yourself* for 2SG), dissimilarities in case marking, and the positions of object pronouns relative to the verb of which they are arguments. The conflict site for pronominal placement is summarized in Table 17.

---

[139] Where both dative and accusative clitics appear sequentially, the dative le/les is expressed as *se*.

Table 17: Conflict site between Spanish and English for pronominal placement

	Pronoun position	
	Spanish	*English*
Finite verbs	Pre-verbal/proclisis *Lo metieron en su carro.*	Post-verbal *I felt it.*
Non-finite verbs	Variable: (i) Preverbal *Lo tenía que manterner.* (ii) post-verbal *Queremos empezarlo.*	*We can do it under the table.*

An adjacency requirement holds between the clitic and its host, which may be a main verb or an auxiliary (Franco 2000), thus, there is a clitic-verb dependency between the clitic and predicate (Sportiche 1999:681). Clitics can be considered arguments of the verb[140] and therefore whether an accusative or dative clitic is selected will be based on the argument taking properties of the verb or predicate, which in this case is the complex predicate [*Hacer* + V] where the English-origin verb introduces arguments and *hacer* carries the functional load[141]. The variable position of the clitic is a function of the finiteness of the predicate which is expressed on *hacer*. This verb-clitic dependency can be directly observed in data in the pattern of variable clitic placement and thus serves as a diagnostic in assessing the behavior of BCVs with regard to Spanish or English grammar[142]. Therefore, the prediction is that if the English-origin verbs in BCVs are code-switches then we might expect to see object pronouns expressed in English and, more importantly, we would expect them to occupy slots consistent with the English system rather than the Spanish one. These would

---

[140] For an overview of alternate views of clitics as affixes, case markers, agreement markers and specificity markers see González López (2008).

[141] I do not mean to imply that the dependency is purely syntactic, given that the clitic requires a phonological host and cannot stand alone (RAE and AALE 2010:311; RAE 2009: 1207).

[142] I am essentially treating clitics as bound morphemes due to their affix-like properties (e.g. they cannot stand alone), though it should be noted that there is no general agreement in the literature as to their status since they share some characteristics with lexical words (e.g. their position with respect to the verb is not rigid) (see Ordoñez 2012 and González López 2008 for overviews of this issue).

almost invariably be post-verbal. If the English-origin verbs are integrated into the Spanish system (i.e., borrowings), then the pronominal clitics that are arguments of these English-origin verbs that appear in BCVs should occupy positions in accordance with the Spanish verbal system and, consequently, reflect that relation between object pronouns and the verbs of which they are arguments. In other words, given that *hacer* carries the functional load in BCVs, when it is finite the clitics should appear pre-verbally and when non-finite variability should be observed in post/pre-verbal position as in Spanish. In (82) below, for example, the clitic appears post-verbally where *hacer* is non-finite and in preverbal position in (83) where *hacer* is finite.

82. No tienes que   hacerlo   **mix**
      no   have.to2SG.PRES   to   do.INF=CL   mix

'You don't have to mix it.'
{speaking about language}

(BSEK/62N/62N/00:21:51/)

83. pero   uno   lo   hace   **mix**,   no   sé
      but   you   CL   do3SG.PRES   mix   no   know1SG.PRES

'but you mix it, I don't know'
{discussing language}

(BSEK/11NJ/11NJ_12NJ_13NJ_2/00:03:18)

4.4.4 Coding protocol

Fifty-five (30%) of the 182 tokens of BCVs appear with object pronouns. These were coded for clitic type, finiteness of *hacer* (finite/non-finite), and position of the clitic (pre/post-verbal). Note that the tokens are coded first for presence of object clitics and then once those tokens that appear with object clitics are isolated, then they are coded for clitic position. No decision is made regarding the status of the English-origin item, only once the pattern of clitic placement has been ascertained for all three datasets (mixed, Spanish and English) can a comparative analysis be

carried out to determine if the English-origin items pattern with Spanish or English. Each of these factor groups is exemplified below where the relevant section is underlined:

Finite use of hacer, object pronouns, proclisis

84. tal vez   si   querés   pasar   bastantes   cosas así,
    maybe   if   want2SG.PRES bring.INF   a.lot   things like.that

<u>te hacen</u>   **charge** un   fee para pasar   ese y   así
CL do3PL.PRES   charge a   fee to   bring.INF that and like.that

'Maybe if you want to bring over a lot of things, like that, they charge you a fee to bring that over and like that'

(BSEK/29W/29W/00:15:27)

85. entraba   allá y   lo   hacía   **transfer**   en   mi
    enter1SG.IMP there and CL do1SG.IMP transfer   in   my

account yo   hacía   el   cheque   y   se   lo
account I   do1SG.IMP the   cheque   and CL   CL

daba   a   Greg
give1SG.IMP to   Greg

'It would go in there and I would transfer it, I would do the cheque and I would give it to Greg' {context: transferring money}

(BSEK/45NJ/45NJ_43NJ_3/01:06:01/)

Non-finite[143] use of hacer, reflexive pronouns

Enclisis

86. para   que   yo   pueda   <u>hacerme</u>   **certify** como un
    so   that I   can1SG.PRES do.INF=CL   certify as   a

professional   make-up   artist
professional   make-up   artist

---

[143] With the exception of two cases (an adverbial participle and a relative clause), when the BCV is non-finite it is embedded and the matrix verb is finite. The matrix verbs were *tener que* (N=2), *poder* (N=1), *querer* (N=1), *tratar de* (N=1), *acabar de* (N=1) *ir a* (N=1).

'So that I can get certified as a professional make-up artist'

Proclisis

87.
me quiero		hacer	**certify**	como	un	make-up	artist
CL want1SG.PRES		do.INF	certify	as	a	make-up	artist

'I want to get certified as a make-up artist'

Additionally, there were three tokens that had both the direct and indirect object clitic, as exemplified in (88) below. These were coded separately and will be addressed in the results section.

88.
Y	me	dice,	ehm,	te	lo	hago	**promise**	que,
and	CL	say	ehm	CL	CL	do1SG.PRES	promise	that

que	si	tú	te	lo	pintas	yo	me
that	if	you	CL	CL	dye2SG.PRES	I	CL

lo	pinto.
CL	dye1SG.PRES

'And he tells me, ehm, I promise you that, that if you dye yours {hair} I will dye mine'

4.4.5 Results

A summary of the distribution of pronominal object clitics is presented below in Table 18.

Table 18: Distribution of object pronouns associated with BCVs

Object Pronoun	N	%
le	6	11
lo	37	67
Reflexive or other *se*	9	16
Other	3	5
Total	55	100

There were 45 finite BCV (i.e., where *hacer* was finite). The corresponding object pronouns were categorically pre-verbal. With regard to non-finite BCVs, clitics appeared both pre-verbally and post-verbally with a preference for the former. This variability is exemplified in (89) where the clitic is post-verbal and (90) where it appears pre-verbally with the same BCV (*hacer certify*) used by the same speaker. Results are summarized and compared to Spanish monolingual verbs in Tables 19 and 20 below.

89. para   que   yo   pueda   <u>hacerme</u>   **certify** como   un
     so   that   I   can1SG.PRES   do.INF=CL   certify   as   a

     professional   make-up   artist
     professional   make-up   artist

     'So that I can get certified as a professional make-up artist'

                                             (BSEK/05N14/05N14/ 00:26:25)

90. <u>me quiero</u>   hacer   **certify** como   un   make-up   artist
     CL   want1SG.PRES   do.INF   certify   as   a   make-up   artist

     'I want to get certified as a make-up artist'

                                             (BSEK/05N14/05N14/00:25:46/)

Variability in proclisis and enclisis is found for both non-finite monolingual Spanish verbs and BCVs as shown in Table 20. Variability is part of the normal Spanish pattern, as exemplified in examples (92)-(93) and in Table 20 below. While the rates themselves differ, the ranking of preverbal versus postverbal position is the same for both datasets. Both Spanish and mixed data show the same preference for preverbal position of the clitic when the verb is non-finite at rates of 88% and 60% respectively. There was no statistically significant difference found between the two datasets with regard to this observed variability. Furthermore, while the tokens are few, the enclisis observed in (89) *(pueda hacerme* certify), for example, is consistent with at least one study on Latin American

Spanish. Myhill (1988:233-235) found that clitic climbing in Latin American prose is rare in affirmative contexts with the matrix verbs *such* as poder, when it means ability (instead of possibility). Thus, this post-verbal position is unrelated to the English system which places object pronouns post-verbally. In fact, there is no evidence that the verb's pronominal arguments are positioned in keeping with the English system.

Table 19 presents the distribution of object pronouns for finite BCVs and verbs in monolingual Spanish. As noted above, when the verb is finite in Spanish, as in (91) (*metieron*), the clitic position must be preverbal and this is evidenced in the data shown below. The object pronouns associated with finite BCVs mirrors the Spanish pattern for finite verbs as all object pronouns are categorically pre-verbal.[144]

Table 19: Proclisis/enclisis when the verb is finite

Clitic position	*Hacer* in a BCV N %	Spanish verb N %
Pre-verbal	45 100	83 100
Post-verbal	0	0
	45	83

91. <u>lo   metieron</u>   en   su   carro
     CL put3PL.PRET   in   his   car

   'They put it in his car'

                                                        (BSEK/12WJ/11WJ_12WJ_13WJ/00:54:17)

Variability in proclisis and enclisis is found for both non-finite monolingual Spanish verbs and BCVs as shown in Table 20. Variability is part of the normal Spanish pattern, as exemplified in examples (92)-(93) and in Table 20 below. While the rates themselves differ, the ranking of preverbal versus postverbal position is the same for both datasets. Both Spanish and mixed data

---

[144] In order to extract a comparison dataset, the first 10 Spanish verbs that had clitics associated with them were extracted for each speaker from the same twenty-minute segment that was used to determine the language-mixing rate and from which monolingual Spanish nouns were extracted.

show the same preference for preverbal position of the clitic when the verb is non-finite at rates of 88% and 60% respectively. There was no statistically significant difference found between the two datasets with regard to this observed variability. Furthermore, while the tokens are few, the enclisis observed in (89) (*pueda hacerme certify*), for example, is consistent with at least one study on Latin American Spanish. Myhill (1988:233-235) found that clitic climbing in Latin American prose is rare in affirmative contexts with the matrix verbs such as *poder,* when it means ability (instead of possibility)[145]. Thus, this post-verbal position is unrelated to the English system which places object pronouns post-verbally. In fact, there is no evidence that the verb's pronominal arguments are positioned in keeping with the English system.

Table 20: Proclisis/enclisis when the verb is non-finite

	Hacer in a BCV N %		Spanish verb N %	
Pre-verbal	6	60	21	88
Post-verbal	4	40	6	22
	10		27	

X^2 (36)=2.16; *p* = .14

92.
cuando	comienza,	ehm,	dar	las	frutas	lo	tenía que
when	start3SG.PRES	ehm	give.INF	the	fruits	CL	have.to1SG.IMP

mantener	¿no?	Hacerlo	spray,	después		hacer	harvest,
maintain.INF	no	do.INF=CL	spray	after		do.INF	harvest

es	trabajo
is3SG.PRES	work

'When it starts, ahm, to give fruit, I had to maintain it, no? Spray it, then harvest, it's work.'

(BSEK/38W/38W_1_2/00:18:59)

---

[145] According to Myhill's (1988) data clitic climbing (CC) "occurred seven out of forty-six times when the finite verb was *tener (que),* which indicates obligation. Six of the seven cases of CC were very similar in meaning: a person was forced to take an action resulting in permanent separation from someone (four cases involved moving away to live somewhere else, and the other two cases involved abandoning a child). … there was an overwhelming statistical tendency to use CC with tener if and only if this meaning was conveyed" (Myhill 1988:233). These results should nonetheless be interpreted with caution since the data is taken from literary works.

93.

todavía	no	hemos	hecho	eso,	quiero	empezarlo	este
still	no	have1PL.PRES	do.PTCP	that	want	start.INF=CL	this

año	que	viene,
year	that	come3SG.PRES

'We still haven't done that, I want to start it this coming year'

(BSEK/45NJ/45NJ_43NJ_3/01:47:51/)

Returning now to the BCVs where both the accusative and dative clitics preceded the verb, these tokens are examined to determine if their ordering is consistent with Spanish. When multiple pronouns precede the verb, they follow a canonical hierarchical order: dative pronouns precede accusative ones (RAE 2009:1230). Consistent with this norm, this dative-accusative ordering of clitics is observed in the monolingual portions in the example below (*te lo pintas*; *me lo pinto*). Where both dative and accusative clitics precede BCVs, the clitics also follow the dative-accusative ordering mirroring the ordering in monolingual Spanish.

94.

Y	me	dice,	ehm,	te	lo	hago	**promise**	que,
and	CL	say	ehm	CL	CL	do1SG.PRES	promise	that

que	si	tú	te	lo	pintas	yo	me	lo
that	if	you	CL	CL	dye2SG.PRES	I	CL	CL

pinto.
dye1SG.PRES

'And he tells me, ehm, I promise you that, that if you dye yours {hair} I will dye mine'

(BSEK/13WJ/11WJ_12WJ_13WJ/00:03:46)

The variability in clitic placement is also indicative of several features of BCVs. First, their variable positioning when *hacer* is finite in BCV constructions, as exemplified in (92) above (*lo tenía que mantener*) and (95) below (*empezarlo*), parallels the variability found with simplex Spanish verbs

pointing to the constituent or "fused" nature of the BCVs (Wilson and Dumont's 2014)[146]. Coupled with the data, presented previously, which shows that hacer predicates jointly with the English-origin verbal item, this is strong evidence for BCVs as constituents and as the functional equivalents of simplex verbs. It also indicates that BCVs are functionally parallel to the synthetic strategy of borrowing verbs where morphology is attached directly to the English–origin verb stem. Finally, the fact that the clitic attaches to *hacer* further confirms its status as functional (Wilson and Dumont 2014; Fuller Medina 2013, 2010, 2005b).

95. todavía   no   hemos   hecho   eso,   quiero empezarlo   este
    still   no   have1PL.PRES do.PTCP   that   want   start.INF=CL   this

    año   que   viene,
    year   that   come3SG.PRES

    'We still haven't done that, I want to start it this coming year'
                                                    (BSEK/45NJ/45NJ_43NJ_3/01:47:51/)

The above results show that the English-origin verbal item is integrated into Spanish; therefore, it is unlikely that it is a code-switch or produced from a "third grammar." However, recall that some researchers also suggest that BCVs can be analysed as instances of creolization (Gardner Chloros 2009; Pfaff 1979), or that they exhibit Creole features and might even be undergoing a process parallel to the development of the Palenquero habitual marker *asé* (Smith 2013), presumed by some to derive from *hacer* (Balam et al. 2014). The comparison of mixed discourse to Creole languages, in particular, Spanish-English data is not a new one (Lipski 2015, 2008) and the results of the preceding analyses together with documented features of Belizean Spanish can help shed light on the view of BCVs as an outcome of creolization. However, rather

---

[146] They cite Torres Cacoullos (2009:153) who, in a study of [auxiliary + gerund] constructions, argues that as the construction becomes fused it begins to behave as a unit instead of a periphrasis consisting of two main parts with the clitic attached to the non-finite verb. Clitics "climb" to a proclitic position mirroring their behavior with finite (simplex) verb forms.

than taking up the general claims of creolization, which, as shown below, are often vague, I aim to address the specific arguments put forth in support of creolization.

Gardner-Chloros (2009:34), for example, considers BCVs to be a construction that lies between code switching and creolization. She also considers them to show "features of creolization as they involve grammatical convergence and an analytic approach to vocabulary" (Gardner-Chloros 2010:198). On the other hand, because Pfaff (1979:301) analyzes BCVs as derived from the Spanish causative construction, she considers them to reflect a process of "semantic extension from an existing syntactic structure … reminiscent of the expansion process of creolization." For their part, Balam et al. (2014) take up Gardner-Chloros' argument of BCVs as an "analytic approach" and consider BCVs to more succinctly express meaning as compared to monolingual Spanish counterparts [147]. They also point to the preverbal position of *hacer* coupled with phonological reduction (word final *r*-deletion) to support claims of creolization. They compare *hacer* in BCVs to the Palenquero pre-verbal aspectual marker *asé,* which is potentially derived from *hacer*[148], due to the similar surface forms as a result of r-deletion on *hacer* in BCVs. Additionally, they consider *r*-deletion a Creole-like feature (Balam et al. 2014:259) [149]. As noted, the vagueness of the claims, and a concomitant lack of clarity with regard to what these various authors mean by creolization, makes a creolization hypothesis difficult to test empirically. However, the more specific arguments that these authors appeal to, such as the morpho-phonological reduction of *hacer* can be addressed.

First, Pfaff's notion of semantic expansion of the Spanish causative structure to BCVs is not supported by the growing consensus in the generative literature on the structure of BCVs as

---

[147] Balam et al. (2014) cite *hacer agree* versus *ponerse de acuerdo* 'to agree'. I leave this point aside given that there are numerous examples (e.g. *hacer promise* versus *prometer* 'to promise') where the Spanish equivalent is a simplex verb rather than a periphrastic one and it is unclear why a BCV, which is a periphrastic form, would be considered more succinct than the simplex verb.

[148] There is not general agreement that this is the case (see Smith 2013 and references therein).

[149] Balam et al. (2014) hold that *r*-deletion on infinitives is characteristic of Spanish Creoles and Afro-Hispanic varieties.

light-verb constructions analogous to monolingual light verb constructions (Rao and Den Dikken 2014 on English-Telugu; Fuller Medina 2013, 2010 on Spanish-English; González-Vilbazo and López 2011 on Spanish-German; Nakajima 2008 on Japanese-English) and is not based on the Spanish causative (Fuller Medina 2005a,b)[150]. The question of BCVs falling somewhere in between code-switching and creolization has been addressed in the previous sections by showing that the English element is best analyzed as a borrowing and not a code-switch.

In terms of phonological reduction and the similarities to Palenquero *asé,* the available data simply does not provide clear support at this time. While phonological reduction may be observed when an item moves from lexical to more functional use (Campbell 2013:131), which might be expected when a full verb shifts to an invariant tense-aspect marker, a general pattern of word final *r*-deletion, not limited to verbs, has been documented in Belizean varieties of Spanish. Hagerty (1979) first noted this feature and it is also reported by Balam (2013:296) to occur at a rate of 59% in Northern Belizean Spanish. Impressionistic analysis of the current corpus reveals a similar pattern even with mono-syllabic verbs such as *dar* 'to try'. Thus, the phonological reduction observed with *hacer* appears to be part of a pervasive pattern in varieties of Belizean Spanish making it difficult to distinguish between regular phonological reduction and a creolization or grammaticalization process. The similarity to *asé* appears to be superficial.

In terms of analogy to *asé* as a preverbal marker, if *hacer* is undergoing morph-phonological reduction into a general preverbal marker we might expect to see evidence of decategorialization as it grammaticalizes, i.e., the loss of "morphological and syntactic properties that would identify it as a full member of a major category such as noun or verb" (Hopper and

---

[150] Moravscik (1975) was one of the first to suggest that BCVs consisted of what she termed a "generic" verb and a "foreign" element. For variationist studies that have uncovered BCVs as minor patterns in their data see Sankoff et al. (1990) on Tamil-English; Torres Cacoullos and Aaron (2003:321-323) on Spanish-English light verb constructions using various light verbs with nominal complements; and Eze on Igbo-English (1998:190,1997).

Traugott 2003:107). This would mean that in the synchronic distribution, at least some occurrences of invariant forms of the verb which lack agreement in person and number or lack tense and mood morphology should be present. The distribution of BCVs presented previously (examples 66-76 and Table 13 above) in fully conjugated forms with the requisite morphological markings for person, number, tense, and mood in different clause structures shows that *hacer* has not lost its morphological and syntactic properties. This suggests that BCVs are not undergoing morpho-phonological reduction or any visible change which might lead to a an invariant grammatical morpheme similar to *asé*[151]. The verb *hacer* in BCVs still appears to be a full-fledged Spanish verb and does not appear to function like *asé*, an invariant, optional preverbal marker.

Caution is warranted when drawing comparisons between creolization and bilingual discourse given that the terms creole and creolization are often used to refer to outcomes of contact and language mixing without any process of true creolization having taken place (Lipski 2014, 2008). This effectively renders the term vacuous. Furthermore, Spanish-English mixed discourse need not be viewed as exceptional but rather as a normal process of utilizing the linguistic mechanisms of Spanish, in the case of borrowing, for example, to adapt and integrate English concepts and lexical items to Spanish (Otheguy and Stern 2011: 86, 92). It appears that this is exactly the case for BCVs.

The potential Kriol influence on varieties of Belizean Spanish, in particular in mixed discourse, is certainly an appealing possibility and an obvious one that should be investigated as much research is needed in this area. Spanish is in contact with Kriol in Belize and this fact was evidenced in the current in the form of overt Kriol features in mixed speech. However, Kriol influence (and any concomitant connection to grammaticalization or creolization) remains to be

---

[151] It should also be noted that BCVs with *hacer* in the infinitive form, the only form eligible to undergo phonological reduction leading to similarities with *asé*, only occurs at a rate of 28% of BCVs in the current data.

empirically tested as contact-induced effects should not be assumed to be inevitable (Poplack and Torres Cacoullos 2015; Poplack and Levey 2010). Furthermore, it would be important to distinguish general processes of language change and variation from what constitute true creole features and processes (Poplack and Tagliamonte 2001:5). Establishing the underlying grammar of Belize Kriol and clear operationalization of creolization would provide the basis for a comparative analysis to reveal if BCVs indeed evidence features of Belize Kriol grammar or processes deemed to be specific to creole languages. For example, a comparative quantitative analysis comparing the rates and constraints governing phonological reduction on *hacer* as compared to the general patterns in Spanish, and in particular, to non-BCV uses of *hacer* could potentially aid in illuminating whether or not a distinct process of morpho-phonological reduction is taking place with *hacer* in BCVs. I leave this for further research.[152]

In summary, the categorial status of the English-origin items in BCVs was shown to be verbal. Two diagnostics of verbal morphology and variably clitic placement showed that the English-origin verb was integrated into the Spanish grammatical system. Spanish morphology was expressed categorically on *hacer* and there was no evidence of English verbal morphology and the variability in clitic position with regard to BCVs paralleled the variability found in monolingual Spanish. When the BCV is non-finite (i.e., *hacer*) the clitic either 'climbs' to proclisis with the finite matrix verb or can attach to the non-finite verb and when *hacer* was finite the clitic was

---

[152] There are two other possibilities worth noting, found in English and in Kriol, which may account for the origin of the BCV structure. The Kriol verbal system consists of bare verbs and variable use of invariant preverbal markers. The [preverbal marker + V] frame would appear similar to the BCV construction in the same fashion as the English [do + verb] frame. English has a long history of *do* support and this structure may have been calqued to form the [*hacer* + V] structure of BCVs. And, in fact, either case would be a plausible scenario under some analyses of language contact which propose that structures can be borrowed or transferred particularly under circumstances of prolonged contact (Thomason and Kaufman 1988). While it is possible that the presence of these structures in the languages in contact with Spanish could contribute to the choice of BCV as the preferred strategy for borrowing verbs, and this remains to be empirically tested, neither of these frames is likely the basis for BCVs. First, Spanish borrows verbs in this same fashion from other languages in cases where there is no contact with English or Kriol (cf. Suárez-Molina (1996) on Spanish-Maya BCVs). Second, as noted previously, BCVs are prevalent cross-linguistically and, therefore, cannot be attributed to English or Kriol structures.

categorically preverbal. Again there was no evidence that the English-origin verb behaves according to English grammar with respect to the verb-clitic relation. The object pronouns never appeared post-verbally as independent words and were never expressed in English. They only appeared post-verbally when the Spanish verb permitted it, i.e., when it was non-finite. Since BCVs are consistent with Spanish grammar in these two areas, then it is also unlikely that BCVs are manifestations of some third system, third grammar or "bilingual syntax" (Muysken 2000), code-switches, or manifestations of creolization. BCVs appear to be nothing more than Spanish speakers making use of the resources of Spanish to borrow English verbs showing "clear mastery of Spanish lexical and syntactic resources" (Otheguy and Stern 2011: 92). By operationalizing clear criteria for distinguishing code-switching and borrowing and by setting the communities' vernaculars as the benchmarks for quantitative comparison, the results of the investigation of BCVs converge in indicating that the English-origin components of these constructions behave as borrowings.

4.5 Aspectual Constraints and Grammaticalization

The second objective of this chapter was to determine if stativity conditions the formation and use of BCVs. The same principles and tools of the variationist method used in the preceding analyses can be brought to bear on this question, first, by clearly operationalizing aspectual criteria so that statives are operationally defined, and second, by implementing a quantitative analysis of the distribution of aspectual classes of the English-origin verbal item. As noted earlier in Chapter 2, very little quantitative work has been done on BCVs, therefore, the nature of BCVs is not fully

understood[153]. Therefore, in the ensuing sections lexical aspect is operationalized, stative verbs circumscribed and data is analyzed to determine if there is a restriction on the class of English-origin verbs that can be borrowed via BCVs. If stative English-origin verbs are not found in the distribution of BCVs then this points to the possibility that statives cannot be borrowed via BCV. Given that diachronic patterns of a grammaticalizing construction or the lexical history of an item can be found in its synchronic distribution (Torres Cacoullos and Walker 2009a,b), the distribution of English-origin verb classes in BCVs should give insight as to the nature of *hacer* and its compatibility with statives in BCV constructions. This analysis of the distribution of stative and dynamic verbs is a first step in testing aspectual restrictions.

The hypothesis that BCVs are aspectually restricted has its basis in both Spanish-English data and cross-linguistic data. In the Spanish-English literature the BCVs reported have tended to be with dynamic verbs (e.g. Jenkins 2003). Silva-Corvalán (1994:47-48) suggests that in the verb system that has developed for Spanish-English speakers in Los Angeles, the perfective/imperfective distinction has eroded for statives while non-statives maintain the distinction and appear in BCVs in the preterite and [estar + participle] constructions in the imperfect. Moreover, Jenkins (2003:202) suggests that the choice of *hacer* as a helping verb (cross-linguistically) may be attributed not only to its functional properties but also to its lexical nature "since the semantic function of any non-stative verb relates to *doing* something" (Jenkins 2003:202-203). Further evidence to suggest that BCVs may be aspectually restricted comes from cross-linguistic data, where some languages use a *do*-verb for more dynamic verbal borrowings and a second 'helping' verb for stative or less dynamic verbal borrowings (see Chatterjee 2014 on Bengali-English; Romaine 1995 for Punjabi-English; for an overview of language pairs see Matras

---

[153] As the discussion in Chapter 2 showed, much work cross-linguistically has been qualitative and work on Spanish-English BCVs currently constitutes a relatively small body of work where few studies to date, if any, employ the quantitative methods of the comparative variationist approach.

2007:47-48; Muysken 2000:184-220). Finally, the assertion that *hacer* is grammaticalized or is grammaticalizing also raises the question of aspectual constraints (Balam et al., 2014; Wilson and Dumont 2014; Fuller Medina 2013, 2010; Wilson 2013; Jenkins 2003). Grammaticalization generally refers to the process where a lexical item (or construction) develops grammatical function (Lehman 1982) or moves along a cline to becoming more grammatical over time (Hopper and Traugott 2003). Even as an item becomes more grammatical, however, "older meanings may still continue to constrain newer "emptier" ones," i.e., bleached forms (Hopper and Traugott 2003:3,76). Thus, *hacer* may retain some of its lexical meaning of *do* which may constrain its functional properties making it incompatible with the holding of a state. In addition, as an item becomes more functional, this change is generally observed with activities first spreading to states last (Bybee et al. 1994) [154].

Addressing the question aspectual restrictions also sheds light on how BCVs might be constrained. On the one hand, we know that monolingual light verb constructions, at least in English, show fairly consistent distributional patterns with regard to the light verb's complements. The light verb *do* in English takes events as its complements; *make* takes 'things' as complements, and when the complement of the light verb is a state, then the light verb can be *make* or *cause* (Harley 2005:64). The structure for these types of constructions is considered to consist of a little *v*P[155] and VP structure, where the light verb occupies the spec of little *v* and the complement occupies V[156]. This structure is considered to be the same for simplex verbs in which case there is no spell out of little *v* (in languages such as English and Spanish). Additionally, little *v* has different

---

[154] Where verbs become auxiliaries, for example, they often retain some of their lexical meaning. This is well-documented in the grammaticalization of future temporal reference forms in Spanish where the periphrastic future [ir + V] disfavors stative verbs due to the retention of the motion allative meaning of *ir* 'to go' (Torres Cacoullos 2015; Aaron 2006; see also Hopper and Traugott 2003:68-69 for the English "go" future).

[155] For further details on the little *v* hypothesis see Chomsky (2001, 1995). For details on BCVs as little vP structures see González-Vilbazo and López (2012, 2011) for German-Spanish, Nakajima (2008) for Japanese-English and Fuller Medina (2013, 2010) for Spanish-English.

[156] For details on the compositional nature of verbs see Harley (2009), Ramchand (2007) and (Cuervo 2003:11-32)

"flavors" or types that are not compatible with all Vs (Harley 2009; Folli and Harley 2007; Cuervo 2003), thus, accounting for the distributional patterns of light verbs and their complements (one *does a dance* but not *makes a dance*). It follows then, that if BCVs are little *v*P structures, as has been proposed in the literature (Fuller Medina 2013, 2010; González-Vilbazo and López 2011; Nakajima 2008), then there may be restrictions on the English-origin verb types in BCVs. If there are restrictions on the type of verb that can be a complement of *hacer* in BCVs then this may be evident in patterns that emerge from quantitative analysis of BCVs.

Two studies have examined aspectual restrictions on BCVs. Fuller Medina (2005), motivated by the rarity of stative BCVs reported in the Spanish-English literature, and Jenkin's (2003) observations regarding the lexical nature of *hacer* being used to form complex non-stative predicates, tested aspectual restrictions on BCVs in northern and western Belizean Spanish (Fuller Medina 2005a,b). Both production data and results from judgment tasks were analyzed. A categorical preference for dynamic verbs in BCVs was found in production data (elicitation tasks)[157], while analysis of the judgment tasks found no statistically significant difference in the acceptability of [+/-dynamic] verbs in BCVs (Fuller Medina 2005a). Nevertheless, the dataset was small and aspect was operationalized in general terms by appealing to Vendler's (1957) four-way classification (states, accomplishments, activities, and achievements) and reduced to a binary factor group [+/-dynamic] for both production data and judgment tasks.

In a second study testing the same hypothesis, and which also used both production and judgment-task data, 5% (N=553) of the BCVs in the production data were reported to be stative (Balam et al. 2014). Results of the judgment tasks produced no statistically significant difference with regard to statives, a finding consistent with Fuller Medina (2005a). It should be noted, however, that in this second study the production data was reanalyzed so that statives were

---

[157] Two examples of statives are reported in Fuller Medina (2005b) which were gleaned from participant observation.

combined with a number of other categories including change of state verbs, reverse psych verbs, psychological and perception verbs, and a category 'other' (9%) that consisted of "verbs difficult to classify" (Balam et al. 2014:255) [158]. Once these disparate categories were combined into the category "stative," they accounted for 50% of the data which led Balam et al. (2014) to claim that there are no aspectual restrictions on BCVs. Yet, the verb categories they collapsed into one category are not all statives[159]. Stative verbs are characterized by homogeneity rather than changes (Rothstein 2008, Levin 2009) therefore the inclusion of change of state verbs is questionable, as is the inclusion of miscellaneous verbs from their category labeled "difficult to classify"[160]. In the end, because neither Balam et al. (2014) nor Fuller Medina (2005a,b) operationalized lexical aspect in a well-defined or motivated fashion and because there were slight differences in what was considered stative, the results of these studies fail to elucidate the question of aspectual restrictions on BCVs. This is not to say that operationalizing aspect is a straightforward endeavour. There remains much discussion in both the theoretical (Rothstein 2004) as well as in the variationist literature (Walker 2010a) as to how aspect should be approached; thus, it remains an elusive area of the grammar (LeBlanc 2010:66). Accordingly, the issues highlighted in the previous studies on BCVs reflect, in part, the difficulty in defining lexical aspect in general, and stative verbs, in particular. Some of these difficulties are discussed below, followed by the solution adopted for the current analysis. Consequently, in addition to testing aspectual

---

[158] This was done presumably because they consider all these verbs to lack agentive action. States generally have an experiencer as their subject but not in all cases. This highlights the need for operationalizing aspect and providing clear definitions of what will be considered stative *a priori*.

[159] The authors also provide few details as to the rationale for the verb classes they choose. Citing Aaron's (2010) study on future reference they claim to choose verb classes in keeping with the variationist literature (Balam et al. 2014:249). On the contrary, however, there is no standard in the variationist literature for verb classes. Factors are based on hypotheses being tested, and in the case of future temporal reference, motion versus non-motion verb categories, for example, are relevant for analyzing the development of the "go" future because of the motion meaning encoded in *go* which constrains in the initial stages the type of verb that can occur in a periphrasis with *go*.

[160] See also Marín and McNally (2011) who argue that Spanish reflexive psych verbs, often assumed to be stative, do not all denote states; the *aburrirse* ('to be bored') class does, while the *enfadarse* ('to get mad') class is best analyzed as achievements. Such a study lends support as well to the analyses of statives not forming a homogenous class and lying instead on a continuum (e.g. Levin 2013).

constraints, a second objective is to provide a transparent and replicable methodology for carrying out this type of analysis with production data[161].

4.5.1 Operationalizing lexical aspect

One of the difficulties with stativity is that most discussions about statives revolve around prototypical statives such as *know* and *love,* yet statives are not necessarily a homogenous class of verbs[162] and involve a range of verbs often not considered. In fact, Bach 1986 (as cited by Rothmayr 2009:4) divides statives into static and dynamic states. States might be thought of as lying on a continuum: prototypical verbs such as *know* may lie on one end and verbs such as *wonder* termed more "dynamic" might be closer to the other end of the continuum presumably because such verbs require that the subject actively expend energy (Levin 2013).

Second, and perhaps because of the reasons outlined above, many of the tests, such as compatibility with the progressive, aimed at identifying statives, fail to consistently do so (Levin 2009). Some statives actually do occur in the progressive (Dowty 1979) despite being conventionally considered to resist occurring in the progressive (Biber et al. 1999). In fact, most tests indicate where statives do not occur rather than where they do (Olsen 1997: 36). Furthermore, while I assume lexical aspect to be a property of verbs, this is also a source of disagreement in the literature[163].

The prevalent approach in the literature is that the aspectual characterization of a sentence is built compositionally. Information about the internal composition of an event is obtained via the interaction of the verb, its internal arguments, and adjuncts. Lexical aspect, at the level of the verb, can be considered the first building block (Sharma and Dao 2010; Rothstein 2004; Olsen

---

[161] I leave aside the issue of judgment tasks for now as I have discussed some of the general challenges in Chapter 2 and I also discuss elsewhere issues specific to the use of this tool in the Belizean context (Fuller Medina 2005b: 51, 88). The approach that is proposed here for the production data could potentially still be applied to design of judgment tasks if such a task were to be employed by other researchers.

[162] For a full discussion of stative verbs see Rothmayr (2009)

[163] See discussions in Rothstein (2004:29-33) and Olsen (1997:15)

1997; Smith 1991). At the VP level, the verb along with its internal objects are determinants of the telicity of the event (Jakendoff 1996; Krifka 1989; Tenny 1992; Verkuyl 1972 among others). Finally, at the level of sentential aspect or viewpoint aspect all of the previous elements (verb, objects, adjuncts) interact with the progressive or perfective to determine the overall aspectual reading of the sentence (Smith 1991). Of course, pragmatic factors also enter into the composition of aspectual meaning where telic readings can be coerced from atelic verbs (Rothstein 2004). Because the hypothesis being tested here only concerns the aspectual properties of the lexical item that can MERGE with *hacer*, I do not address other levels of aspect. These are not immediately relevant due to the fact that at the VP level the internal arguments of the verb are considered to delimit telicity, but not stativity, and that the role of sentential aspect does not come into play until later in the derivation after first MERGE.

Having outlined some of the complexities of aspect and the difficulty in defining statives, it should be noted that there are some characteristics of statives that are generally agreed upon. First, states are considered to be homogenous, having a subinterval property which indicates that if the predicate is true for a given time interval, then at any sub part of that interval it will be true (Rothmayr 2009:3). Thus, statives are characterized by continuity rather than any change in state and consequently also have the property of durativity. Additionally, statives are said to lack internal structure (Rothstein forthcoming). But how can this be operationalized and how can possible aspectual restrictions on BCV constructions be assessed?

A straightforward resolution to the ongoing issue of categorically defining stative verbs may not be readily available, and in addition, diagnostics are of limited use unless grounded in a theory of aspect (Levin 2009). Therefore, following Walker (2010a,b) I advocate one way of operationalizing or circumscribing stativity that makes use of independent criteria. Walker's (2010b) approach is based on Olsen's (1997) model of aspect and also makes use of Levin's verb classes

154

(1993), both of which I adopt. I assume that lexical aspect is an inherent property of the verb (Rothstein 2004:33; Olsen 1997, 1994) and, following Leblanc (2010), that at the lexical or word level, the aspectual distinction in the sense of Aktionsart is one of stativity, i.e., lexical verbs are either stative or non-stative. This binary opposition corresponds to the demarcation hypothesized to govern the verb classes that can enter into a construction with BCVs. Recall that as the grammaticalization of a lexical item progresses there is a general tendency for it to co-occur with statives last. Thus, for the current study, the binary opposition is of most interest.

4.5.2 Olsen's model of aspect

According to Olsen's model (1997, 1994), aspect is comprised of three privative features: telicity, dynamicity, and durativity. Each verb has an underlying representation that consists of some combination of these features being marked (or not marked) on the verb. In this way, the features are part of the inherent property of the verb. The [+ feature] on the verb is considered to be consistent and "uncancellable"; on the other hand, when the verb is not marked for one of the features, lexical constituents and pragmatic context may lead to the verb being interpreted as telic, durative, or dynamic even if it is not marked [+ feature] for any of those features (Olsen 1994:39). Verbs are not marked for [-feature] in this model. Thus, a stative verb would be marked [+durative] but it could have a dynamic reading as a result of arguments and adjuncts that shift the aspectual meaning of the lexical verb. In contrast, verbs marked [+dynamic] cannot have a stative reading because the verb is marked for dynamicity; this feature is "uncancellable" and cannot be neutralized. It is only when the verb is not marked for a feature that there can be a shift in the reading of the verb. In this way we can account for stative verbs with activity, accomplishment and stage-level predicate readings (Ibid: 196).

Presented below is Olsen's summary (1997, 1994) of the privative features according to

the traditional Vendlerian fourway distinction of Aktionsart as well as two additional categories discussed by Olsen (semelfactives and stage-level predicates).

Table 21: Lexical aspect of verbs according to Olsen's (1994) privative features

	Telic	dynamic	Durative
States			+
Activities		+	+
Accomplishments	+	+	+
Achievements	+	+	
Semelfactives		+	
Stage-level predicates	+		+

The main opposition between these categories is one of dynamicity, where all the categories are marked [+dynamic] except for states and stage-level predicates. States are only marked for one privative feature, [+durative], which is shared with activities, accomplishments, and stage-level predicates. Thus, states are verbs that are marked [+durative] and which are not accomplishments or activities[164], i.e., they are unmarked for [+dynamic]. This criteria is used for isolating those verbs which may be stative and which are subsequently checked against Levin's verb (1993) classes of *perception*, *psychological states*[165], *existence*, and *keep* verbs. This is not meant to be absolute; some verbs that can be argued to be stative may be left out, for example, but this is one way of delimiting statives as a first step to assessing lexical aspect of BCVs in a quantitative fashion rather than on a case-by-case basis. The verbs belonging to each of the classes chosen from Levin (1993) are considered to be homogenous in their behavior; therefore, we can assume that they share some inherent properties, including aktionsart. The four classes identified as stative here

---

[164] For the purposes of the current study I leave aside any distinction between stage-level predicates and states for now.
[165] Not including *amuse* verbs since these tend to indicate a change of state, even if it is an internal change (Levin 1993:191).

generally refer to internal states and have a property of durativity. According to the theoretical approach adopted, and the ensuing criteria, the question of aspectual restriction on BCVs can be re-framed as follows:

> Can verbs that are marked [+durative] (and not marked [+dynamic]), and that belong to the four verb classes of: (i) *perception*, (ii) *psychological states*, (iii) *existence*, and (iv) *keep*, be found as complements of *hacer*?

In the section that follows, data is coded for aspectual class according to how this has been operationalized above and, subsequently, the distribution of statives in the data is assessed.

4.5.3 Coding protocol

Each English-origin verb in the BCV was coded for dynamicity and durativity. Each token could be coded as either marked [+ dynamic] or unmarked for this privative feature. Likewise, tokens could be coded as marked [+durative] or unmarked for this feature. The verb *enjoy* in the BCV *hacer enjoy*, 'to enjoy' in (96), for example, is unmarked for dynamicity but marked [+durative] for durativity and was coded as stative. It was then checked in the above-named verb classes and appears in Levin's (1993:191) class of "*admire* verbs" which includes other verbs such as *like*, *love*, *treasure* and *tolerate*.

96. Todo   el   tiempo lo   hacemos   **enjoy** así   con   los
    all   the   time   CL   do1PL.PRES   enjoy like   with   the

    friends   que   salimos
    friends   that   go.out1PL.PRES

'We enjoy it all the time, like with friends that we go out with'

(BSEK/29W/29W/00:06:24)

Verbs such as *drive* in (97) below are marked [+ durative, + dynamic] while a verb such as *hacer top* 'to top' is marked [+dynamic] only. Thus, both verbs are considered dynamic verbs

97. iba   hacer **drive** para   acá
     go3SG.PRET   do.INF drive   to   here

    'She was going to drive here'

<div align="right">(BSEK/11WJ/11WJ_12WJ_13WJ/ 58:18.8)</div>

98. va   estar   bueno que   haga   **top**   pero, yo,   para
     go3SG.PRES   be.INF good   that do3SG.PRES top   but   I   for

     mí,   mi   punto de   vista, John a   top   the   polls
     me   my   point of   view   John will top   the   polls
     again
     again

     'It would be good for him to top {the polls}, but, me, for me, from my point of view, John. will top the polls again.'

<div align="right">(BSEK/ 45NJ/45NJ_43NJ_3/01:23:03/)</div>

All but four of the stative verbs (*experience, understand, focus,* and *deal*) were confirmed as belonging to the verb classes in Levin (1993) which were previously identified as constituting stative verbs for the current analysis. Therefore, by appealing to Levin's (1993) verb classes, the status of 80% of the data coded as stative, i.e., marked only for the privative feature [+durative], could be confirmed as such.

4.5.4 Results

The distribution of statives and [+dynamic] verbs is presented in Table 22 below.

Table 22: Distribution of stative/non-stative BCVs

	Tokens	
	N	%
Stative	19	10
Dynamic	163	90
Total	182	100

The results show that 90% of the BCVs are dynamic and that statives are relatively infrequent (10%). Thus, BCVs appear to be dynamic in their majority. This distribution suggests that the appearance of statives in these constructions may be restricted in some way. The verb *hacer* may continue to retain some of its lexical meaning of "to do/make" (Fuller Medina 2013; Jenkins 2003), which may translate into an aspectual sensitivity to the type of verb that can occur in BCVs. Such a finding is not inconsistent with a grammaticalization hypothesis. First, as has been noted, *hacer* shows semantic bleaching and an almost exclusively functional role in BCVs, typical of light verbs, in contrast to its main verb counterpart. Second, the infrequent appearance of statives in BCVs is consistent with grammaticalization trajectories, as when an item is developing a new function, it generally begins with activities and then spreads to states last (Bybee et al. 1994). This is reminiscent of the development of English *do* into a support verb which initially had a prohibition on stative verbs but lost it over time (Roberts 1993)[166]. The above pattern may reflect a synchronic picture of this type of trajectory. When the group of stative verbs is examined more closely, additional evidence emerges to support the idea that the use of stative verbs in BCV constructions may be at an incipient phase.

Recall that statives can be considered to be on a continuum of more to less stative (or more dynamic). If this is the case, then it is possible that the verbs in the current data are closer to the more dynamic end of the continuum. First, some of the verbs have both a dynamic use and a stative use. In (99) below, the verb *taste* is often considered a stative verb (consider *The tacos taste good*) and, according the criteria used, was coded as such, but in the example below, it has a dynamic interpretation denoting the act of eating tacos to assess their taste. This may reflect a case of a stative verb receiving an active reading, which is accounted for in Olsen's model. It could

---

[166] See also previous references given for the development of the "go" future in Spanish and English for further illustration of grammaticalizing items.

also be the case that there are two different verbs *to taste*. Whichever the case may be, verbs of

this type permit both a dynamic and stative use.

99. los judge hacían **taste** los tacos
    the judge do3PL.IMP taste the tacos

    'The judges tasted the tacos'

(BSEK/62N/62N/18:26.7)

Second, all the verbs coded as stative in the current data show the possibility of occurring in the

progressive. While the test of the progressive is not reliable for identifying statives, it is still useful in

telling us about the group of verbs labeled statives here. That all the verbs do not resist the

progressive, in the way that prototypical statives such as *know* do, suggests there is some quality

about these verbs that is different, possibly that they are at the more dynamic end of the stative

continuum. If we consider the verbs *trust*, *enjoy*, *understand*, and *tolerate* which are found in BCVs

in the current corpus (though not in the progressive), we can see that they can take the progressive

as exemplified below.

100. Voters are trusting him to do that[167].
101. I am enjoying the conference.
102. I am understanding all the lectures.
103. I am tolerating the pain well.

These sentences may not appear in the corpus but neither are they your typical made-up

sentences. These are utterances that are easily overheard in spontaneous speech.

It appears then that the stative English-origin verbs that occur in BCVs all have in common

some "more dynamic" characteristic (like the *wonder* verbs noted above from Levin 2013) in

addition to their stative identity. These are precisely the types of verbs we would expect to see

participate first in a process of grammaticalization where a dynamic verb is expanding in its

---

[167] (http://www.saultstar.com/2015/10/21/justin-trudeau-is-ready---column)

functional use to all verbs. Furthermore, it is the youngest speakers who produce almost all the stative verbs in BCVs and if the use of statives is at an incipient phase then we would expect to see this use among younger speakers. This may be epiphenomenal of the data given that younger speakers tend to be high language mixers and, as a result, may simply produce more BCVs. However, since older speakers produced almost no statives, this question is one that would have to be examined further.

Thus, while the data shows that statives can occur in BCVs, they do so infrequently. In addition, it appears that it is only a special class of statives which appear as complements of *hacer* in these constructions, possibly because *hacer* bears a sensitivity to the lexical aspect of its verbal complement and is only just beginning to take stative verbs as complements in the formation of BCVs. To further confirm the above analysis, diachronic data would help to establish whether or not this is a change in progress where further bleaching of *hacer* may take place. Moreover, determining the overall frequency of statives in the corpus would help to ascertain whether the infrequent appearance of stative BCVs is simply a reflection of overall infrequent use of statives or whether they are only infrequent in BCVs. In addition to overall rates, it would be important to determine the distribution of different types of stative verbs. If in monolingual Spanish, for example, there is a broad distribution of both less dynamic and more dynamic statives, but only the latter appear with BCVs, then this would suggest that the above analysis is not epiphenomenal of the corpus. I leave these questions for further research.

In summary, the following points can be made: (i) *hacer* shows evidence of bleaching in the sense that it has minimal semantic load and is highly functional (ii) there exists a possible limitation on *hacer* taking stative English-origin verbs as complements, as only a limited class of statives seem to be occur in BCVs; (iii) *hacer* appears to retain some lexical meaning and (iv) it shows no morphological reduction indicating that it may become an invariant form, therefore, it is

161

not fully grammaticalized.

4.6 Discussion

Returning now to the research questions of this chapter, two conflict sites were used to assess the status of the English-origin verbal items with a focus on those found in BCVs: verbal morphology and variable clitic placement. It was found that some English-origin verbal items showed Spanish morphology directly attached to the English stem (*trayar* 'to try'). There were also English-origin verbs with English morphology (*tenemos… signed* 'we have… signed'). The former could be considered borrowings while the latter are likely code-switches, though both categories were very infrequent and did not constitute a major pattern in the data. All other English-origin verbal items appeared with *hacer* in a BCV (the major pattern in the data), in which Spanish morphology was always expressed on *hacer*, and English verbal morphology was not detectable, nor were the features of the Kriol verbal system, though this was not quantitatively assessed. In addition, object pronouns were expressed in Spanish and occupied positions in relation to the finiteness of the BCV consistent with monolingual Spanish as spoken by the same speakers. Thus, I concluded that both the direct and indirect diagnostics used to assess BCVs in Belizean Spanish support the view that the English-origin verbs are to be considered borrowings. While some authors believe that code-switching and borrowing cannot be distinguished (e.g. Myers-Scotton 2002) and that the status of BCVs, in particular, cannot be ascertained with respect to borrowing or code-switching (Wilson 2013), or that BCVs may represent a third grammar or separate "bilingual syntax" (Romaine 1995; Muysken 2000), the preceding analysis has shown that when assessed quantitatively, and by using conflict sites as diagnostic tools within the comparative framework, their status can indeed be confirmed as borrowings.

The second research question regarding aspectual restrictions was also addressed through quantitative analysis. It was found that statives, as defined in this study, do occur in BCVs but are highly infrequent. Furthermore, those stative English-origin verbs in BCVs may constitute a more "dynamic" set of stative verbs. This would be consistent with an analysis of grammaticalization if it is the case that the synchronic distribution is reflecting a process of expansion in which *hacer* is moving from restricted use with dynamic verbs to use with all verbs in order to borrow verbs from English into Spanish. This suggests that there is no categorical restriction on statives, but rather that BCVs may be sensitive to the aspectual property of the English-origin verbal item that it takes as a complement.

This chapter also addressed some methodological questions in the study of BCVs. The first issue was how to define BCVs. Much of the literature gives reasonable indications of how the English-origin item is to be categorized but a detailed description of how BCVs can be defined is rarely presented. Lexical categories can serve more than one function; therefore, verbs can act as nouns and nouns can act as verbs. Since we still know relatively little about BCVs, especially in Spanish-English data, it is important to clearly demarcate the object of study, particularly if comparisons are to be made with corresponding categories in monolingual data. I proposed a series of discourse-based tests to determine whether the English-origin item was indeed verbal or nominal (or some other category) and by applying these tests, most of the data was accounted for and confirmed as verbal. It also became apparent that English-origin items that could be identified as nouns are infrequent in light verb constructions with *hacer.*

The use of BCVs in the Spanish spoken in Belize was found to be highly productive and is the main strategy for borrowing verbs in both Spanish-speaking communities targeted in this study. This is a finding that differs from other U.S. Spanish-English settings where BCVs are reported as a minor pattern (Toribio et al 2012; Silva-Corvalán; Pfaff 1979; though see Wilson and Dumont

2014:12), reported to be limited to Southwest Spanish (Jenkins 2003) or not reported at all for others (e.g. Poplack 1980 on Puerto Rican Spanish). This finding also contrasts greatly with the typologically similar language pair French-English where verbs were reported to be categorically borrowed synthetically into the –*er* class of verbs (Poplack et al 1988). This points to the limited predictive power of typology in determining the verbal borrowing strategy a community will adopt and to the fact that the borrowing strategy is constrained, at least in part, by community-based usage norms. The mere fact that a bilingual strategy might be possible and available to speakers does not mean that it will be used to any significant degree[168].

The second methodological issue was defining and operationalizing aspect. As a first step I have followed variationist approaches (Walker 2010a,b; Leblanc 2010) and grounded the approach in a model of aspect that can account for some of the variable behavior of statives but that still assists in narrowing down which verbs might be considered stative. This has helped to disentangle, to some extent, the broad question of aspect in order to identify and test a specific hypothesis at the level of the first building block of aspect: lexical aspect. This also provides a division of verb classes which is more strongly grounded both theoretically and empirically and, as a consequence, potentially offers a more easily replicable methodology. This approach is not meant to be absolute. Rather, it is a first step as it not only allows a protocol for classifying production data, but such a protocol can also be applied to the selection of verbs for analysis in other types of studies such as the judgment tasks for those researchers who work within the experimental paradigm. At the same time, this approach leaves room for further development to assess issues of grammaticalization and aspect shift of statives in BCVs. Expanding the verb classes that are included as stative might be a second step. Verb classes such as *container subject* and *raw material subject* classes (Levin 1993) would be worth exploring, for example, since

---

[168] Community patterns are further elaborated in chapter 6.

164

these show alternations between accomplishments and states (Olsen 1997:39). In addition, *amuse* verbs were initially excluded as these tend to show a change of state (Levin 1993:191) which is inconsistent with the generally held notion of states as homogenous. Nevertheless, it would be important to explore this class as well since some of these verbs show alternations between change of state verbs and stative verbs. While I attempted to address the question of constraints on lexical aspect specifically, it would be important to determine the pattern of different verb types with BCVs as well, i.e., whether they are used primarily with activities, accomplishments or achievements given that this synchronic distribution can also give insight into the nature of *hacer,* and whether or not it is in a process of expansion.

Furthermore, whereas the synchronic data may evidence persistence of previous function and use of a grammaticalizing item, an important step would be to compare the use of BCVs across age groups in an effort to capture the diachronic development using the apparent time construct (Bailey et al. 1991) or, alternatively, accessing diachronic data to assess earlier uses of BCVs. The behavior of monolingual Spanish light verbs, in particular with *hacer*, as well as the overall frequency of statives would also have to be established to gain a complete picture with regard to grammaticalization and the observed sensitivity of BCVs to statives.

One final note on methodology merits comment. The variationist method gives primacy to vernacular data and holds that this type of data is best obtained via the sociolinguistic interview. Numerous studies, as discussed in chapter 2 and 3, have successfully used these methods to elucidate many aspects of language change and variation. To add to this discussion, we can compare the methods used in Fuller Medina (2005a,b) and the current study. It provides a unique opportunity to contrast the two methodologies since they are both quantitative studies based on production data and analyze the same variable from the same regions in Belize with comparable speakers. In contrast to the current study, the data in Fuller Medina (2005 a,b) was obtained via a

semi-directed interview in which pictures depicting the activities denoted by the BCVs reported in the literature (e.g. *hacer bless*) were used as prompts. Speakers were asked to describe or tell a story about the picture[169]. Eleven participants of the eighteen who completed the elicitation task produced a total 34 tokens of BCVs[170]. The current study gleaned more than five times that data with a comparable number of speakers (N=12) from the same communities. The latter dataset provides more robust data and allows patterns of use to be identified. Thus, both the quality of the data and the quantity of data differ in non-trivial ways. This highlights the utility of accessing language as it is used in well-defined communities via tools such as the sociolinguistic interview, which is specifically designed to tap into vernacular speech. Furthermore, despite claims to the contrary (Wilson 2013: 132), it shows that large numbers of BCVs can, in fact, be captured in production data provided that their use is robust and that the vernacular of high language mixers in a well-defined speech community is tapped.

---

[169] It was unknown at the time how frequent BCVs were in Spanish spoken in Belize and whether they were used in both Spanish-speaking regions; this, coupled with the fact that morphosyntactic variables can sometimes be infrequent, led to this type of elicitation task being chosen as the tool for collection of production data.
[170] Type/token frequency 23/34.

CHAPTER 5 ENGLISH-ORIGIN NOUNS IN SPANISH DISCOURSE

5.1 Introduction

Nouns account for the highest number of single non-Spanish items in the data, and because they often appear at equivalence sites, where English and Spanish share the same word order, their status with regard to code-switching or borrowing tends to be ambiguous. Spanish and English share the same word order for DPs [DET + N], PPs [P+ N], and generally share SVO order, though Spanish allows post-verbal subjects. In (104), for example, the English-origin noun in *mis relationships* appears in the same post-determiner position as in the monolingual Spanish *mis amigos* 'my friends'. Consequently, examples such as *relationships* in (104), may be manifestations of a code-switches or borrowings. This chapter aims to ascertain the system membership, or status, of these English-origin nouns in order to elucidate community patterns of language mixing. At the same time, this analysis addresses the overarching question of whether or not single-word and multiword alternations are produced from two different processes or one underlying process. Three conflict sites are identified between Spanish and English: (i) gender agreement, (ii) number agreement and (iii) post-verbal subjects. These are discussed in detail below. In the examples below the English-origin noun appears in **boldface**, while Spanish nouns are underlined.

104.
y	así	trato	de	hacer	con	toda	mis
and	like.that	try1SG.PRES	to	do.INF	with	all	POSS.PL

relationships,	yo	como	digamos,	con	mis	amigos o
relationships	me	like	say1PL.PRES	with	POSS.PL	friends or

algo
something

'And I try to do it that way with all my relationships, I, like let's say, with my friends or something.'

The variationist comparative method provides some key tools for revealing the system membership of English-origin nouns in Spanish discourse. By making use of conflict sites, and by examining the quantitative patterning of nouns in all three datasets (mixed discourse, monolingual Spanish, and monolingual English), it can be shown whether the ambiguous English-origin nouns are behaving in ways consistent with Spanish or English grammar. This approach allows a detailed perspective of the underlying grammars that bilingual speakers are using, in particular, the grammar that is operative when they use English-origin nouns in Spanish discourse.

Three conflict sites between English and Spanish are used as diagnostics to assess the English-origin nouns. The first is gender agreement- Spanish employs gender agreement between nouns and their determiners and modifiers while English does not. Consider (105) below where a feminine determiner (*la*) is used with *banda* and a masculine determiner (*el)* is used with *colegio*.

105. | solo | like, | puros | **boys** | así | de | <u>la</u> | <u>banda</u> |
     |------|-------|-------|----------|-----|------|---------|----------|
     | only | like | pure | boys | like | from | DET.FEM.SG | band.FEM |

y	de-l	<u>colegio</u>
and	from-DET.MASC.SG	college.MASC

'Not so much, just, like, only boys, like from the band and from school.'

A second conflict site is number agreement between the determiner (or other modifier) and noun. In Spanish, if the noun is plural, then the determiner must be plural as well, while in English agreement between the determiner and the noun occurs in limited contexts. Note, for example, the use of the plural form *mis* 'my,' in (104), which agrees with *relationships*. This is consistent with Spanish grammar, and if this is found to be a pattern in the mixed data, then this would indicate

that the English-origin nouns are integrated into Spanish grammar since they are being subject to Spanish agreement rules. Finally, Spanish permits postposed subjects while English does so only in very limited contexts (Quirk et al 1985). The other conflict sites of gender and number agreement rely on agreement morphology observable on the determiners and the nouns they modify while the final conflict site of subject position makes use of typological differences between Spanish and English word order.

Each of these conflict sites is described in detail below. This is followed by a description of the compilation of the comparison datasets. Once the datasets have been described, the coding protocol, analysis, and results are presented in turn for each of the conflict sites. In keeping with the assumption that borrowing is a morphosyntactic process, optionally accompanied by phonological and social integration of the loanword, only morphosyntactic criteria are addressed here[171].

5.2 Conflict Sites

5.2.1 Conflict site 1: Gender agreement

Gender agreement represents the first conflict site used to determine if English-origin nouns in Spanish discourse are being integrated into Spanish grammar or if they retain English grammatical properties. Gender agreement is not a part of the English system, therefore, English nouns do not carry a gender feature. On the other hand, Spanish requires gender agreement between the determiner and the noun; consequently, either masculine or feminine determiners are assigned to

---

[171] As noted in Chapter 2 and Chapter 3, numerous studies have established that nonce uses of non-native words in a recipient language behave in the same ways as those loanwords that have been fully integrated into a recipient language to the point of achieving dictionary attestation. See: Poplack and Dion (2012), Poplack et al. (1988), and Poplack and Meechan (1998a).

Spanish nouns, based on the grammatical or biological gender of the referent[172]. If the noun denotes an animate entity, then the gender of the determiner will be masculine or feminine in accordance with the biological sex of the entity. Inanimate nouns receive canonical or non-canonical gender marking. Canonical gender consists of marking words ending in 'o' with masculine determiners (*el/los*, *un/unos*, *este/estos*) and words ending in 'a' with feminine determiners (*la/las*, *una/unas*, *esta/estas*). Non-canonical marking refers to gender assignment that does not depend on the ending. This includes consonants, vowels other than *–o* and *–a*, and those nouns which end in *–a* but are masculine, such as *el día* 'day,' or those that end in *–o* but are feminine, such as *la mano*, 'hand.' In addition, Spanish is also said to have the masculine as the default gender which is used when the lexical item does not bear enough information to determine gender (Eichler et al. 2012). More specifically, in Harris' (1991:43-44) analysis, there is only one privative gender marker on nouns - feminine - rather than a binary [+/-feminine] feature. As a result, there is no marking for (grammatical) masculine gender on nouns and what is termed masculine only means that the noun lacks gender marking (i.e., feminine). In other words, default gender, then, is the absence of gender on a noun. Eichler et al. (2012:236) provide the following examples as evidence of the default masculine in Spanish. In (106), taken from Harris (1991), the masculine is used to modify *para*, 'for,' while in (107) where both the feminine and the masculine nominals are modified, "masculine agreement on the adjective is triggered" leading to the masculine form of the modifier being used. Noun-verb compounds and nominalized verbs are also modified by masculine determiners, as illustrated in (108) and (109), respectively. Finally, in (110), the use of the masculine for superordinate concepts is also taken as evidence for the masculine as the unmarked form in Spanish.

---

[172] Agreement is required between all modifiers and the nouns they modify but I leave aside adjectival modification for the time being.

106. Tienes demasiados (m.) "paras" en ese párrafo.
     'You have too many "fors" in that paragraph.'

(Harris 1991:43)

107. los   lobos   y   las   gatas   blancos
     DET.MASC.PL   wolf.MASC.PL and   DET.FEM.PL cat.FEM.PL white.MASC.PL[173]
     'the white wolves and cats'

(Eichler 2012:236)

108. a. el saber
        'knowing
     b. El fumar es nocivo para la salud
        'Smoking is bad for your health'

109. el lanzallamas
     throw-flames(f)
     'flame-thrower'

(Eichler 2012:236)

110. el primo-- la prima-- los primos
     the-MASC cousin-- the-FEM-- the-PL.MASC cousins

(Eichler 2012:236)

While gender is considered an inherent feature on nouns, in Harris' analysis, nouns are marked for gender (feminine) or not marked. "I interpret unmarked or default gender literally as the absence of any information about gender in lexical entries. […] Genderless words like *para* cannot transfer or be checked for either value of a binary feature like [± feminine], […]," so 'masculine' is just a label applied to the absence of the feminine marker (Harris 1991:44, 60). In (4) above, *para* 'for', is genderless, and as a consequence, according to Harris, Spanish assigns the 'masculine' default (el/los, un/unos, etc.) to genderless nominals.

From this analysis, it can be understood that in monolingual Spanish, speakers apply feminine determiners when nouns are marked for gender (feminine), either due to the biological sex of the animate referent and/or due to grammatical gender on the noun, otherwise, when nouns are unmarked speakers apply the masculine. These options are also presumably available to bilingual speakers, and (i) they may treat English-origin nouns as genderless (unmarked for

---

[173] With the exception of this gloss, glossing is retained from original source for these examples.

feminine) and apply the default masculine as they normally would to Spanish unmarked nouns; (ii) they may treat them as marked feminine or unmarked for gender by analogy with the Spanish equivalent, thus fulfilling what is known as the "analogical criterion" (Otheguy and Lapidus 2003); and/or (iii) they may assign a masculine or feminine determiner based on biological sex of animate referents.

Previous research has found that speakers may choose more than one strategy. English-origin nouns in Spanish discourse, or mixed DPs[174], as they are often referred to in experimental and theoretical analyses, have been found to be assigned masculine or feminine determiners based on the following factors[175]: according to analogical gender, i.e., the Spanish equivalent of the English item (Liceras et al. 2013, 2008, 2006; Morin 2006; Moyer 1993; Poplack et al. 1982; Jake and Myers-Scotton 2002); according to physiological gender or biological sex of the referent, (Morin 2006; Smead 2000; Poplack et al. 1982) or according to the masculine default, i.e., a preference for all English-origin nouns to be assigned masculine determiners (Aaron 2014; Liceras et al. 2013; 2008; Valenzuela et al. 2009; Franceshcina 2001[176]; Roca 1989; White et al. 2004).

It should be noted that the application of the analogical criterion has been largely reported for experimental data and for native speakers of Spanish[177], while production data tends to show evidence of a default masculine (Aaron 2014; Otheguy and Lapidus 2003) or an interplay of

---

[174] Because these English-origin nouns often carry a Spanish determiner, they have been referred to as mixed DPs, code-mixed DPs, code-switches, or simple lone English-origin nouns.

[175] Some researchers argue that the phonological shape or final phoneme of the English word is a factor in assignment of the gender of the determiner (e.g. Dubord 2004; Smead 2000; Poplack et al. 1982); however, this line of inquiry is not pursued further as it is arguable if such a rule is robust in monolingual Spanish (RAE 2009:82), and, in addition, the current data did not lend itself to this exploration since the variability needed to test this factor was simply not present in the data. Additionally, while the above factors are presented as if independent from each other, some researchers have found that multiple factors impact gender assignment and no one factor is responsible (see Jake and Myers-Scotton 2002; Poplack et al. 1982 among others).

[176] For near-native and non-native speakers.

[177] As compared to what are termed heritage and L2 speakers. Liceras et al (2013) report that there some evidence of application of the analogical criterion in Moyer's (1993) Gibraltar data.

competing of factors as described above (Smead 2000; Poplack et al. 1982).[178]. There appears to be some correlation between application of the analogical criterion and fluency in Spanish though it is not clear-cut (Liceras et al. 2008). Bilingual speakers may treat borrowings as lacking 'gender information' and apply the default masculine (Harris 1991) regardless of other factors[179] or they apply the analogical criterion.

The most visible evidence of integration of English-origin nouns into the Spanish system would be the application of the analogical criterion, and we can think of this as strong evidence since it would reflect the same distinctions made in monolingual Spanish. Determiner marking according to biological sex of the referent could also be taken as strong evidence in support of a borrowing analysis of English-origin nouns in Spanish discourse. If the English-origin nouns in Spanish discourse are marked in their majority or, categorically, with masculine determiners, then this can be interpreted as the application of the default masculine consistent with monolingual Spanish when a noun is not marked feminine or, in Eichler et al.'s (2012) terms, when there is insufficient evidence to determine gender.

If, on the other hand, these nouns were single-word code-switches, then we could expect an absence of the application of the concord rule even where the contexts for it to be applied are available in the data. On its own, the use of the masculine form for all nouns would not provide enough evidence to justify treating these nouns as code-switches, i.e., behaving according to an English (genderless) system, since the default masculine is part of the Spanish system. Other morphological markers on the determiner and noun can be examined for application of Spanish

---

[178] For a recent summary of experimental data with respect to this, see Liceras et al. (2015:29-30).

[179] An alternative view suggests that the preference for the masculine can be analyzed as a contact effect where bilinguals "adapt" by simply not applying a masculine/feminine distinction to English-origin nouns in Spanish discourse (Otheguy and Lapidus 2003). As the analysis and results will show, the current data lends support to the masculine default given that some examples of its use could be found in the unmixed Spanish dataset.

agreement rules which would not be evidenced in English. This is further explored via number agreement, as described in the following section.

5.2.2 Conflict site 2: Number agreement

Spanish requires number agreement. Determiners must agree in number with the nouns they modify (RAE 2010) as in (111) below where *aldea* is singular and as such occurs with a singular determiner, *la*, while *caballos* is plural and takes a plural determiner, *los*. Examples are underlined.

111. que   le   hacía   trensa a   <u>los</u>   <u>caballos</u>…,
      that   CL   do3SG.IMP   braid   to   DET.MASC.PL   horses.MASC

      en   <u>la</u>   <u>aldea</u>
      in   DET.FEM.SG   village.FEM

      'that he used to braid the horses' manes…, in the village …'
                                                  (BSEK/60NJ/59NJ_60NJ/00:15:38.046)

English, on the other hand, requires agreement only in limited contexts, i.e., indefinite articles and demonstratives. Both singular and plural nouns may have a singular determiner and, in fact, English does not have a plural form for definite determiners or possessives[180]. In (112) below *polls* is plural but occurs with a determiner that is invariant with regard to number.

112. but I think Greg will top the polls (BSEK/45NJ/45NJ_43NJ_3/01:22:08)

Where agreement between plural nouns and plural determiners is present with English-origin nouns in Spanish discourse, this is taken as evidence of nouns being integrated into Spanish, i.e., borrowings. Where a determiner that is singular in number (e.g. *el*) is used irrespective of number on the noun, this is taken as evidence that the nouns are single-word code-switches, i.e., behaving according to English grammar where definites and possessives are

---

[180] When the possessor is plural, English has the forms *our* and *their*, however, these forms do not agree with the object being possessed whereas in Spanish agreement with the noun must take place regardless of the number of the possessor. Cf. Our book. Our books. *Nuestro libro. Nuestro**s** libro**s**.*

invariant with respect to number. Accordingly, the failure of the Spanish number agreement rule will be taken as evidence of a switch to English.

5.2.3 Conflict site 3: Subject position

One final conflict site between Spanish and English is subject-verb word order. As noted above, Spanish freely permits post-verbal subjects, while subject-verb order is predominantly SVO in English. In fact, Torres Cacoullos and Aaron (2003:314) found post-verbal subjects to occur in their Spanish data at a rate of 46%, while no examples of post-verbal subjects were found in their English dataset. If the English-origin nouns Spanish discourse are subjects and are found to be post-verbal, then this will constitute additional evidence of integration into Spanish, pointing to their status as borrowings rather than code-switches. This last diagnostic differs from the first two in that post-verbal subjects are not obligatory for all verbs in Spanish.

In summary, if these single items are code-switched, we would expect to see evidence of English grammar, such as the absence of gender and number agreement between the Spanish determiner and the noun. Evidence in support of a borrowing analysis would include the application of a gender-agreement rule, number agreement between the Spanish determiner and the English noun, and the presence of post-posed subjects.

5.3 Comparison Datasets

In keeping with the comparative variationist framework as applied to language mixing, three comparison datasets were created with the purpose of assessing the English-origin nouns in Spanish discourse. These include monolingual Spanish, monolingual English, and mixed speech, i.e., single English-origin nouns in Spanish discourse. By establishing the patterns of use for gender and number agreement as well as variability of subject position for each dataset, the patterns in mixed speech can then be compared to those in Spanish and English to determine if

they mirror the patterns of one or the other. If the mixed speech shows patterns parallel to English, then an analysis of these nouns as code-switches is supported since the nouns will be evidencing English grammar rather than Spanish. If the mixed speech patterns with Spanish, however, then the English-origin nouns can be analyzed as borrowings since parallel patterns show integration into Spanish. The comparison datasets are compiled from the current corpus so that mixed data is assessed against the communities' varieties of Spanish and English rather than against a set of prescriptive rules or standard varieties. This compilation is described in the ensuing sections.

Mixed Speech

The dataset of English-origin tokens in Spanish discourse, whose status as code-switches or borrowings is to be determined, was created by extracting every English noun occurring in otherwise Spanish discourse, such as in (113) below.

113. El   **microwave**   cayó,   like,   solo   cayó   así,
     DET.MASC.SG   microwave   fell   like   alone   fell   like that
     'The microwave fell, like, it just fell like that.'

                                           (BSEK/29W/29W/11:00.1)

Spanish monolingual

To create the monolingual Spanish comparison dataset, the first eligible 60 nouns for each participant were extracted from the first twenty minutes in the body of the interview.[181] These tokens had to occur in monolingual segments of Spanish, as in (114) below, where all underlined items were eligible for extraction and analysis. However, if tokens occurred in the vicinity of mixed speech, following Torres Cacoullos and Aaron (2003), these tokens were not included. Thus

---

[181] The body of the interview is considered here to be the segment between the first 15 minutes of the interview and the language component at the end of the interview. Typically, this is about 45 minutes long with the exception of interviews that were shorter or longer than an hour. The first 60 nouns were extracted regardless of the amount of speech time; in this way, the same number of tokens were extracted per speaker, with the exception of one interview which was less than 30 minutes, in which case all the eligible Spanish nouns were extracted.

tokens such as *amigos* in (115*)* below were not included since they immediately follow a mixed segment. Consequently, 665 tokens were retained for the analysis of Spanish nouns in Spanish discourse, hereafter referred to as the Spanish dataset.

114. me contaban sí, <u>historias</u> pero nunca lo vi, de, <u>del duende</u>, sí, bastante, que le hacía trenza a <u>los caballos</u>, pero hay <u>unos historias</u> más miedosa que dan, ehm, en <u>la aldea</u> donde vivo ahorita yo. Ehm, había <u>un viejito</u> y atrás de <u>el monte</u> el fue y vio el Loco sin Cabeza {I: ok}. Con <u>calentura</u> llegó.

    'They used to tell me stories, yes, stories, but I never saw it, about, about the Duende, yes, a lot, that he used to braid the horses' manes…, in the village where I live now, I ehm, there was an old man and he went behind the bushes and saw the Crazy Headless man)
    (BSEK/ 60NJ/59NJ_60NJ/00:15:38.046)

115. y así trato de hacer con toda mis relationships, yo como digamos, con <u>mis amigos</u> o algo
    'And I try to do it that way with all my relationships, I, like let's say, with my friends or something.'
    (BSEK/08W/08W_03/00:17:27)

English Monolingual

Despite the copious non-Spanish multiword fragments in the data, the monolingual English dataset is much smaller than the Spanish and the mixed speech datasets. This is due to the fact that 50% of the multiword fragments consist of two word fragments made up largely of a noun with a modifier, as in the examples below. In addition, these often appear modified by a Spanish determiner and adjective, as in (116), or are determinerless, as in (117), and, as a result, they are ambiguous in the same way as the single English-origin nouns in Spanish discourse are. Consequently, these nouns cannot be included in the English dataset and are analyzed separately in Chapter 6 where multiword fragments are discussed.

116. me quiero hacer certify como un **make-up artist**
    'I want to get certified as a make-up artist.'
    (BSEK/05N/05N14/00:25:46)

117. yo estaba detrás de la política haciendo campaign como **campaign manager**

'I was in politics in the background campaigning as a campaign manager'

(BSEK/45NJ/45NJ_43NJ_3/15:19.5)

Tokens were only extracted from multiword fragments which contained sufficient information to indicate internal consistency with English, i.e., unambiguous code-switches (Sankoff et al. 1990). In this way, only nouns consistent with the English system were included. Tokens that were included in the monolingual dataset are in boldface in (118). A total of 127 tokens make up the monolingual English dataset.

118.   but I'm a human being and we can do it under the **table**.

(BSEK/45NJ/45NJ_43NJ_3/01:34:08/)

These three datasets are analyzed separately for gender and number agreement between the determiner and the noun they modify, as well as subject position. But before presenting those results, the diffusion of these English-origin nouns in Spanish discourse is presented.

5.4 Analysis and Results

5.4.1 Diffusion of English-origin nouns in Spanish discourse

Presented below is the diffusion of the lexical types that occur in the mixed data, i.e., single English-origin nouns in Spanish discourse. In the sections that follow, the coding protocol adopted for each conflict site is presented, followed by the results of analyses carried out for each of these respective areas.

The type/token frequency of the nouns in English discourse was 265/506. Following Poplack et al. (1988), in order to determine the diffusion of lexical types in the corpus, lexical types

were coded for idiosyncratic, recurrent, and widespread use[182]. The English-origin nouns used only once in the corpus are normally categorized under nonce use (Torres Cacoullos and Aaron 2003); Poplack et al. 1988)[183]. Nouns were considered to be idiosyncratic if they were used more than once by the same speaker. The descriptor "recurrent type" generally refers to use of lexical type by multiple speakers. To take into consideration the two varieties of Spanish spoken in Belize, this category was subdivided into local and supralocal recurrent use. Local recurrent use refers to the noun being used by two or more speakers in the same region (*either* north or west). Supralocal use, on the other hand, entails use of the lexical type in both northern *and* western Belize, with at least one occurrence in each region though many were used by at least two speakers in each region. The difference between supralocal and local recurrence is whether recurrence is attested only in one of two regions or if the lexical type is recurrent in both regions. This distribution is summarized in Table 23 below.

Table 23: Diffusion of English-origin nouns in Spanish discourse

	Lexical types		Tokens	
	N	%	N	%
Nonce (used only once)	188	70	188	37
Idiosyncratic (multiple use by one speaker)	26	10	123	24
Recurrent-local (≥ 2 speakers in one region)	19	7	60	12
Recurrent-supralocal (≥ 1 speaker in each region)	32	12	135	27
	265		506	

---

[182] For other detailed treatments of the diffusion of English words in Spanish discourse see also Poplack and Dion (2012), Clegg (2010), and Torres Cacoullos and Aaron (2003). These studies included "attested use" as confirmed by local and international dictionary entries, as well as compilations of Anglicisms in the case of Torres Cacoullos and Aaron (2003). These categories are not presented here since an underlying assumption in the current study is that the process of borrowing is a morphosyntactic one, only optionally accompanied by social diffusion and integration (Poplack et al. 1988). In addition, there are currently no local dictionaries or other such resources for Belizean varieties of Spanish.

[183] The categories used are defined relative to the corpus. Poplack et al. (1988: fn14) define "widespread" as the use of one lexical type by more than 10 speakers, as they were analyzing a large corpus. Torres Cacoullos and Aaron (2003: 299) do not distinguish recurrent from widespread, and define the latter as the use of a lexical type by two or more speakers.

Nonce items account for 70% of lexical types and 37% of the tokens, indicating that the use of single English-origin words is a productive process in the Spanish spoken in Belize. Idiosyncratic uses account for 10% of lexical types and reflect the repetition of the same lexical item, likely a function of the speaker narrating details of one particular topic. Recurrent types, which are repeated by two or more speakers, are also relatively small categories with local recurrent use at 7% (12% of total data) and recurrent supralocal use at 12% (27% of the total data). While it is nowhere near the category of nonce uses, it is the second largest category of lexical types, indicating diffusion of lexical types in both varieties of Spanish in Belize. A large proportion of these words recur (30% for both local and supralocal) or are repeated through idiosyncratic use, which is a feature that is not shared with code-switches since code-switches tend not to recur. The distribution in Table 23 shows that some of the English-origin nouns might enjoy some level of integration socially as the supralocal recurrent use indicates that the same lexical types are used in both regions. A larger sample would be needed to confirm this hypothesis.

5.4.2 Gender

5.4.2.1 Coding protocol

Masculine d*eterminer*

119.
mi	prima		era	el	ex-,	pues	era	el	**mayor**
POSS	cousin		be	the	ex-	well	was	DET.MASC.SG	mayor
en	ese	tiempo,		ahora	es	ex-mayor		no?	
in	that	time		now	is	ex-mayor		no	

'My cousin was the ex-, well she was the mayor at that time, now she's ex-mayor, no?'

(BSEK/45NJ/45NJ_43NJ_3/01:15:00)

120.
aquí	los	tratan	más	suaves	siento	los		**teachers**,	te
here	CL	treat	more	gently	feel	DET.MASC.PL		teachers	CL
lo	van	a	decir	en	inglés		y	en	español
CL	go	to	say	in	English		and	in	Spanish

'Here the teachers treat them more gently, I feel, they will say it in English and in Spanish.'

(BSEK/11WJ/11WJ_12WJ_13WJ_2/00:02:42/)

Feminine determiner

121. La muchacha de allá de, María Josefina se llamaba la **accountant** del, la Fundación
      'The girl from there, from, Maria Josefina was the name of the accountant from the Foundation.'

(BSEK/45NJ/45NJ_43NJ_3/01:06:41)

Analogical Gender

Included in the analysis were all English-origin nouns marked with an overt determiner, that had a masculine and feminine form as indicated above and for which the gender of the Spanish equivalent was available. The Spanish definite and indefinite articles have masculine and feminine forms whereas nouns marked by the possessive *mi* 'my,' for example, were not included since *mi* is invariant regardless of the gender of the noun when it precedes the noun (*mi amigo/amiga*). Each noun was then coded masculine or feminine according to the gender of the Spanish equivalent so that 'advantage,' in example (122) below, was coded as feminine while the determiner was coded as masculine. Nouns such as 'teachers' in (126) below were not included since this noun has both a feminine and a masculine form depending on the referent. In keeping with the prescriptivist tradition, its plural form is used to refer to both male and female teachers. In this case, it appears that the speaker is referring to teachers in general and not to a group of male teachers. These data were then cross-tabulated to determine if speakers were applying the analogical criterion in assigning masculine or feminine determiners.

Feminine Spanish equivalents

122. yo   creo   que   ese   es   un   **advantage**
      I   believe that that   be   DET.MASC.SG advantage

      que   tenemos

that   have

'I think that that is an advantage that we have.'
Spanish equivalent: la ventaja (FEM)

Masculine Spanish equivalents

123. | Los | muertos | entran | en | la | casita | y | están | allí |
|-----|---------|--------|----|----|--------|---|-------|------|
| DET | dead.people | enter | in | DET | little.house | and | be | there |

comiendo	ice-cream {laughter}	por	eso	no	le
eat	ice-cream	why	that	no	CL

compraban	el	**ice-cream**
buy	DET.MASC.SG	ice-cream

'The dead enter the little house and are there eating ice-cream that's why people never used to buy ice-cream.'
Spanish equivalent: el helado (MASC)

Biological Sex

Each animate noun was coded for biological sex if it was known from discourse. The noun 'accountant' repeated here in (124) was coded as feminine since it refers to a woman, while 'producer,' in (126) was coded as masculine since it refers to Daniel. Items such as 'teachers' in (126) were not included as noted above.

124. | La | muchacha | de | allá | de, | María | Josefina | se | llamaba |
|-----|----------|----|------|-----|-------|----------|----|---------|
| DET | woman | from | there | of | María | Josefina | CL | name |

la	**accountant**	del,	la	Fundación
DET.FEM.SG	accountant	from-DET	DET	Fundación

'The young woman from there, Maria Josefina was the name of the accountant of the foundation.'

125.	Daniel	que	es	el	**producer**,	el	music	producer
	Daniel	that	be	DET.MASC.SG	producer	DET	music	producer

'Daniel who is the music producer, the music producer.'

<div align="right">(BSEK/08W/08W_03/00:12:16/)</div>

126.	aquí	los	tratan	más	suaves	siento	los	**teachers**,	te
	here	DET	treat	more	gently	feel	DET.MASC.PL	teachers	CL

	lo	van	a	decir	en	inglés	y	en	español
	CL	go	to	say	in	English	and	in	Spanish

'Here the teachers treat them more gently, I feel, they will say it in English and in Spanish.'

<div align="right">(BSEK/11WJ/11WJ_12WJ_13WJ_2/00:02:42/)</div>

5.4.2.2 Results

Table 24 presents the results of the cross-tabulation of English-origin nouns (according to the gender of the equivalent Spanish noun) and determiners assigned to the nouns. What is immediately apparent is that the Spanish equivalents are both masculine and feminine yet masculine determiners predominate regardless of the gender of the Spanish equivalents. There were five nouns whose equivalent in Spanish is feminine, and these were marked with a feminine determiner. These tokens were produced by two speakers talking about bicycles (*bicicleta)* and while it points to the application of the analogical criterion, this is the only token where these speakers show gender agreement by analogy with the Spanish equivalent[184].

Table 24: Distribution of masculine and feminine determiners assigned to English-origin nouns according to analogical gender

	Spanish determiner		
Gender of the Spanish equivalent	Feminine N %	Masculine N %	Total

---

[184] Less than 3% of all tokens appeared with a Spanish determiner (*el*, *la*, *un*, *una*) were marked with a feminine determiner, therefore, following Guy (1988), no multivariate statistical analysis was carried out as rates were less than 5% of the data. Since there is always the possibility that most of the English-origin nouns were animate and had masculine referents or that the Spanish equivalents were prescriptively masculine and that gender was being assigned according to the Spanish equivalents, these individual factors were still assessed (see tables 24 and 25).

		5	4	119	96	124
Feminine		5	4	119	96	124
Masculine				117	100	117
Total		5		236		241

The distribution of masculine and feminine determiners according to the biological sex of animate referents is summarized in Table 25 below. There were very few female referents, therefore not enough variability to draw conclusions regarding biological sex as a determining factor in gender assignment. Two out of three tokens with female referents were marked with the masculine determiner, pointing to the use of the default masculine. In fact, the two nouns that were coded as feminine but marked with the masculine determiner refer to the superordinate concept which would be marked in Spanish by the masculine determiner. In monolingual Spanish it is perfectly acceptable to say, for example, "su primer hijo fue una niña." The examples below show similar usage. The context for example (127) is the speaker is showing pictures of her work as a makeup artist where she made up the same person as a man (*un male*) and then as a woman (*un female*). In (128), the speaker is referring to midwives as a group rather than to any specific group of midwives. She uses the masculine determiner despite real-world information which suggests that midwives are usually women.

127. oh   ese   es   un,   no   me   acuerdo   que   estaba tratando
      oh   that   be   DET   no   CL   remember   that   be   try

      de   hacer, como   un   **male**, un   **female**
      to   do   like   DET.MASC.SG male   DET.FEM.SG   female

      'Oh that is a, I don't remember what I was trying to do, like a male, a female.'
                                                                (BSEK/05N14/05N14/00:32:33)

128. llaman   a   los   **midwives**   en   las   casas
      call   to   DET.MASC.PL midwives   in   the   houses

      'They call the midwives to the homes.'
                                                                (BSEK/05N14/05N14/00:02:26)

It should be noted that male referents are never marked with feminine determiners; this is consistent with the use of masculine as default according to a feature checking analysis which predicts that the derivation would crash if a masculine determiner were used with a noun marked [+feminine] (Liceras et al. 2008; White et al. 2004). Thus, it is not only the preference for the masculine determiner that points to the use of masculine as default but, also, the fact that feminine determiners are not used for male referents or with English-origin nouns (except for *bike*).[185]

Table 25: Distribution of masculine and feminine determiners according to biological sex of the referent

Biological sex of the referent	Determiner			
	Feminine N %		Masculine N %	Total
Female	1 33		2 67	3
Male			9 100	9
Total	1		11	12

The main hypothesis with regard to these single English-origin items is that if speakers are integrating the English-origin nouns into Spanish, then we would expect to see the same pattern of gender assignment as with monolingual Spanish nouns in their respective varieties. However, in contrast to the pattern found in the mixed data, there is no overall preference for the masculine determiner in the monolingual Spanish data as seen in Table 26 below. Where a noun is feminine, the feminine determiner is applied 94% of the time, and when the noun is masculine, the masculine determiner is used 95% of the time. These speakers are therefore appropriately applying native Spanish gender agreement rules to Spanish nouns that carry a gender feature, i.e., there is

---

[185] *a la nené la puso en una canastita, creo que en frente, que tenía la **bike.*** 'and she put the baby in a little basket, that the bike had, I think in the front part.' **(**BSEK**/**11WJ/11WJ_12WJ_13WJ/00:55:12)

enough information on the noun to determine if a feminine determiner should or should not be used (assuming Harris' (1991) analysis of nouns being unmarked for gender or marked feminine). These speakers show a preference for the masculine determiner only with English-origin nouns.

Table 26: Distribution of masculine and feminine determiners according to prescriptive gender on Spanish nouns in Spanish discourse

	Determiner		
Prescriptive gender of the Spanish nouns	F	M	Grand Total
F	179 **94**	11 5	190
M	9 6	163 **95**	172
Grand Total	188	174	362

This preference for the masculine determiner is most clearly seen in Table 24 since there is near equal distribution of nouns that have feminine Spanish equivalents and masculine Spanish equivalents. Thus, the data allows for the choice of applying the analogical criterion; however, speakers simply do not apply it. It was hypothesized that if English-origin nouns were integrated into Spanish via determiner marking that gender would be assigned as per Spanish rules by application of either the analogical criterion or they could apply the default masculine (Harris 1991) in the same way that monolingual speakers do when a noun is unmarked for gender, a finding which would be consistent with a preference for the masculine determiner reported for other Spanish-English bilinguals. (in the U.S. Aaron 2014; Otheguy and Lapidus 2003; see also Jake and Scotton 2002). I would suggest that speakers are still applying a gender agreement rule, though not one based on the analogical criterion; English-origin nouns are unmarked for gender, but this provides no conflict for Spanish since they can still undergo default gender agreement just like any other Spanish nominal which does not carry a gender feature. Whether these nouns are considered to not bear enough information to determiner gender (Eichler et al. 2012), or whether

they are considered to be unmarked for feminine like any native Spanish noun not marked feminine (Harris 1991), when confronted with English-origin nouns, speakers apply this default masculine rule in the same way that it is applied to Spanish nominalized forms that have no gender feature.

Of the different types of evidence discussed previously in support of the masculine default, the current data provides examples of the masculine default when referring to superordinate concepts in the Spanish dataset. By way of further illustration, consider (129), below: when the speaker refers to teachers in general, he uses the masculine plural form in monolingual Spanish, as does the speaker in (130) when the English equivalent of *maestros* 'teachers,' is used[186]. Additionally, when the same speaker refers to a specific doctor who happens to be a woman, they use the feminine form as seen in (131), but when they speak of the doctors as a group they use the masculine form (132). Thus these same speakers show evidence of using the masculine default in monolingual Spanish.

129. | A veces | los | **maestros,** | ellos | son | aburrido |
| sometimes | DET.MASC.PL | teachers | they | are | boring |

'Sometimes the teachers, they are boring.'

(BSEK/35W/35W/00:14:16/)

130. | aquí | los | tratan | más | suaves | siento | los | **teachers**, |
| here | CL | treat | more | gently | feel | DET.MASC.PL | teachers |

'Here the teachers treat them more gently, I feel.'

(BSEK/11WJ/11WJ_12WJ_13WJ_2/00:02:42/)

131. | Así | le | dijo | la | **doctora**, | así | fría |
| like | CL | told | DET.FEM.SG | doctor.FEM | like | cold |

'Just like that the doctor told her, cold like that.'

(BSEK/45NJ/45NJ_43NJ_3/00:56:10)

---

[186] Real-world information indicates that most teachers in this community are women, yet when referring to teachers as a group, the masculine form is used.

132. Pero   según   **los**   **doctores**,   el   chamaco
     but   according.to DET.MASC.PL   doctors.MASC.PL   DET   kid

     está   bien
     is   fine

'But according to the doctors, the kid is fine'

(BSEK/45NJ/45NJ_43NJ_3/01:00:55)

The application of the masculine determiner to English-origin nouns in Spanish discourse can be taken as evidence that these nouns are integrated into Spanish. First, we see the application of a Spanish rule: apply the default masculine to nouns which do not carry any gender information to indicate feminine gender. Second, gender assignment has been shown to be consistently uniform in the community from the first use of a non-native item (Poplack and Dion 2012; Poplack et al. 1982). In the current data, the near categorical use of the masculine suggests that there is consistency in the assignment of the gender of the determiner regardless of the social diffusion of the lexical item. Recall that that 63% of all tokens were repeated in the corpus either through idiosyncratic use or recurrent and widespread use while, and 37% of the nouns are nonce uses.

5.4.3 Number agreement (DET + N)

Spanish requires determiners, demonstratives, and possessives to agree in number with the nouns they modify, while English only requires agreement between the possessed object and demonstratives and indefinite articles (Stanley Whitley 2002)[187]. This is summarized in Table 27 below.

Given that English and Spanish differ in the requirement of number agreement between the determiner and noun with regard to definite determiners and possessive pronouns, if the

---

[187] In Spanish modifiers must also agree with the nouns they modify, I deal only with determiners here.

English-origin nouns that are modified by Spanish definite articles and possessive pronouns show

number agreement, then a rule consistent with Spanish grammar is being applied, suggesting that

they are integrated into the Spanish system and are borrowings rather than code-switches.

Table 27: Number agreement in Spanish and English

	SPANISH	ENGLISH
Definite article singular plural	el libro los libros	the book/ books
Possessive singular plural	mi, su,tu,nuestro, vuestro libro mis, sus, tus, nuestros, vuestros libros	my/your/our/his/her/their book my/your/our/his/her/their books
Indefinite singular plural	un libro unos libros	a book some books
Demonstrative singular plural	este libro estos libros	this book these books

5.4.3.1 Coding protocol

All English count nouns[188] that were modified by a Spanish definite article or possessive pronoun

were extracted and coded for plural or singular number. Each determiner was also coded for plural

or singular form. In (133) below, both the determiner and the noun were coded plural.

133.	más more	mis my	cosa thing	mía mine	es is	como how	cuidar care.INF	los DET.MASC.SG	**babies,** babies
	porque because	los the	babies babies	son are	cute cute	y and	todo everything		

---

[188] Nouns such as 'news' were not included as this noun does not occur in the singular form; on the other hand, nouns like make-up do not have a plural form and these were also excluded.

'My thing is more like to take care of babies, because babies are cute and everything.'

(BSEK/62N/62N/00:10:08/)

5.4.3.2 Results

As can be seen in Table 28 below, speakers apply the Spanish agreement rule to English-origin nouns as they use singular determiners with singular nouns and plural determiners with plural nouns, categorically so in the first case, and 94% of the time in the latter case. In the monolingual Spanish data, singular agreement is also categorical, and plural agreement is near categorical at 99%[189].

Table 28: Cross tabulation of determiners and count nouns in mixed data

	PLURAL DETERMINER		SINGULAR DETERMINER		Total
	N	%	N	%	
PLURAL NOUN	58	100	0		58
SINGULAR NOUN	4	2	179	98	183
Total	62		179		241

The monolingual English data had only singular determiners, and possessive pronouns were plural only in reference to the 'possessor' but, as expected, were not marked for the objects being possessed. Table 29 below shows that where the objects were plural they occurred with a singular determiner or possessive pronoun, consistent with English grammar.

Table 29: Distribution of English plural and singular DET + N

	SINGULAR DETERMINER		PLURAL DETERMINER	
	N	%	N	%
SINGULAR NOUN	31	100	N/A	
PLURAL NOUN	12	100	N/A	
TOTAL	43			

---

[189] There was one exception with one token where the noun was transcribed as singular: no se no se cómo son sus pensamiento {laughing} 'I don't know what their thinking is'. (BSEK/60 N/59N_60Nr/00:25:52/).

Thus, while in Spanish singular determiners almost never co-occur with plural objects, they do in English. The English-origin nouns that occur in isolation in Spanish discourse pattern closely with Spanish as they do not show any instances of singular determiners with plural nouns and can be taken to be borrowings and not single-word code-switches.

5.4.4 Post-verbal subjects

One final conflict site that provides evidence of integration into Spanish of the English-origin nouns is subject position. It is well known that Spanish allows both SVO and VSO word order (RAE 2010:297, 637-639) while English is more rigid in its word order[190]. Furthermore, single English-origin nouns in New Mexican Spanish discourse have been shown to occur post-verbally in concordance with Spanish syntax (Torres Cacoullos and Aaron 2003:314). If post-verbal subjects are found in the mixed data, then this would be evidence of the Spanish system being operative rather than an English (or even a Kriol) one. Table 30 below shows rates of postposed subjects in the three datasets.

Table 30: Post-verbal subjects

	Spanish		Mixed		English	
Subject position	N	%	N	%	N	%
post-verbal	21	17	7	16	0	
pre-verbal	103	83	36	84	7	100
Total	124		43		7	

X^2 (166)=0.01; p = .92

English shows no post-verbal subjects, consistent with the grammar of that language, while post-verbal subjects are present in the mixed data at rates virtually identical to those found in the monolingual Spanish data. No statistically significant difference is found between the latter two datasets. In addition, the post-verbal subjects in the mixed data appeared with the same verb types which are said to typically allow post-verbal subjects in monolingual Spanish (Rivas 2013). These

---

[190] See Quirk et al. (1972:770-771) on the limited contexts where post-verbal subjects can occur in English.

included *estar* (N=6), unaccusative verbs such as *llegar, caerse, venir* (N=7), "say" verbs (N= 4) and other intransitive verbs (N= 4). Post-verbal subjects are exemplified below for both Spanish and mixed data.

134. Después llegó   la   **policia**
     after   arrived DET.FEM.SG   police.FEM.SG

     'Then the police arrived.'

                                        (BSEK/12WJ/11WJ_12WJ_13WJ/00:50:30)

135. Aquí   los   tratan más   suaves, siento, los   **teachers**,   te
     here   DET   treat   more gently feel   DET.MASC.PL   teachers   CL

     lo   van   a   decir en   inglés   y   en   español
     CL   go   to   say   in   English   and   in   Spanish

     'Here the teachers treat them more gently, I feel, they will say it in English and in Spanish.'
                                        (BSEK/11WJ/11WJ_12WJ_13WJ_2/00:02:42/)

While the tokens in subject position only account for a small portion of the data, they pattern identically with Spanish, further supporting the above analysis of single English-origin nouns in otherwise Spanish discourse as borrowings rather than single-word code-switches. In addition, these findings parallel those of Torres Cacoullos and Aaron (2003) where they found English subjects to be categorically pre-verbal, and where rates of post-verbal subjects in monolingual Spanish and mixed speech were virtually identical.

5.5 Summary

Several diagnostics were applied to the data in an effort to determine whether speakers are engaging in borrowing or code-switching when they use single English-origin nouns in Spanish discourse. The former implies the operation of only one grammar, that of the recipient language, and, the latter, the simultaneous use of Spanish and English where the speaker switches from one

grammar to the other. The use of the masculine determiner with English-origin nouns, regardless of the gender of the equivalent Spanish translation suggests that speakers are using the Spanish option of the default masculine, i.e., treating the English words in the same fashion as Spanish nouns that are unmarked for gender (as in Harris' (1991) analysis) or that do not have sufficient information for gender (Eichler et al. 2012). When number agreement was examined, nouns were once again shown to be integrated into the Spanish system where plural determiners were used with plural nouns and singular determiners were used with singular nouns reflecting Spanish number agreement. Subject position also provided some evidence to bolster these findings as English-origin nouns in Spanish discourse were found to exhibit the same variability in subject position relative to the verb as is found in the Spanish data. Taken all together, this evidence suggests that the English-origin nouns are borrowings, given their morphosyntactic integration into the Spanish system. The very presence of concord and post-verbal subjects militates against an analysis of an English system governing the single English-origin nouns in Spanish data, thus reinforcing the borrowing analysis.

Other Spanish-English production data show multiple factors conditioning gender assignment on English-origin nouns for contexts similar to Belize in terms of long-term Spanish-English contact (Dubord 2004; Smead 2000). Multiple factors were not found as determinants of gender assignment on English-origin nouns in the current data. These findings indicate that speakers may apply the rules of Spanish grammar in community specific ways (Poplack et al. 1982). In addition, while bilingual ability and/or language dominance may be a determinant in gender assignment strategy, this was a question unaddressed in the preceding analysis. In experimental data, the masculine determiner has been reported to be the choice of English-dominant/Spanish L2 speakers, while Spanish-dominant native speakers of Spanish show a preference for mixed DPs that meet the analogical criterion (Liceras et al. 2013, 2008). The

speakers included in the current analysis were not found to be homogenous with respect to language dominance, as discussed in Chapter 3, however, there was no noteworthy variability in gender assignment strategy; they all used the masculine default. The masculine default may simply be the community norm. Nonetheless, the sample size in the current study must be considered and it points to the possibility that the differences in bilingual ability in this subset of speakers may not large enough for a distinction, if any, to be discerned in gender marking strategy. Assessing the borrowing patterns of the low language-mixers who are Spanish-dominant would be needed to determine if bilingualism and language dominance are correlated with differential strategies for gender marking of English-origin nouns in these communities.

CHAPTER 6 COMMUNITY PATTERNS OF LANGUAGE MIXING AND CONSTRAINTS ON

CODE-SWITCHING

6.1 Introduction

This major goal of this is chapter is to analyze English-origin multiword fragments which constitute 32% (N= 634) of the other-language material. The first objective is to determine the overall patterns of code-switching used by Belizean Spanish speakers given that different types and patterns of language mixing have been identified in the literature (Poplack 2004, 1987; Muysken 2000; Lipski 1985). Different patterns, for example, have been found in diverse communities even for the same language pair (Poplack 2004; Poplack et al. 1988). In addition, different patterns of code-switching and borrowing are reported to correlate strongly with factors such as bilingual ability (Bullock and Toribio 2012; Treffers-Daller 1994; Berk-Seligson 1986; Lipski 1985; Poplack 1980; McClure 1977), therefore the global patterns of language mixing in the current data are analyzed in detail to determine where individual speakers are situated vis à vis community patterns. All the speakers in the data analyzed here have been identified as high language mixers but, as noted in Chapter 3, they are not homogenous in terms of language dominance and self-reported level of bilingualism. Additionally, while they all code-switch, they are not all high code-switchers as defined for the corpus and discussed further below[191]. A second objective is to assess whether or not the equivalence constraint and free morpheme constraint are upheld in the data once intrasentential switches are identified and extracted. Related to intrasentential switching, the hypothesis that discourse markers introduce code-switches or act as pivots for code-switching (Aaron 2004) is also assessed. The chapter is organized in the following way: global patterns of code-switching in the

---

[191] Recall that language mixing is used as a cover term for both borrowing and code-switching and the language mixing rate, as defined for the corpus, refers to a count of *all* non-Spanish material in the data. The language mixing rate is a relative measure as compared to speakers in the same corpus. Thus, the terms high or low *code-switcher* are further defined in terms relative to the twelve speakers analyzed. As a result, high language mixers may potentially be low code-switchers as compared to their peers in the set of 12 interviews analyzed here. Normalized rates are established for the twelve speakers and discussed further in section 6.5 Community patterns and regional differences.

data are discussed first, followed by an analysis of intrasentential code-switching and discourse markers as code-switching triggers. As will be seen, a large portion of the data turned out to be ambiguous multiword fragments as, despite their multiword status, they exhibited some characteristics of borrowings. These are then analyzed in detail. Finally, regional differences and bilingualism are discussed in relation to patterns of language mixing followed by a summary discussion of the chapter's major findings.

6.2 General Patterns

As noted previously in Chapter 2, language mixing between typologically similar language pairs, or even the same language pair, may yield different patterns in different bilingual communities (Poplack 2004; Budzhak Jones 1998, 1996; Poplack 1987; see also Muysken 2000:221-249) highlighting the fact that linguistic constraints are not the only contributing factor to patterns of language mixing. In fact, linguistic constraints more accurately predict the type of mixing that is not permitted rather than the pattern that will be found in a given community (Budzhak Jones 1998,1996; Poplack et al. 1989; Poplack 1987). Poplack's (2015, 2004) typology of language mixing patterns is repeated below in (i) –(v) or convenience and discussed further below.

(i) Smooth Intrasentential code-switching or equivalence based code-switching (also referred to as "skilled" code-switching)
(ii) Flagged Switching
(iii) Tag switching (or what Muysken (2014) calls backflagging)
(iv) Constituent insertion
(v) Borrowing

These types of language mixing are not necessarily mutually exclusive, however, as more than one pattern can be found within the same community. Furthermore, these patterns appear to be determined by extralinguistic factors such as bilingual ability, social stigma, high levels of societal

bilingualism, the social status of the respective languages, or level of contact between the languages in question (see Poplack et al. 1988 and Poplack 1980 for detailed discussions).

In Puerto-Rican Spanish, for example, Poplack (1980) found the dominant pattern to be (i) smooth Intrasentential code-switching[192] but also found (iii), tag switching, and "noun switching" which is essentially pattern (v): borrowing[193]. In Canadian French-English data, patterns (ii), "flagged" switching, and (v), borrowing, were found (Poplack 1987). Despite the fact that both language pairs are typologically similar, and therefore share a number of equivalence sites where intrasentential switches could occur, this type of switching is predominant only in one community and virtually absent in the other. Discussed below are patterns relevant to the current analysis.

According to Poplack (1987:54; 1980: 583), "smooth" or equivalence based code-switching, is characterized by the smooth alternation between L1 and L2 without any flagging such as false starts, hesitations, or lengthy pauses. In addition, an absence of these types of "flags" is presumed to evince a general unawareness on the part of the speaker with regard to their code-switching, thus, these types of switches tend to also be unaccompanied by metalinguistic commentary or repetition or translation of the adjacent segment. These switches occur within the clause or intrasententially and require the speaker to navigate the syntactic boundaries when alternating between languages. The equivalence constraint predicts that switches will only occur at points where the alternation does not violate the syntax of either language, i.e., at equivalence sites between the grammars or where the grammars are equivalent; therefore, this type of intrasentential switching has also been called equivalence based switching. It is widely accepted that this type of switching is in the domain of speakers who are highly bilingual (Bullock and Toribio 2012; Treffers-Daller 1994; Myers-Scotton 1993; Berk-Seligson 1986; Lipski 1985; Poplack 1980; McClure 1977).

---

[192] See also Torres Cacoullos and Travis (2013) on New Mexican Spanish.
[193] Though at the time 'noun-switching' was not yet analyzed as borrowing.

197

Smooth intrasentential switching constitutes the ideal data for assessing constraints since speakers must navigate syntactic boundaries. This type of switching is exemplified below in (136) with data from Poplack (1980:597). English portions being discussed in examples appear in **boldface.** Where Spanish items or portions in examples are being discussed, these are underlined.

136. No tienen ni tiempo **sometimes for their own kids and you know who I'm talking about** (04/17).
    'They don't even have time sometimes for their own kids and you know who I'm talking about.'

In contrast to this "smooth" code-switching is "flagged" switching (Poplack 1987) where speakers produce hesitation or fillers before the switch. In addition, Poplack (1987) considers metalinguistic commentary, repetition or translation to also "flag" the code-switch which not only draws attention to the switch but also interrupts "the speech flow at the switch point, rendering a grammaticality requirement irrelevant" (Poplack 2004:593). For this reason, "flagged" switching is generally considered inappropriate for assessing constraints on code-switching. Flagged switching has been reported for Canadian French-English data (Poplack 1987) where high social stigma may be the contributing factor for "flagging" at the discursive level or what Poplack (2004:593) calls "functional flagging" as in (137) (*mobile home*) below, where the speaker translates the English item into French. "Flagged switching" has also been reported for Finnish-English data (Poplack et al. 1989), however, in this community the flagging consists of phenomena such as hesitations, pauses and ratification markers which Poplack (2004, 1993) associates with production difficulties and the absence of a community norm of borrowing and code-switching. Example (137) below exemplifies translation flagging from her French-English data (Poplack 1987: 61).

137. J'ai acheté une roulette, **mobile home** là, une maison mobile (GF/83)
    'I bought a trailer, a mobile home, a mobile home'

Though not listed as a pattern in Poplack's typology, intersentential switching is also considered a type of code-switching. This type along with pattern (iii), extra sentential or tag switching (e.g. *you know*), does not require navigating syntactic boundaries since intersentential switching occurs between independent clauses and tag switching tends to occur at clause boundaries. These types of switching might be found more readily in the speech of those speakers that have lower bilingual ability (Bullock and Toribio 2012; Lipski 1985; Poplack 1980; McClure 1977)

When a language pair shares typological similarity, there is a strong possibility that code-switching will be found in the speech of bilinguals since there would be numerous equivalence sites or permissible switch points between the two languages (Poplack 2015, 2004). However, typology alone cannot predict what patterns of language mixing will be found in any given community (Poplack 2004). In addition, as noted above, different social factors, such as bilingual ability, correlate with certain patterns and not others, it is therefore important to ascertain the distribution of the types of language mixing and major patterns of mixing in the communities under study. This is of particular interest here since Spanish in Belize is highly understudied and it is not known how it might compare to other bilingual communities or if the two dialectal regions will pattern in the same ways. So far the analysis has shown that Belizean Spanish speakers engage in borrowing nouns, consistent with many bilingual situations and that they also borrow verbs via BCVs with apparent unusual frequency. Multiword fragments are now examined.

The coding schema utilized here is adapted from Poplack's (1980) study of Puerto Rican (henceforth PR) Spanish-English bilinguals and from a similar study on French-English speakers (Poplack 1987). In addition, I adopt Poplack's (2004) distinction of functional flagging versus other flagging such as false starts, hesitation and so forth. The kind of smooth, skilled switching found with the PR speakers was reported to be minimal with the French-English bilinguals in Ottawa Hull

(Poplack 1987:60) [194] and therefore, the French-English data was coded in detail at the discourse level for discourse-pragmatic functions or "functional flagging" as well as other "'flagging" which includes hesitations, false starts etc.[195] I use the categories from both these studies to code the current data as this permits an overview of the data with regard to types of linguistic code-switching as well as the potential discourse-pragmatic functions of code-switching that might be present in the community patterns of language mixing. In the first instance, I separate out all alternations from Spanish to English such as in (138) and (139) that have a false start, pause filler, or hesitation at the boundary of the alternation from Spanish to English. As argued by Poplack (2004), these interrupt speech flow and may render the syntactic boundary irrelevant.

Non-smooth (preceding the alternation to English)

Pause fillers
138.   no, es, *ehm*, **mixed with science.**
　　　'No, es, ehm, mixed with science.'

(BSEK/29W/29W/00:00:37)

False starts
139.   En ***payp-,* el paypal account**
　　　'In *payp-,* el paypal account'

(BSEK/45NJ/45NJ_43NJ_3/01:05:46)

Table 31 [196] below summarizes the distribution of smooth and non-smooth alternations from Spanish to English. Speakers engage in smooth alternations from Spanish to English unmarked by

---

[194] It is important to remember that in Poplack's (1980) landmark study, single words and multiword fragments were both subsumed under code-switching. This chapter deals only with multiword switching with the exception of the discussion of borrowing rates as a community pattern and discourse markers as pivots. In addition, the PR data was not coded at the discourse level in the same way as the Ottawa-Hull data. In fact, translations, metalinguistic commentary, proper nouns and switches at a turn boundary were considered "non-code-switches" and excluded from the data (Poplack 1980: 598). While Ns for these categories are not reported, given that the variationist method aims to identify and analyze major patterns in the data, these categories were most likely relatively small in comparison to the smooth, skilled switching.

[195] Poplack (1987:60) included eleven categories in exemplifying the distribution of switches in the French-English data: expression/"mot juste," meta linguistic commentary, bracketing English segment, repetition/translation/explanation, reported speech, proper name, change in interlocutor, false start, turn boundary, sentential and intrasentential.

[196] 35 multiword fragments were analyzed in Chapter 4 as BCVs and are not included here. In addition, 39 tokens were excluded on the basis that the token was difficult to parse even after re-listening to the audio and examining more

hesitations, false starts, or pause fillers 94% of the time, which is comparable to the PR data

(Poplack 1987, 1980).

Table 31: Distribution of smooth and non-smooth alternations from Spanish to English

	N	%
Smooth alternation	527	94
Non-smooth	33	6
TOTAL	560	

Once the non-smooth alternations were removed from the data, I coded the remaining multiword

fragments for what I call *linguistic code-switching* which includes extrasentential, intrasentential,

and intersentential code-switching. This is in contrast to functional flagging or *discourse functions*

of code-switching as reported by Poplack (1987:60) which include translation/repetition, reported

speech, metalinguistic commentary, proper names, expressions, bracketing, and switches at turn

boundaries or changes in interlocutor. All of these categories are explained in detail and

exemplified further below.

A fourth category emerged from the data that did not serve discourse pragmatic functions

but were not easily coded for linguistic code-switching as they seemed to show characteristics of

both borrowing and code-switching. These are therefore termed *ambiguous MW fragments* which

are contrasted with unambiguous code-switches. Both these categories are also defined in detail in

the coding protocol described in the following section.

6.2.1 Coding protocol

The difference between intrasentential switching and intersentential switching is defined differently

by different researchers and as a consequence there is a lack uniformity in what is included under

---

context. This is illustrated in the following example where text with a question mark inside curly brackets indicates the transcriber's guess as what was said:
 {cuándo está?} the last campaign {?} public Central Park {BSEK/45NJ/45NJ_43NJ/41:03.5).

intrasentential switching (Deuchar 2010:703). I have chosen to adopt a clause-based approach and unambiguous intrasentential code-switching is defined as multiword fragments (hereafter MW) that contain both lexical and functional items from the donor language, that are internally consistent with the donor language (Sankoff et al. 1990:95)[197,198] and that occur within the clause or between a main clause and subordinate clause. Included here are adverbial, relative and complement clauses traditionally considered subordinate clauses. The first two clause types generally modify elements within the main clause (or the entire clause) while complement clauses are arguments of a verb in the main clause. Also included are those clauses that are marked by subordinators (e.g. *porque/because*). These clauses were infrequent, however, following Palmer (1987), to distinguish them from coordinated clauses, if the subordinate clause could prepose the main clause then they were considered subordinate and coded as intrasentential. Examples of unambiguous Intrasentential code-switching are shown below in (140)-(141).

Unambiguous intrasentential code-switching

140.	pero	es	algo	que	yo	soy	**so**	**sure**
	but	be3SG.PRES	something	that	I	be1SG.PRES	so	sure

	about	**it**	que	sólo	le	hago	blurt-out
	about	it	that	only	CL	do1SG.PRES	blurt.out

---

[197] This definition is not intended to suggest that single word switches are somehow prohibited, in fact, the tools of the variationist method allow for the identification of potential single word switches (on single word switches identified in variationist studies see Turpin 1998:230; Poplack and Meechan 1995: 212; Meechan and Poplack 1995: 184, 186-187; Poplack et al. 1988:53).

[198] Some multiword fragments are surrounded by Spanish discourse as in (140), however, many MW fragments appear at the end of a clause as in (141), or the speaker continues in English for several clauses before changing back to Spanish, often at an intersentential boundary. This means that some tokens as in (140) could be coded for two switches, the first switch from Spanish to English and the second, when they switch back into Spanish from English while others would only be coded once, the switch from Spanish into English such as (141). Given that this thesis does not investigate differences in Spanish-English switch sites versus English-Spanish switch sites, and, in order to avoid coding the same token twice, only the alternation from Spanish into English is coded. Additionally, the coding schema is intended to stay close to those found in Poplack (1980) and Poplack (1987) and in the former, it is not clear how the coding was done for language directionality. It is unclear, for example, if tokens such as (140) above were coded both for both the switch into English and the switch back into Spanish. As a consequence, the decision was made to analyze only the alternations from Spanish to English in this thesis.

'but it's something that I am so sure about it that I just blurt it out'

<div align="right">(BSEK/08W/08W_2/00:14:46/)</div>

141. | Habíamos | dicho | que | **we** | **want** | **to** | **downsize.** |
| --- | --- | --- | --- | --- | --- | --- |
| have1PL.IMP | say.PTCP | that | we | want | to | downsize |

So	**we**	**are**	**downsizing**.
so	we	be	downsize

'We had said that we want to downsize. So we are downsizing.**'**

<div align="right">(BSEK/45NJ/45NJ_43NJ_3/01:33:13/)</div>

Ambiguous code-switching, on the other hand, refers to MW fragments that do not contain function words and that consist largely of two-word fragments even though they occur intrasententially. Due to the short length, the absence of function words and the overlap with Spanish word order in many cases, their internal consistency is difficult to assess and they are analyzed separately. Examples of alternations involving ambiguous MW fragments can be found in (142) and (143).

Ambiguous multiword fragments

142. Fui con un **private doctor**, con Doctor Cano
    'I went to a private doctor, to Doctor Cano'

<div align="right">(BSEK/05N14/05N14/00:03:57/)</div>

143. porque mi pelo siempre estaba ***healthy, healthy***[199]
    'Because my hair was always healthy, healthy'

<div align="right">(BSEK/13WJ/11WJ_12WJ_13WJ/00:02:32)</div>

---

[199] The repetition of the adjective to intensify the meaning of the adjective is reminiscent of reduplication in Creoles (Greene 1999) but is not necessarily a pattern restricted to Creoles. There were only two examples of this type of reduplication in the data.

Intersentential switches were considered to be all those instances where the speaker changed to English at the start of a new clause following a Spanish clause, this includes where two independent clauses were coordinated.

Intersentential code-switching

144. Después digo: ¿Será?   ¿Será?   **It's**   **a**   **big**
      after   say1SG.PRES   could.it.be   could.it.be   it's   a   big

      coincidence
      coincidence

      'Then I say: Could it be? Could it be? It's a big coincidence.

(BSEK/58N/04N_58_N/01:00:38/)

Extra-sentential switches include interjections (*oh my God!*) and MW discursive fragments as in (145) below. They serve a discourse-pragmatic function and, as "freely moveable constituents", they are less bound by the syntax and do not require the speaker to resolve any potential conflicts between the two grammars (Poplack 1980:598).

Extra sentential switching

145. porque   son   personas   que   admiro,   **you**
      because   be3PL.PRES   people   that   admire1SG.PRES   you

      know.
      know

      'because they are people that I admire, you know.'

(BSEK/08W/08W_03/00:24:05/)

At the discourse level, some, but not all, of Poplack's (1987:60) categories were found in the current data, these include: translation (146), reported speech (147), proper nouns[200] (148), and

---

[200] Proper nouns are normally excluded (Torres Cacoullos and Aaron 2003: 296; Poplack 1980), however, for the purposes of comparison they are included here as in Poplack (1987: 60)

turn boundary/change in interlocutor (149-150)[201]. It should be noted that what I have termed *linguistic code-switching* and *functional flagging* are not mutually exclusive categories. In (146), for example, the speaker changes to English between a preposition and an NP, which is a permissible switchpoint as per the equivalence constraint but it is followed by a translation/repetition. Therefore, rather than being included in one of the above categories pertaining to linguistic codeswitching it was coded for discursive function as per Poplack's (1987) categories.

Translation/repetition

146. Nosotros dejamos a ese cuate tres meses en **intensive care**. **Three months in intensive care** el chamaquito.

    'We had that guy in intensive care for three months. Three months in intensive care that kid.'
    <div align="right">(BSEK/45NJ/45NJ_43NJ_3/01:01:09/)</div>

Reported Speech

147. y les digo: "¿y abrieron la ventana?" dice: "no", dice. **"Well don't do that."** le digo: "no vayan a abrir la ventana"

    'And I say to them: "and did you open the window?" he says: "no", he says. "Well don't do that". I say to him: "don't open the window".'
    <div align="right">(BSEK/05N14/05N14/00:18:45)</div>

Proper nouns

148. y un día llegaron allá en **House of Culture**
    'And one day they arrived there at the House of Culture.'
    <div align="right">(BSEK/58N/04N_58_N/00:44:52)</div>

Turn-taking/ change in interlocutor/answer to interviewer question

149. Interviewer: ¿Qué estabas haciendo **study** aquí?
    Participant**: 'IT, information technology'**

    Interviewer: What were you studying here?
    Participant**:** 'IT, information technology'

---

[201] One category not included as a "flag" is discourse markers. Poplack (1987) considers discourse markers to flag code-switches, I do not make this assumption at the outset given that this is something that must be empirically assessed in the current dataset. Additionally, as the discussion will show, discourse markers are infrequent in the MW fragments.

150.   {laughter} **stop laugh!** {directed to another bilingual participant in the room}
<div align="right">(BSEK/38W/38W_1_2/00:03:56/)</div>

By coding for many of the same categories as reported in previous variationist studies on language mixing (Poplack 1987,1980) it becomes possible to see overall patterns in the current data and, more importantly, the proportion of the different types of switching. This additionally permits cross-community comparison.

6.2.2 Results

Table 32 below summarizes all types of linguistic code-switching found in both communities. The largest category of MW alternations is comprised of ambiguous MW fragments (45%). This category consists largely of two-word English-origin fragments and will be discussed separately below. Intrasentential switching (17%) and intersentential switching (14%) are the next largest categories and occur at approximately equal rates. Extra-sentential or tag switching is minimal at 2% of the data and is accounted for almost exclusively by one speaker (8/13 tokens).

Table 32: Overall distribution of types of code-switching in both communities

	N	%
TYPE OF ALTERNATION		
LINGUISTIC CODE-SWITCHING		
Ambiguous	234	45
Intrasentential (unambiguous)	**90**	**17**
Inter Sentential	**76**	**14**
Extra sentential	13	2
DISCOURSE FUNCTIONS		
Reported speech	38	7
Proper noun	37	7
Turn boundary/change in interlocutor	28	5
Translation/repetition	11	2
Grand Total	**527**	

The above distribution now permits a brief comparison to the Spanish-English data from New York Puerto-Ricans and the Canadian French-English data. The pattern of linguistic code-switching exhibited by Belizean Spanish speakers is in stark contrast to the French-English speakers who were reported to engage in intrasentential switching only 1-6% of the time, depending on the community (Poplack 1987: 60). The largest categories reported for discursive functions in the French-English data (Poplack 1987:60), i.e., expression/"mote juste" (13-22%) and metalinguistic commentary (9-36%) are absent in the current data. Furthermore, no single category for discursive function in the current data exceeds 7%. Thus, Belizean Spanish speakers appear to be more similar to the PR speakers who engage in equivalence or skilled and smooth code-switching and who engage in flagged switching only 6% of the time (Poplack 1980).

Before moving on to analyzing code-switching patterns by region, it is worth noting that functional flagging, when found at low rates as in the current data, may simply be a characteristic of bilingual speech in general. Bilingual speakers have been found to translate or repeat words or segments with apparent unawareness into the L2 when there were originally uttered in the L1 Lipski (1985: 29)[202]. This is exactly what is evidenced in the current data. First, the translation data is sparse and two speakers (45NJ and 08W) account for almost all the tokens. In contrast to the Canadian French-English speakers, translation in the current data is not used to draw attention to the switch or to switch back into Spanish after what might be termed a "momentary lapse" (Poplack 1987:64). The use of translation to switch back into French is exemplified in (151) below for French-English mixed speech (Ibid).

151. J'ai acheté une roulette, **mobile home** là, une maison mobile (GF/83)
'I bought a trailer, a mobile home, a mobile home'

---

[202] Recall as well that these types of discursive functions were also noted to some extent in the Puerto Rican data (Poplack 1980). Rates were not reported as they were largely removed from analysis precisely because they were considered non-code-switches (see Poplack (1980: 598) on "non-code-switches").

In the current study, speakers tend to partially translate or repeat the English-origin fragment using mixed speech rather than monolingual Spanish as in (146) above and in (152) below. The segment *they were going to crash* gets repeated and translated into Spanish as *iba a hacer crash.* There does not appear to be a consistent attempt to "stay" in or switch back to Spanish, to avoid mixed speech, or to translate the English fragment into monolingual discourse as has been reported for French-English speakers.

152. **They were going to crash,** iban a hacer **crash**
     they be   going to   crash go3PL.IMP to do.INF crash

     'they were going to crash, they were going to crash'

     (BSEK/45NJ/45NJ_43NJ_3/01:03:35/)

This is a strategy also observed in other data from Belizean Spanish speakers (Fuller Medina 2005a:134) where speakers show apparent unawareness that they are engaging in mixed speech. In example (153) below (Fuller Medina 2005a), even where there is metalinguistic commentary in the preceding clause, the speaker settles on a bilingual verb without any flagging immediately preceding the bilingual verb or additional translation to the Spanish verb. There is no indication of an awareness that the bilingual compound verb constitutes mixed speech. This suggests that speakers are treating the individual words *crash* and *wink*, in the bilingual compound verbs, and even possibly *intensive care*, in the examples above, as Spanish items. Rather than being a repair strategy or flag, these examples may simply be repetition.

153. ¿Cómo se dice esto? {pointing to picture} **Wink**. Está haciendo **wink** el ojo. (08BZSI)
     'How do you say this? Wink. He's winking his eye.'

It appears that when functional flagging is the prevalent pattern, and includes metalinguistic commentary such as English bracketing (e.g. *like they say in English)*, as in the French-English communities, then these types of functional flagging do serve the purpose of drawing attention to

the switch and indicating the speaker's awareness of the switch as noted by Poplack (1987). Other

types of flagging have been found to signal production difficulties and possibly indicate that code-

switching is not a community norm (Poplack et al. 1989). Thus, while some of these discourse

pragmatic functions can be indications of production difficulties or social stigma, they can also be

expected, at low rates, as a normal part of bilingual speech. These features may even be

epiphenomenal of the interview process itself as reported speech, for example, is a part of

recounting narratives and speakers may (or may not) switch when reporting speech. In the

example below, for instance, there are four segments of reported speech as the speaker recounts

a story; only one of four segments is in English. Thus, code-switching is not necessarily used for

the express purpose of reporting speech; though it can be, just as it is used in other contexts

without necessarily being used to serve a specific discursive function.

154.   y les digo: "¿y abrieron la ventana?" dice: "no", dice. **Well don't do that.** le digo: "no
       vayan a abrir la ventana"
       'And I say to them: "and did you open the window?" he says: "no", he says. "Well don't do
       that". I say to him: "don't open the window".'

That different communities employ these "flags" in dissimilar ways also forms part of the

community norm. The function of such "flags" is revealed by detailed quantitative analysis in which

the types of fagging are determined as well their respective rates of use in relation to overall

patterns. Since linguistic code-switching, in particular intrasentential code-switching and

ambiguous multiword fragments, constitute the major patterns in the data, this is discussed in the

following two sections followed by a detailed discussion of potential regional differences with

regard to linguistic code-switching, borrowing.

6.3 Unambiguous Intrasentential Switching and Constraints on Code-switching

A major premise of this thesis is that single English-origin items and multiword English-origin fragments must be kept separate in the data until their status can be determined with regard to borrowing and code-switching. This is because "we cannot falsify the prohibitions against switches at certain points, nor can we substantiate positive predictions about switches at other points, unless we can prove that the attesting examples are not borrowings." (Sankoff et al. 1990: 93). Single English-origin items have now been analyzed as borrowings and unambiguous intrasentential code-switches have been identified and extracted from multiword fragments. The current data can now be examined for adherence to the equivalence constraint and the free morpheme constraint. For convenience, these constraints are repeated below:

Free Morpheme Constraint (hereafter, the FMC)

Codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme. This constraint holds true for all linguistic levels but the phonological (Poplack 1980 585-586)[203]

"Included under this constraint are idiomatic expressions, such as cross my fingers [sic] and hope to die and Si Dios quiere y la virgen ('God and the Virgin willing') which are considered here to behave like bound morphemes in that they show a strong tendency to be uttered monolingually." (Poplack 1980:585)

Equivalence Constraint (hereafter, EC**)**

"Code-switches will tend to occur at points in discourse where juxtaposition of L1 and L2 elements does not violate the syntactic rule of either language, i.e., at points around which the surface structures of the two languages map onto each other. According to this simple constraint, a switch

---

[203] This means that if a non-native lexical item is rendered in one syllable with native phonology and another syllable in non-native phonology, this is considered transference of phonological patterns and not a "switch" between bound morphemes. For example, in the production of *excuse* with aspiration of /s/ before a voiceless consonant in the first syllable with English pronunciation of the second syllable is transference of the aspiration tendency of /s/ in Caribbean Spanish (Poplack 1980:586). The issue of phonology is further clarified: "…a switch may not occur between a bound morpheme and a lexical form unless the latter has been phonologically integrated into the language of the bound morpheme." (Sankoff and Poplack 1981:5).

is inhibited from occurring within a constituent generated by a rule from one language which isn't shared by the other." (Poplack 1980: 586)

There is general agreement that code-switches may take place at major constituent boundaries: NP, VP, PP, (Halmari and Regetz 2012; Poplack 2004, 1980; Eze 1997; Wei 1994; Woolford 1983; Lipski 1985) and, in addition, because Spanish and English share many equivalence sites with respect to word order, code-switching can also take place within the major constituents if it is between determiner and noun, preposition and noun, verb and object and so forth. There are, however, some conflict sites between Spanish and English related to word order where, according to the EC and FMC, code-switches would be prohibited. The adjective–noun boundary, for example, is considered to be a prohibited site (Poplack 1980, Timm 1975) because, with the exception of a limited number of qualitative adjectives, Spanish adjectives are prenominal while the reverse is true in English. Another boundary where switches are predicted to fail to occur is between the verb and object pronoun (Lipski 1985: 21; Poplack 1980; Wentz and McClure 1976; Pfaff 1979: 299; Timm 1975). This is because the clitic-verb dependency as well as the syntactic position of object pronouns differs for English and Spanish as discussed previously in Chapter 4. Furthermore, it was proposed early in the literature that clitic pronoun objects must be "realized in the same language as the verb to which they are cliticized, and in the position required by the syntactic rules of that language" (Wentz and McClure 1976 and Timm 1975 as cited in Pfaff 1979:303). Consistent with this constraint, Pfaff (1979) found that unadapted English-origin verbs did not occur in the data with objects that were pronouns. A third conflict site where code-switches are predicted to be prohibited is the site of negation (Lipski 2008: 233,1985: 22; Timm 1975). English utilizes *do*-support and Spanish does not. This is exemplified below in (155) where *do*-support is used in English (*don't do that*) but in Spanish it is not (*no vayan*). These last two conflict sites would be prohibited both by the EC and the FMC. Recall that clitics can be considered bound

morphemes and negation has been analyzed as a bound morpheme for Spanish as well (Zagona 1988 as cited in MacSwan 2014).

155. "Well **don't do** that" le digo, <u>no vayan</u> a abrir la ventana (BSEK/05N14/05N14/00:18:45)
      "Well don't do that". I say to him: "don't open the window".'

In order to assess the data for adherence to the constraints on code-switching, each unambiguous code-switch was coded for the switchpoint from Spanish to English. The coding procedures are outlined below.

6.3.1 Coding protocol

Following Eze (1997:161), the switch boundaries were coded for the lexical or functional category of the items that appear at the boundary rather than the constituents involved in the switch. Thus, each alternation was coded for both the Spanish category preceding the switch and the English category at the onset of the switch. In some cases a number of switches occurred between the same categories such as in (157) where the preceding category was a Spanish determiner and the following one an English-origin noun or noun phrase. In other cases a variety of categories in Spanish preceded the same type of English category, such as discourse markers, exemplified in (159). There were nineteen different switchpoints found in the data, some of which are exemplified below and followed by summary results.

Preposition and following category

156.

frijol,	carne,	eh,	con	un	poquito	de	**cabbage**	**or**	**lettuce**
beans	meat	eh	with	DET	bit	of	cabbage	or	lettuce

'beans, meat, eh, with a bit of cabbage or lettuce'

(BSEK/58N/04N_58_N/00:00:07)

Determiner + N/NP[204]

157.
es		como	un	**informal**	**way**	**of**	**English**
be3SG.PRES		like	DET.MASC.SG	informal	way	of	English

'it is like an informal way of English'

(12WJ/11WJ_12WJ_13WJ_2/00:11:03/)

Verb + adverb

158.
voy	a	ir	**on**	**my**	**own**,
go1SG.PRES	to	go.INF	on	my	own

'I'm going to go on my own'

(BSEK/45NJ/45NJ_43NJ_3/01:16:56)

Preceding category and discourse marker

159.
Oyí		que	había		unos	casos	**like**,	**one**	**after**
hear1SG.PRET		that	be3SG.IMP		some	cases	like	one	after

the	**next,**	**one**	**after**	**the**	**next**
the	next	one	after	the	next

'I heard that there were some cases like, one after the next, one after the next.'

(BSEK/58N/04N_58_N/00:49:09/)

Subordinator + following category in a dependent clause

160.
Yo	ya	sabía		porque	**I**	**work**	**the**	**figures**
I	already	know1SG.IMP		because	I	work	the	figures

"I already knew it because I worked the figures.'

(BSEK/45NJ/45NJ_43NJ_3/01:17:57)

---

[204] There were three tokens where the switch took place between a determiner and English-origin noun as exemplified in (i). Arguably, these English-origin nouns could be borrowings with the switch taking place after the English-origin noun. Tokens of this type have been reported in the literature (Torres Cacoullos and Aaron 2003; Eze 1997) though they were infrequent in the current data.

(i)   Pero no en el **windshield, it's against the law to put it on the windshield**.
'But not on the windshield, it's against the law to put it on the windshield'.
(BSEK/45NJ/45NJ_43NJ_3/01:32:11/)

Verb + NP

161. Acá   es   **free**   **internet**   **services**   **everywhere**.
     here   be3SG.PRES   free   internet   services   everywhere

     'Here it is free internet services everywhere.'

                                              (BSEK/58N/04N_58_N/00:04:28)

Complementizer + following category in a complement clause

162. Habíamos   dicho   que   **we**   **want**   **to**   **downsize.**
     have1PL.IMP   say.PTCP   that   we   want   to   downsize

     So   **we**   **are**   **downsizing**.
     so   we   be   downsize

     'We had said that we want to downsize. So we are downsizing**.'**

                                              (BSEK/45NJ/45NJ_43NJ_3/01:33:13/)

6.3.2 Results

As noted previously, there were 90 intrasentential switches identified in the data which is the largest category at 17% (if ambiguous switches are not considered). This finding is consistent with the literature on Spanish-English data where code-switching was found to be a discourse mode (Torres Cacoullos and Travis 2013; Poplack 1987, 1980). Table 33 below summarizes the boundaries at which the switches from Spanish to English take place. Most take place between a Spanish preposition and following category[205] (11%) and at the determiner-noun/noun phrase boundary (11%) followed by the verb-adverb boundary (9%). Other favoured switchpoints are between a verb and a noun (8%) and before or after a subordinator (8%). There were also a number of examples at various switch boundaries that did not fit the other categories because they were switches between some type of dependent clause and an independent clause such as at a

---

[205] The categories that followed the Spanish preposition were NP (N=5), time expression (N=2), and three tokens of the same idiomatic expression (en **black and white**).

boundary between two pronouns, as shown in (163) below, where the switch is between *mí* and

that[206]. This category was labelled "other".

163.
para	mí,	**that**	**means a**	**lot**	se	puede	decir
for	me	that	means a	lot	CL	can2SG.PRES	say.INF

'For me that means a lot, you can say'

(BSEK/08W/08W_03/00:15:20)

Another category includes switches that take place between various categories and a discourse marker such as *like*. While discourse markers can be considered to be "freely moveable" (Poplack 1980), rather than constituting a true syntactic boundary, these were coded in this way in order to capture any distribution of discourse markers at switch boundaries[207]. In so doing, it becomes possible to test the hypotheses that discourse markers trigger code-switching (Aaron 2004) or that they function as the locus for code-switching (Gardner-Chloros and Edwards 2002: 1442). These account for 6% of the data.

The data below show that many switches occur at major constituent boundaries: PP, NP, V though the majority take place within the PP or NP. This is keeping with the literature where for example, code-switching before an NP, is widely accepted as a permissible switch site (Eze 1997; Berk-Seligson 1986; Poplack 1980; Treffers-Daller 1991) as is the case of switching after a determiner for Spanish-English (Lipski 2008, 1985; Poplack 1980). All of the switch sites are ones predicted to be permissible by the EC. Recall that the EC predicts that switches will take place where "the juxtaposition of L1 and L2 elements does not violate the syntactic rule of either language, i.e., at points around which the surface structures of the two languages map onto each other." (Poplack 1980: 586). When the data was examined further, no switches were found at the

---

[206] Other switch boundaries include NP-PRO, *if* clause-independent clause, and isolated NP and following category.
[207] When examined further according to Poplack's (1980) coding procedure where the discourse marker is not considered to form part of the syntactic boundary, all switches were found to occur at permissible switch sites as predicted by the EC. In (159) above, for example, the absence of *like* would not render the switch ungrammatical according the EC.

prohibited sites which were previously outlined. Thus, there were no switches between object pronouns and the verb, at negation sites, or at the noun-adjective boundary. Thus, the equivalence constraint and the free morpheme constraint are upheld in the data.

Table 33: Switch boundaries for Spanish-English code-switching

	N	%
Preposition + following category	10	11
Determiner + N/NP	10	11
Verb + adverb	8	9
Verb + NP	7	8
Subordinator* + following category	7	8
Other	6	7
Preceding category + discourse marker	5	6
Complementizer + following category	5	6
Noun + verb	4	4
como, 'like' + following category	4	4
Adverb-adverb	4	4
Preceding category + relativizer	3	3
Aux + participle	3	3
NP + preposition	3	3
entonces, 'so' + following category	3	3
Verb + number expression	2	2
Modifier and pronoun or NP	2	2
Homophonous word[208]	2	2
Verb + Preposition	2	2
TOTAL	90	

*Includes both Spanish and English (*porque*, *but*)

---

[208] Muysken (2000:139) identifies words as homophonous if they are either very similar or the same in both languages in the language pair. Where words appear at these boundaries, it is difficult to identify the switch boundary.

6.3.3 Discourse markers as pivot points

As indicated in Chapter 3, the third largest category of single non-Spanish items in the data consists of items such as *so*, *like*, and *right*, which are generally subsumed under the term discourse markers[209] (Sankoff et al. 1997; Schifrin 1987).

164. Todo   estaba **like**   **normal** acá   adentro.
     all   be   like   normal   here   inside

     'Everything was, like, normal here inside.'

                                                                (BSEK/29W29W/00:11:22).

The term discourse marker has been used as an umbrella term to refer to a heterogeneous group of items from various word classes that optionally occur utterance initially, are independent of sentence structure and serve to ensure discourse coherence as well as bracket units of talk (D'Arcy 2007; Schiffrin 2003:57,1987). D'Arcy (2007:394-395) further differentiates between discourse markers and discourse particles, the latter occurring clause medially and serving to aid in "cooperative aspects of communication such as checking or expressing understanding"[210]. When such particles signalling cooperative communication are absent, discourse is said to be perceived as unnatural and even unfriendly (Brinton 1996:3 as cited in D'Arcy 2007: 395). The notion that the use of discursive features, whether markers or particles, aid in creating more 'natural' speech is supported by the findings of Sankoff et al. (1997:213,204) who maintain that the use of such

---

[209] Single items such as *so*, *like*, and *right*, are generally subsumed under the term discourse markers (Sankoff et al. 1997; Schiffrin 1987) and this is the term adopted here in keeping with previous variationist studies (D'arcy 2007; Sankoff et al. 1997). In the current data, these items generally occur in isolation surrounded by Spanish discourse as will be seen in the discussion on discourse markers further below. Some of these items, however, co-occur with other categories as part of a MW fragment and are therefore addressed in the discussion of multiword fragments. Expressions such as "you know" might also be considered to serve discourse pragmatic functions in a similar fashion to single word items, however, their status as tags or extra sentential switches is considered to be unambiguous and they were instead included with interjections under extra-sentential switching following Poplack (1980) (see also Lipski 1985:27).

[210] Schiffrin (1987:32) notes that some markers may occur freely at locations that are difficult to define syntactically.

features appear to assist in producing cohesive discourse, one that is more colloquial in nature, and one that is typical of native speaker discourse[211] (Sankoff et al. 1997).

As a result of the discourse-pragmatic role of discourse markers and their somewhat fuzzy nature in terms of morpho-syntax, lexical class, and functional category, they are notoriously difficult to assess for language membership in bilingual data (Aaron 2004; Lipski 2005; Torres 2002).[212] In the Spanish-English literature they have been termed borrowings (Aaron 2004; Torres 2002), a special form of code-switching (Lipski 2005), borrowings that start out as code-switches (Torres 2002) and as triggers for code-switching (Aaron 2004). Given the lack of morpho-syntactic integration and morphological marking, discourse markers do not lend themselves well to the kind of comparative analysis that was carried out with English-origin verbs and nouns and is not pursued here. Nonetheless, the claim that discourse markers trigger or introduce language alternation is examined in the data.

In an analysis of Spanish *entonces* and English *so* in the speech of Spanish-English bilinguals, Aaron (2004:173) analyzed the rates and functions of both discourse markers in monolingual and mixed speech and found that *so* occurred 16% of the time at a switch point between English and Spanish while *entonces* only occurred at a switchpoint 2% of the time. These findings were taken to mean that *so* may be a trigger for code-switching. In addition, Aaron (2004) suggests that based on frequency in monolingual Spanish, *so* may be an established borrowing in New Mexican Spanish and, consequently, a convergence site at which switching is facilitated between Spanish and English.

---

[211] Sankoff et al. (1997:193) analyzed the use of French discourse markers in the speech of Anglophone speakers of Montreal French. They considered the use of discourse markers to be a feature of local vernacular speech and a measure of integration into the local speech community given that discourse markers are generally not taught as a part of prescriptive language instruction.

[212] In fact, some discourse markers can be found to have counterparts with comparable functions across various unrelated languages, thus, presenting a challenge to a borrowing or code-switching analysis (Levey 2007:204).

In order to test the hypothesis that discourse markers, such as *so*, trigger code-switching, the position of discourse markers relative to non-Spanish material is assessed. If the marker *so* or other discourse markers trigger or introduce code-switching, then a preponderance of discourse markers at sites of alternation between the two languages should be evinced by the data. Table 34 presents the distribution of English and Belize Kriol discourse markers in monolingual stretches of Spanish as well as in mixed speech. For mixed speech, both ambiguous and unambiguous switches are considered[213]. Other discourse markers do not occur at intrasentential switch sites since items such as *right* and *though* occur clause finally, and *you know* (N=2) and *bway* occur as single items surrounded by Spanish discourse.

Table 34: English and Kriol single-word discourse markers in Spanish and mixed speech[214]

discourse markers	Monolingual Spanish		Mixed speech		TOTAL	
	N	%	N	%	N	
SO	83	98	2	2	84	33
LIKE	78	84	14	16	89	35
RIGHT	21	100	n/a		22	09
BWAY (*boy*)	17	100	n/a		17	07
Other (then, though, Gyal)	18	100	n/a		18	07
TOTAL	218	93	16	6		234

The most frequent discourse markers are *so* and *like*, followed by *right* which occurs categorically in clause final position. The only Kriol discursive elements are *bway* 'boy' which is used only by one speaker (45NJ)[215] and *gyal* 'girl,' which appears only twice. What is immediately obvious from

---

[213] When intersentential switches were examined it was found that only in 8/76 instances did a discourse marker appear at the onset of the non-Spanish fragment.

[214] Includes alternation boundaries for both unambiguous and ambiguous code-switches.

[215] The use of *bway* is very common in speech in Kriol and not at all unusual in the Spanish spoken in Belize. In this case it may be part of the idiosyncratic behavior of 45NJ but it is more likely a function of the interview and topic of conversation. First, another male participant was present and 45NJ directed most of his speech to the other participant

Table 34 is that speakers frequently make use of discourse markers but not in MW English-origin fragments. Speakers do not appear to make use of them to introduce switches into English; therefore, for these communities neither *so* nor any other marker appears to be a trigger for code-switching or alternations into non-Spanish multiword fragments. Only 2% of the time does *so* appear at the boundary of a multiword fragment. Instead of appearing at switch points, discourse markers appear almost categorically in isolation surrounded by Spanish discourse. This is not unlike other Spanish-English communities where *so* predominates in monolingual discourse (Aaron 2004; Torres 2002:80). Even in the New Mexican data, *so* appears 50% of the time in monolingual speech as opposed to 16% of the time at a switchpoint (Aaron 2004). If any marker were a candidate for a code-switch trigger in the current data, it would be *like*, however it should be kept in mind that 84% of the time it occurs in Spanish discourse. Overall, English-origin discourse markers do not appear at switch boundaries. Their function as triggers for code-switching is not confirmed.

6.4 Ambiguous Multiword Fragments

Recall from the discussion of Table 32 above that the majority of the MW switches were ambiguous MW fragments. While these occurred intrasententially, they were kept separate since they only partially fit the criteria employed for unambiguous intrasentential code-switching, i.e., multiword fragments that include function words and are internally consistent with the syntax of the donor language. It was found that 84% (196/234) of these tokens consisted of two-word fragments, and these are now examined in more detail to ascertain their status, particularly since they represent a large portion of the data overall (approximately 35% of MW fragments). More importantly, however, two-word fragments are not often analyzed separately in language mixing

---

(43NJ). The only other times *bway* was used was when he was reporting speech that was directed to a man. While other speakers may use *bway* as well, they would not have done so with me unless in reported speech.

studies unless, for example, they do not have dictionary attestation[216] (Torres Cacoullos 2003). In some cases, if the two-word fragment is considered to consist of "more than one word functioning independently" then it is included in code-switching data (Budzhak Jones 1996: 26). However, rather than grappling with the decision as to whether the words are functioning independently, it may be critical to treat such data at the outset as ambiguous and subject it to separate quantitative analysis in order to determine its status. Non-native two-word fragments such as adjective + noun constituents, despite their characterization as MW, may in fact behave as a single unit, i.e., as borrowings (Torres Cacoullos and Aaron 2003; Eze 1997:169; Sankoff et al. 1990:80). If these MW fragments are not true code-switches then they are not relevant for testing constraints; automatically classifying them as code-switches could give a false result with regard modelling code-switching (Torres Cacoullos and Aaron 2003: 320; Budzhak Jones 1996:30-31).

6.4.1 Coding protocol: Constituent type

Each ambiguous MW fragment was coded for switch point from Spanish to English as was done for unambiguous code-switches. In addition, each token was coded for type of constituent of the MW fragment.

Discourse marker and other category

The most frequent discourse marker is *like* which generally precedes another English-origin item.

Discourse marker + adjective

165.	Y	siempre	yo	me	ponía	**like**	**excited**
	and	aways	I	CL	become	like	excited

---

[216] Alternatively, they are included with single English-origin items if they are of the ADJ + N type, conjoins (e.g. *daily bread*) or lists (e.g. *rock and roll*) (Otheguy and Lapidus 2003:211).

y todo	eso,	right?	
and	all	that	right

'And I used to always get, like, excited and all that, right?'

<div align="right">(BSEK/11WJ/11WJ_12WJ_13WJ/58:55.4)</div>

Discourse marker + noun

166.
mi	esposo	le	gusta	sembrar	**like,**	**plants**
my	husband	CL	likes	plant.INF	like	plants

que	da	frutas	para	comer
that	give	fruits	for	eat.INF

'My husband likes to plant, like, plants that bear edible fruits'

<div align="right">(BSEK/05N14/05N14/ 00:12:51)</div>

Discourse marker + time expression

167.
y	**like**	**three**	**weeks**	después,	ehm,	mi	cuñada
and	like	three	weeks	after	ehm	my	sister.in.law

me	llamó
CL	called

'And like three weeks after, ehm, my sister-in-law called me.'

<div align="right">(BSEK/05N14/05N14 /00:22:49)</div>

Noun + "right"

168.
¿los	representa	qué?	¿los	**fish**?	¿**right**?	dos	fish.
what	represents	them	the	fish	right	two	fish

'What represents them? The fish? Right? Two fish?'

<div align="right">(BSEK/08W/08W_03/19:36.0)</div>

Numbers in an NP or adverbial expression[217]

Dollar amounts

169. yo   iba   ganar   como   **seven hundred**   **and**   **fifty**
      I   go1SG.IMP   win.INF   like   seven   hundred   and   fifty

      'I was going to win like seven hundred and fifty.'
                                     (BSEK 13WJ/11WJ_12WJ_13WJ/00:38:03)

Temporal adverbial expressions

170. ya   cuando **eleven thirty,**   los   voy   a   buscar
      when   eleven thirty   CL   go1SG.PRES   to   pick.up.INF

      'When it is eleven thirty, I go to pick them up'
                                       (BSEK/05N14/05N14/00:39:06)

Ordinal number + noun

171. Como   fue   en   el   **third**   **prize**,   y   como   el   boleto
      since   was   in   the   third   prize   and   since   the   ticket

      era   de   ten   piece,
      was   of   ten   piece

      'Since it was the third prize and since the ticket was for ten pieces,'
                                   (BSEK/13WJ/11WJ_12WJ_13WJ/00:38:09/)

Non-temporal adverbial

172. uno   que   se   llama   Crazy   Disc,   te   lleva   **upside down**,
      one   that   CL   call   Crazy   Disc   CL   carries   upside   down

      'One that is called Crazy Disc, it takes you upside down'
                                   (BSEK/35W/35W/00:23:56/)

---

[217] The tokens of the type [ordinal number + N] could also have been simply coded as NPs/modified NPs but it became apparent that numbers were often expressed in English and, given that in some language mixing scenarios numbers are consistently expressed in one language and not the other (e.g. Yakpo (2009) on Pichi-Spanish mixing; see also Matras (2007: 50-53)), I opted to code all MW fragments with numbers in one category.

Name of institutions, or concepts associated with institutions

173. porque   en   **high**   **school** es   un   requirement   y
     because in   high   school   is   a   requirement   and

     mi   esposo   es   un   Spanish   teacher
     my   husband   is   a   Spanish   teacher

'Because in high school it is a requirement and my husband is a Spanish teacher.'

(BSEK/05N14/05N14/00:43:56)

Noun phrases/modified nouns

Adjective + noun

174. Fui   con   un   **private doctor,** con   Dr. Cano.
     went   with   a   private doctor   with   doctor Cano

'I went to a private doctor, to Dr. Cano.'

(BSEK/05N14/05N14/00:03:57/)

Other

175. Y   como   era   un   **market before.**
     And   since   was   a   market before

'And since it was a market before.'

(BSEK/58N/58N/30:01.2)

176. porque   me   pelo   siempre   estaba **healthy,**   **healthy**
     because my   hair   always   was   healthy   healthy

'Because my hair was always healthy, healthy.'

(BSEK/13WJ/11WJ_12WJ_13WJ/00:02:32)

6.4.2 Results

Table 35 summarizes the distribution of ambiguous MW fragments in the data with the largest categories highlighted in **boldface**.

Table 35: Ambiguous MW fragments: cross tabulation of switch points and constituent of the switched fragment

Constituent Switch point	NP	Number Expression	DM and other category	Insti-tution	Other	ADV	TOTAL N %
DET + [Adj +NP], DET + NP	70	16	3		4		**93 40**
Preposition + following category	16	8		24		1	**49 21**
VERB + NP	24	1	1		4		**30 13**
VERB + # expression		**13**					13 6
COMO + following category	7	**4**			1		12 5
discourse marker + preceding or following category			10				10 4
Modifier + NP or other modifier	4	1	2				7 3
Other	2		4		1	1	6 3
como + like			4				4 2
Verb + following category	1		1			**1**	3 1
ADJ-ADJ					3		3 1
Conjunction		1	1				2 < 1
Grand Total	N % **124 53**	N % **44 19**	N % **26 11**	N % **24 10**	N % 10 4	N % 4 2	234

Most of the data clusters around modified nouns as in (174) (*private* doctor) above, accounting for approximately half of all ambiguous switches (53%) These constituents occur primarily after a determiner (N=70), the most frequent switchpoint overall. The most frequent switchpoints are as follows: at the boundary between determiner and NP (40%), preposition and following category (21%) and between the verb and an NP (13%). These are all equivalence sites between Spanish

and English and sites that are predicted by the EC to permit code-switching. In terms of constituents, modified NPs (N=124), some of which are possibly compounds, together with number expressions (N=44), and MW fragments that include a discourse marker (N=26) are the largest categories followed by two-word fragments denoting institutions (N=24).

To assess system membership of these constituents, the same diagnostics applied to single English-origin nouns are applied. If these two word items are borrowings as has been shown for single English-origin nouns in Chapter 5, then just as in the case of English-origin nouns, they should show evidence of integration into Spanish. For the previously analyzed English-origin nouns, evidence of integration was found in the use of the masculine default, number agreement, as well the presence of post-posed subjects. If the ambiguous MW fragments are borrowings then they should also show masculine determiners, where determiners are present, and where the noun is plural and has an overt determiner, number agreement. A fourth conflict site, not previously applied, is that of determiner marking of nouns denoting institutions. In a study of New Mexican Spanish, it was found that institutions tend to be zero marked in English for determination while in Spanish they surface with an overt determiner (Torres Cacoullos and Aaron 2003). Consequently, those ambiguous MW fragments that refer to institutions are also coded for determiner marking. All nouns in the Spanish dataset that denote institutions were also coded for determiner marking. If the ambiguous tokens denoting institutions pattern with the Spanish nouns denoting institutions with regard to determination, then this is evidence of borrowing. The largest categories of the ambiguous MW fragments are assessed in the ensuing sections.

Modified nouns

As previously noted, the largest category of ambiguous MW fragments is modified NPs such as in (174) above. These fragments are largely two-word fragments and were initially categorized as ambiguous because they did not contain function words and only contained a noun

and an adjective. The predominance of switches in this category raises the question of whether or not these two-word switches may be collocations or compounds which speakers are treating as single units in a similar fashion to the single English-origin nouns (Torres Cacoullos and Aaron 2003)[218]. Given that these tokens are almost exclusively of the [adjective+ noun] type, a second possibility is that these two-word fragments are code-switches since they adhere to English word order.

The majority of the examples are consistent with English word order where the adjective is preposed to the noun, while a small percentage (11%) was consistent with both Spanish and English word order (e.g. *half English, '*mitad inglés'*; certain hour*, 'cierta hora'). There is still the possibility that they could be borrowed as constituents, therefore, each token was checked for dictionary attestation using online Oxford dictionaries. Of the 127 tokens 17 were found to be dictionary attested, as in (177), and these can be considered borrowings.

177.   Cuando Toni tenga cinco años pueden conseguir sus **credit cards**
       'When Toni turns five you can get your credit cards'

(BSEK/45NJ/45NJ_43NJ_3/01:03:49)

The remaining 110 tokens were examined for number agreement and gender of the determiner[219]. As with the single English-origin nouns, these two-word fragments surfaced most of the time with an overt determiner. There were 58 tokens that appeared with definite or indefinite determiners which have both masculine and feminine forms (*el/la*, *un/una*). All 58 tokens were categorically marked with a masculine determiner, regardless of the possibility of the Spanish equivalent being

---

[218] Torres Cacoullos and Aaron (2003) question if two-word fragments may also be examples of constituent insertion (Poplack 2004) but note that they all occur at equivalence sites unlike constituent insertion which has been reported to occur at non-equivalence sites. Additionally, constituent insertion has been found to occur at low frequencies and only with specific constructions (Poplack 2004; Poplack and Meechan 1995).

[219] There were no post-posed subjects, therefore subject position could not be used as a diagnostic.

feminine. Therefore, as with single English-origin nouns, the analogical criterion is not applied and the masculine default is used.

These fragments were also assessed for number agreement but only 12 tokens were eligible for plural agreement, as the remaining tokens were either singular with a singular determiner or zero marked for determination. All but five were found to show plural agreement mirroring the pattern of the single English-origin nouns which also evidenced the application of the Spanish agreement rule[220]. According to the diagnostics applied, these modified NPs for the most part behave in the aggregate more like borrowings despite their apparent internal consistency with English with regard to adjective-noun word order. Furthermore, these modified NPs suggest that when speakers do utilize adjectives to modify nouns in mixed speech, they switch the [ADJ + N] constituent to circumvent the conflict site of adjectival word order. English preposes adjectives and Spanish generally postposes them with exception of a limited number of Spanish nouns that have a qualifying function (*grandes logros* 'great achievements,' but not *roja casa* 'red house')[221]. There are no examples in the data where the canonical Spanish [N + ADJ] word order is violated. Instead what speakers seem to do to resolve the conflict between two grammars is to borrow the [ADJ + N] constituent. This is clear evidence of speaker sensitivity to sites of equivalence and conflict sites and, additionally, provides strong support for the equivalence constraint, which predicts switches at equivalence sites and prohibits them at conflict sites such as between an adjective and a noun for

---

[220] Of the remaining five that did not show agreement, four had a singular determiner with a plural noun (*un whole entire stories, un big issues, el two hundred dollars, el September celebrations*) The first two examples, from the same speaker, can be considered anomalies; they reflect neither Spanish or English grammar both grammars require number agreement between the indefinite article and noun. The last two examples do seem to show the non-application of an agreement rule and these may be code-switches.

[221] While adjectives do appear among the single non-native items, these largely appear in copular constructions for which Spanish and English share the same word order or they are postposed to the Spanish noun consistent with Spanish grammar (e.g. *ese día llevaba una blusa **strapless*** 'And that day I was wearing a strapless blouse' BSEK/11W.J/11W.J_12W.J_13W.J/01:39.6). When the adjective is in a predicate nominal construction (*ella es amigable*, 'she is friendly') or when an intensifier precedes an adjective (*ella es muy amigable*, "she is very friendly'), English and Spanish have no conflict in word order.

Spanish and English. An absence of switches at the adjective-noun boundary is predicted by the equivalence constraint but rather than simply an absence of switches at this boundary, what the data shows is a clear strategy to resolve the conflict of grammars. As noted earlier, switching before a conflict site rather than at the boundary where the grammars of the respective languages conflict has been reported in the literature. In Wolof-French for example, the boundary between nouns and definite articles represents a conflict site between these two languages and it was found that any switches involving an NP took place before the full NP and not NP internally, i.e., not after the determiner (Poplack and Meechan 1995:214). Similar data is reported for Igbo-English where modifiers are postnominal and switches are not found NP-internally but rather before the NP which is a site of equivalence between the two languages (Eze 1997). Switching before major constituents is not unusual and, in fact, as noted earlier switching before an NP is a preferred switch site (Eze 1997; Treffers-Daller 1991; Berk-Seligson 1986; Poplack 1980). However, as noted by Eze (1997), the Igbo-English data is not simply reflecting this preference. Prepositional phrases are also major constituents and switches are permissible between Igbo and English both before the PP as well as PP-internally and speakers do switch within the PP but not within the NP (Eze 1997:163). Thus, these speakers switch at permissible switch sites, whether these are before a major constituent or within one, and do not at conflict sites, just as the speakers in the current study do. Belizean Spanish speakers appear to borrow the [ADJ + N] constituent and integrate it into the recipient language as they do single English-origin nouns.

Number expressions

Number expressions constitute the next largest category and these include time expressions, dollar amounts, and [ordinal number + N] constructions. This last category is the largest, accounting for 48% (N=20) of the number expressions and these are essentially modified NPs. They were all

singular; therefore, only the diagnostic of number agreement could be used. Seventy percent appeared with a determiner that could be masculine or feminine and, as with the modified NPs and single English-origin nouns, the masculine determiner was categorically used. Thus, these are most likely borrowings.

Time expressions as exemplified in (178) were infrequent (N= 5) and appear to be code-switches since they are consistent with English and not Spanish. Spanish time expressions require a feminine plural determiner (with the exception of one *O'clock* which takes a feminine singular determiner, *la una*) and English uses the frame [# O'clock]. None of the time expressions use the Spanish frame and instead use English form as shown below

178. Yo   cuando vengo acá   ya   es   **eight** **thirty**, ya   va   dar
     I   when   come   here   already   is   eight   thirty   already go   strike
     nine.
     nine

     'When I come here it is already eight-thirty, going on nine'

                                                                    (BSEK/05N14/05N14 00:39:11.776).

The remaining tokens in this category are infrequent and remain ambiguous since the word order is the same in both Spanish and English and the diagnostics of number and gender agreement are not relevant. For example, in (179) Spanish and English express the year in the same fashion and while Spanish often prefaces the year with a definite determiner, it is not obligatory.

179. Yo   tuve   mi   first,   yo   me   casé   en   two   **thousand one**
     I   had   my   first   I   CL   married   in   two   thousand one

     y   tuve   mi   first   baby   en   **two**   **thousand**   **three**
     and   had   my   first   baby   in   two   thousand   three

'I had my first, I got married in two thousand one and I had my first baby in two thousand three.'

Discourse marker and other category

There were also a number of tokens that appeared with discourse markers of some sort (*right*, *like*, *so*). Some of these discourse markers appear at the beginning of the alternation (180), others at the end (181) and some accompany single words while others accompany two or more words (182). If it is the case that discourse markers are freely moveable in the syntax, then this means that they are not part of a true syntactic boundary that the speaker has to navigate. In this case, examples from the first category with *like*, as in (180), can be considered borrowings based on previous analyses of single English-origin nouns and verbs and modified NPs. This is also true for those examples that appear with the tag *right* and possibly even for the NPs such as *culture nights* which have been shown to most likely be borrowings in the analysis above.

180. | porque | mi | esposo | le | gusta | sembrar, | **like,** | **plants** | que |
| because | my | husband | CL | likes | plant.INF | like | plants | that |

| da | frutas | para | comer |
| give | fruits | for | eat.INF |

'Because my husband likes to plant, like plants that bear edible fruit.'
(BSEK/05N14/05N14/00:12:51.020)

181. | Allá | es | todo | **bike**, | **right** |
| there | is | all | bike | right |

'It's all bike there, right'
(BSEK/11W/J11WJ_12WJ_13WJ/54:56:40)

182. | cada | rato | hacen, | **like,** | **culture nights** | y | así. |
| every | moment | do3PL.PRES | like | culture nights | and | like.that |

231

'They are always doing, like, culture nights and like that'

<div align="right">(BSEK/35W/35W/01:05.1)</div>

Institutions

Those two-word fragments that referred to institutions were coded separately, because these tend to be marked differently in Spanish and English (Torres Cacoullos and Aaron 2003) as exemplified below for Spanish in (183) (*el colegio,* 'high school') and for English in (184) with *high school*.

183. es   difícil,   como   no   lo   agarré en   <u>el</u>   <u>colegio</u>
      is   hard   since   no   CL   took   in   DET   high.school

      entonces es   más   duro
      then   is   more   hard

      'It's hard, since I didn't take it in high school then it's harder'

<div align="right">(BSEK/38W/38W_1_2 combined/00:02:41)</div>

184. porque   en   **high**   **school** es   un   requirement   y
      because in   high   school is   a   requirement   and

      mi   esposo   es   un   Spanish   teacher
      my   husband   is   a   Spanish   teacher

      'Because in high school it is a requirement and my husband is a Spanish teacher.'

<div align="right">(BSEK/05N14/05N14/00:43:56)</div>

Each noun in the Spanish dataset that referred to an institution was extracted and coded for the presence or absence of a determiner. The same was done for the two-word English origin nouns denoting institutions. There were 20 nouns denoting institutions found in the Spanish dataset and these were categorically marked with an overt determiner while the 24 English-origin fragments such as *high school* were categorically zero-marked. This finding is different from what was reported for New Mexican Spanish where English origin nouns that denote institutions in Spanish

discourse were marked with overt determiners (Torres Cacoullos and Aaron 2003).[222] Belizean

Spanish speakers appear to treat the classes of nouns denoting institutions differently from New

Mexican Spanish bilinguals. Though the data is not substantial, this result highlights the need to

assess in detail the patterning of two-word items as well as the need to assess how speakers in

different bilingual communities treat non-native items in the recipient language in general.

The largest categories of ambiguous MW fragments were analyzed and while the

diagnostics employed could only be applied to those tokens that had overt determiners, most of

these show integration into Spanish suggesting that they are borrowings. The exceptions are time

expressions and those nouns which denote institutions. A small number of tokens in this category

remain ambiguous due to factors such as shared word order between the two languages which

impedes identification. Nonetheless, by utilizing conflict sites within a comparative framework, it

was possible to account for the vast majority of the data and disambiguate the status of these MW

fragments.

6.5 Community Patterns and Regional Differences

While the above discussion of Table 32 indicated the overall pattern for Belizean Spanish speakers

in the current sample, it is of interest to assess whether or not there are regional differences. There

is ample evidence that community norms play a crucial role in patterns of language mixing even

when the same language pair are involved (Poplack 2015, 2004, 1987; Muysken 2000; Nait

M'Barek and Sankoff 1988). Furthermore, Spanish as spoken in the north and in the west have

been identified and described as distinct varieties (Cardona Ramírez 2010; Hagerty 1979).

---

[222] Note also in the example above, the noun denoting an occupation, *Spanish teacher,* in a predicate nominal construction is marked with a determiner, also departing from the pattern reported for New Mexican Spanish (Torres Cacoullos and Aaron 2003) where such nouns were zero marked in Spanish discourse. However, such determiner use is not out of the ordinary for fluent Spanish-English bilinguals (Lipski 2008:60).

Consequently, patterns of mixing cannot be extrapolated from typology alone nor from the overall rates of code-switching previously discussed. In addition, the northern speakers in the current sample are from more urban centres having more contact with English and Belize Kriol as compared to the western speakers who are from a more rural area having less contact with these two languages. Notwithstanding the caveat that the interviews analyzed for the northern speakers and the western speakers are 5 and 7 respectively, potential regional differences are analyzed using normalized rates of borrowing and code-switching for each speaker, along with self-reported language dominance and bilingualism as defined in Chapter 3.

Rates of linguistic code-switching are summarized by region in Table 36 below.

Table 36: Distribution of types of linguistic code-switching by region

	REGION				TOTAL	
Type of linguistic code-switch	NORTH		WEST			
	N	%	N	%	N	%
Ambiguous multiword fragments	142	53	92	64	234	57
Intrasentential (unambiguous)	68	25	22	15	90	22
Intersentential	57	21	19	13	76	18
Extra-sentential	3	1	10	7	13	3
TOTAL	270		143		413	

For both communities the largest category is ambiguous MW fragments, though the western speakers (64%) have higher rates than the northern speakers (53%). The next largest category for both communities consists of Intrasentential switching (22% overall), though, the rates are higher among northern speakers at 25% than for the western speakers (15%). Intrasentential code-switching is the category that follows next and, as with the previous category, the communities differ in rates (21% in the north versus 13% in the west). Within each community, however, the rates of intrasentential as compared to intersentential switching are not highly differentiated, these are 25% and 21% for the north and 15% and 13% for the west respectively. Thus, speakers in each community engage in these two types of switching at virtually equal rates. The final type of

code-switching, extra-sentential or tag switching, is minimal at 1% for the northern speakers and 7% for the western speakers. Though, in the west, most of the tag-switching is accounted for by one speaker (8/13 tokens).

Overall, northern speakers appear to be higher code-switchers. They contribute almost double (N=270) the MW fragments that the western speakers produce (N=143) as seen in Table 36. Northern speakers also show higher rates of code-switching in all but one category of code-switching. Their intrasentential code-switches account for 76% (68/90) of all the intrasentential switches; similarly, northerners account for 75% (57/76) of intersentential switches and 60% (142/234) of the ambiguous MW fragments. These higher rates of code-switching and the fact that northern speakers account for most of the linguistic code-switching in the sample, suggests that northerners as a group may be higher language mixers in general and higher code-switchers, in particular. Differences between rates of use between the two groups achieves statistical significance.

The number of hours of recorded data is equal for both groups[223] and the differences in rates of use achieves statistical significance as indicated above for Table 36. This apparent disparity between the two groups could be a true reflection of varietal differences or community norms. Alternatively, these apparent regional differences may obscure the effect of such factors such as linguistic ability or contact with English (and Kriol), which, in the current study, overlaps with the rural/urban distinction. The northern region has historically been less isolated than the west; consequently, northern speakers have generally had more contact with English and Kriol than the communities in the West (LePage and Tabouret-Keller 1985; Hagerty 1979). In addition, bilingual ability is reported to correlate with different patterns of language mixing (Lipski 2008: 88,

---

[223] While there are seven speakers in the group of western speakers, data from three participants was obtained during a joint interview, therefore, the total interview time for each region is approximately equal.

1985:18; Poplack 2004,1980; Poplack et al. 1988; McClure 1977) and, as discussed in Chapter 3, the speakers analyzed here, reported differing levels of language dominance and bilingualism. Therefore, while all speakers are high language mixers, this does not mean that they will be high intrasentential code-switchers. As a consequence, a detailed analysis of the patterns and language use for individual speakers is carried out to reveal how these speakers are situated with respect to the patterns in each region for the current sample and to ascertain if factors such as bilingual ability may be responsible for the observed patterns of mixing (Wilson and Dumont 2014: 9-12; Poplack 1987: 52; Poplack 1980: 606-608).

The classification of speakers as bilingual, trilingual or low bilingual, previously presented in Chapter 3, is now considered together with both overall and normalized rates of borrowing and code-switching [224]. Table 37, below, shows overall rates by region for borrowings and intrasentential code-switching. The total number of intrasentential code-switches are taken directly from Table 36 above, while rates of borrowing are taken from the English-origin nominal and verbal items previously analyzed as borrowings in the preceding chapters.

Table 37: Rates of language mixing (borrowing and types of code-switching) by region[225]

	North N	West N	TOTAL N %
Intrasentential (unambiguous code-switching)	68	22	90 13
Total Borrowing Nouns + verbs	391	321	712 87
Total borrowing + code-switching	459	343	

---

[224] The speakers form a relatively homogenous group with respect to other extralinguistic factors. All speakers had completed high school or were currently in high school (in the case of the teenage speakers) and none had completed post-secondary education. All but one speaker is between the ages of 15 and 35, which corresponds to the group labeled "young" in Chapter 3. The remaining speaker is in the group "middle" (36-55 years old).

[225] The ambiguous multiword fragments analyzed above are not included here given that the majority appeared to be borrowings and those tokens that appeared to be code-switches were infrequent and require further analysis before including them with unambiguous code-switches.

Overall, both groups produce more borrowings than intrasentential code-switches but northerners produce more than double the code switches than their western counterparts despite a comparable number of borrowings. Thus, it appears that the western speakers may have a tendency to be "borrowers" rather than "code-switchers," the former mechanism not requiring high levels of bilingual proficiency while the latter does (Bullock and Toribio 2012).

In order to assess any correlations between region and bilingualism, and to help determine if the apparent regional differences reflects a true distinction between the two groups, a normalized rate over a twenty-minute period was calculated for each speaker for borrowings (nouns + BCVs) as well as for intrasentential code-switches. This was done by obtaining for each participant: (i) the total speech time from the annotation statistics in ELAN[226], (ii) the total number of borrowings and (iii) the total number of intrasentential code-switches. Normalized rates were then calculated as exemplified below for normalized rates of borrowing[227]:

$$\frac{\text{Total borrowing X 20 (minutes)}}{\text{Total speech time}} = \text{normalized rate}$$

By calculating rates of borrowing and code-switching over a time period that is kept constant for each speaker, it becomes possible to compare their language mixing behaviour to each other as well as according to bilingual ability and region. These data are presented in Table 38 along with self-reports of language dominance and bilingualism previously presented in Chapter 3. For

---

[226] Participant speech is transcribed into segments, called annotations, in ELAN. ELAN provides the total annotation time in the annotation statistics for each speaker which has been transcribed and this corresponds very closely to their actual speech time. The length of the interview does not enter into the calculation since some interviews were joint interviews.

[227] Normalized rates can also be calculated for a particular word count, per 1000 words, for example. I chose to follow a time based calculation, however, since speakers were originally classified as low, medium or high language mixers according to number of non-Spanish items over a twenty-minute period (see also Poplack (1981) for calculations of rates in terms of switches per minute).

convenience, I repeat here the description of how the distinctions in language dominance and bilingualism are operationalized for the current study. If a participant reported non-restricted use of two languages and language dominance in two languages, then they were regarded as bilingual, and as trilingual if they reported non-restricted use of all three languages (Spanish, English and Belize Kriol) and no one language of the three as dominant. If, on the other hand, a speaker reported restricted use in one or more languages and dominance in one language only, then this speaker was considered "low-bilingual" regardless of the language they reported as their dominant one.

As seen in Table 38, all speakers borrow at relatively robust rates with normalized rates for borrowing ranging from 26 (speaker 35) to 78 (speaker 29). While all speakers code-switch, the code-switching rates are more differentiated among speakers, with eight of the speakers code-switching only 3 times or less in twenty minutes. The remainder of the speaker code-switch between 7-15 times per twenty minutes (these rates are indicated in boldface). I will call the first group (N=8) *low code-switchers* and the second group (N=4) *high code-switchers.* Note that there were no speakers that switched between 3 and 7 times per twenty minutes[228].

Despite the overall lower rate of code-switching for western speakers in Table 37, the normalized rates for individual speakers in Table 38 show that while the majority of low code-switchers are found in the west, low code-switchers can also be found in the north. When bilingual ability and language dominance are examined, it appears that the low rates of code-switching are tied to bilingualism rather than region. Of the six low code-switchers from the west, four are also low bilinguals and they are all Spanish-dominant. Recall that low bilinguals are those speakers that have restricted use of two of the three languages and who operate primarily in one. Note that code-

---

[228] A two-tailed t test revealed the differences in normalized rates between low and high switchers to be significant (p= .00003).

switching rates are similar for all the low-bilinguals regardless of the language identified as their dominant one.

Table 38: Normalized rates of borrowing and intrasentential code-switching by region, language dominance, and bilingualism for individual speakers

Region	Participant	Dom Lang		TOTAL BORR	**BORR RATE/20 mins**	TOTAL CS	**CS RATE/ 20 mins**
N	**45J**	**B**	**Bilingual[229]**	**63**	43	**22**	**14**
N	**58**	**C**	**Trilingual**	**110**	65	**27**	**15**
N	**05**	**C**	**Trilingual**	**148**	72	**14**	**7**
N	60J	E	low bilingual	18	29	2	3
N	62	E	low bilingual	52	42	3	2
W	**08**	**T**	**Trilingual**	**58**	44	**12**	**9**
W	**35**	**T**	**Trilingual**	**46**	39	**1**	**< 1**
W	29	B	Bilingual	45	78	1	2
W	38	S	low bilingual	40	26	3	2
W	12J	S	low bilingual	18	39	2	2
W	11J	S	low bilingual	64	44	2	1
W	13J	S	low bilingual	50	**51**	1	1
				712		90	

KEY:
Dom Lang= Dominant language. B=both English and Spanish; E=English; S=Spanish; T= all three: Spanish, English and Belize Kriol.
BORR= borrowing; CS= code-switching; RATE= normalized rate.
Region: N= north, W= west
Participant: J=joint interview,

But what of the northern low code-switchers? Table 38 shows that the two low code-switchers from the north are also low bilinguals, though rather than being Spanish-dominant, they are English-dominant. Thus, 75% (6/8) of low code-switchers are also low-bilinguals while 100% of high code-

---

[229] This speaker reported little use of Kriol and indicated that they did not really care for it, yet their interview was one that revealed the most use of Kriol features. This speaker, for example, accounts for all the tokens of *bway*.

switchers are bilingual or trilingual. None of the high-code-switchers are low-bilinguals. This correlation is precisely what is reported in the literature; code-switching is considered to require proficiency in both languages in the language pair and therefore to be the domain of highly proficient bilinguals (Treffers-Daller 1994; Poplack et al. 1988; Poplack 1980; Berk-Seligson 1986; McClure 1977).

Most high code-switchers are from the northern region but they are also bilingual or trilingual. Thus, bilingual proficiency may be a necessary pre-requisite for code-switching but may not necessarily predict language mixing behaviour. Two speakers (35 and 29) in the west, one bilingual and one trilingual, do not pattern with the trilingual high code-switcher from the same region (speaker 08). It is possible that community norms of the west somehow militate against code-switching, however, with only three bi/trilingual speakers it is difficult to tell if community norms are the determining factor in their code-switching patterns. Another possibility is that mobility is a factor in the linguistic behaviour of speaker 08 who is more mobile than speakers 29 and 35 as he happens to travel regularly within Belize for work. As a consequence, speaker 08 has higher levels of contact with English and Kriol than other speakers and may even be less sensitive to community norms due to regular travel outside the community. Similar findings were reported for Puerto Rican Spanish-English speakers as those speakers who worked off the block had higher rates of intrasentential code-switching than those who did not work off the block (Poplack 1980: 613)[230]. More data from high language mixers would shed light on this issue, in particular from western speakers. I leave these question for further research when such data can be analyzed.

In summary, those speakers in the sample that are low-bilinguals are largely borrowers while those identified as bi/trilingual speakers both borrow and code-switch but are generally high code-switchers. This latter group is found primarily in the north. An important point here is that

---

[230] See also Poplack et al. (1988) on how borrowing patterns correlate with working outside the community.

whether the speaker was Spanish-dominant or English-dominant did not directly correlate with code-switching or borrowing but rather it appears that it is the level of bilingualism, i.e., knowledge of both languages, which is associated with code-switching. If one language in the language pair was highly restricted, then these speakers were unlikely to be high code-switchers and while not all tri/bilinguals were high code-switchers, only tri/bilingual speakers are found in the group of high code-switchers. While more data is needed to confirm the finding, it is keeping with the literature, thus highlighting the assertion that intrasentential code-switching requires knowledge of both languages (Bullock and Toribio 2012; Poplack 2004, 1980; Treffers-Daller 1994; Myers-Scotton 1993; Lipski 1985; McClure 1977). Such a finding also reinforces the view that both languages are operative when code-switching takes place while only one is operative when borrowing does. If they were not both operative with code-switching, for example, then speakers would not need knowledge of the two languages, and, alternatively, if both languages are active during the borrowing process then the same correlations between bilingual ability and code-switching should be evident between borrowing rates and bilingual ability, but this is not corroborated by the data. Furthermore, all twelve speakers behave uniformly with regard to the linguistic integration of single English-origin items as discussed in Chapter 4 for verbal items and for nominal items discussed in Chapter 5, despite the differences in bilingualism and language dominance which emerged from self-reports. Even where Spanish-English speakers might be hypothesized to show differential behavior, i.e., opting to employ the analogical criterion in assigning Spanish determiners to English-origin nouns, Belizean Spanish speakers showed uniform behaviour in gender assignment. Yet, speakers in the sample show differential behaviour, largely according to bilingualism, when it comes to linguistic code-switching, at least with regard to rates of use of intrasentential code-switching.

6.6 Discussion

This chapter was dedicated to analyzing the non-Spanish MW fragments in the data in order to determine community patterns of borrowing (discussed in previous chapters) and code-switching. In addition, switch sites in intrasentential code-switching were analyzed to determine if the data adheres to the predictions of the equivalence constraint and the free morpheme constraint and if any special mechanism, such as discourse markers, trigger code-switching. It was first necessary to identify the appropriate data for testing constraints, i.e., intrasentential code-switching. In so doing, it was possible to show the overall distribution of types of code-switching which indicated that the main pattern of language mixing for this cohort of high language mixers is smooth intrasentential code-switching analogous to other Spanish-English data (Poplack 1980). A large portion of the data made up of ambiguous MW fragments also emerged and these were analyzed to determine whether they were code-switches or borrowings. They appeared to be mostly borrowings, which parallels results found for two-word multiword fragments in New Mexican Spanish (Torres Cacoullos 2003: 320). Some of the ambiguous MW fragments, such as time expressions and nouns denoting institutions, were shown to be intrasentential code-switches. A small number of tokens remain ambiguous such as where ambiguous two-word English-origin items shared the same word order with Spanish and/or were not overly marked with a determiner, therefore, their status as code-switches or borrowings could not be confirmed.

These results were found by a systematic analyses of the data using linguistically principled definitions of borrowing and code-switching which were drawn from the literature. In addition, conflict sites between Spanish and English were used to determine if data was consistent with Spanish grammar and therefore integrated into Spanish, or, if the data remained consistent with English grammar and, therefore, represented a code-switch. Thus, by demarcating the object

of study, as was done in previous chapters, patterns of mixing present in the current data could be elucidated. At the same time, this detailed analysis revealed anomalous tokens in the data that were highly infrequent as well as a small number of tokens that could not be identified by the diagnostics used in this study. The outcome of the analyses is the identification of regular patterns, distinguished from those anomalous occurrences that may be speaker error or attributed to idiosyncratic use. Methodologically, it is evident that all English-origin material in the data cannot be assumed to be the same based on single or multiword status given that speakers appear to treat some multiword fragments as borrowings. Additionally, those items that are borrowings must be separated from those that are code-switches if grammatical constraints on code-switching are to be tested. Finally, the analysis revealed that English-origin two-word fragments may not be treated in the same in different communities (e.g. *high school* may be treated as a borrowing or a code-switch).

At this point, single-word items have been identified as borrowings and MW fragments have been identified as various types of linguistic code-switching and functional or discursive code-switching. One of the overarching objectives of the current study was to identify the English-origin material in the data. It has been shown that these consist of at least two different phenomena, i.e., code-switches and borrowings. Furthermore, the analysis supports the assertion that code-switches and borrowings are distinct phenomena. The distributional differences between intrasentential code-switches and borrowings are distinct. Multiword fragments appear at syntactic boundaries predicted to permit code-switching and are not found at boundaries predicted to prohibit code-switching. Additionally, their distribution across alternation boundaries is far more varied than single English-origin items as they occur at more than 19 switch sites which consistent with other data reported in the literature (Poplack and Dion 2011; Poplack 1980; Eze 1997; Pfaff 1979). In contrast, single English-origin items such as nouns tend to appear in the syntactic slots

243

allocated for nouns in structures such as determiner phrases in subject or objet position (Torres Cacoullos and Aaron 2003; Sankoff et al. 1990). Consequently, English-origin nouns are limited to the boundaries following a determiner, verb, or preposition only, while English-origin verbs appear overwhelmingly after *hacer* in a BCV construction. This distribution is further evidence of the distinct status that single English-origin items, or borrowings, hold as compared to code-switches (see also Poplack and Dion 2012). The characteristics of code-switching and borrowing can be summarized as follows:

Code-switching
- Restricted distribution across syntactic boundaries, tendency to be limited to sites of equivalence
- Non-native multiword fragments are internally consistent with the language of switch. Single-word switches are consistent with morphosyntax of language of switch.
- High level of bilingualism required
- Severely limited or no recurrence

Borrowing
- Distribution across the same syntactic positions as the corresponding monolingual categories
- Consistent with grammar of recipient language
- Lower level of bilingualism required.
- Range of recurrence and social embeddedness: nonce use, recurrent, widespread in the community, dictionary attestation.
- Range of phonological integration.

While single English-origin items have been identified in the literature as the most difficult to define (Poplack and Dion 2012; Winford 2009; Poplack and Meechan 1998; Poplack and Meechan 1995), in reality, the short multiword fragments (mostly modified NPs) appear to present a greater challenge since they exhibit features of both the donor language (word order for modified nouns) and the recipient language (number agreement, default masculine). This is quite different from the single-word items, which show clear indices of full integration into Spanish. It seems important to examine such data on a larger scale as it presents a better testing ground than single-word switches for the hypothesis that code-switches become borrowings as a result of frequent use

(Myers-Scotton 2002). While this seems an unlikely scenario given the preceding analyses that borrowing and code-switching are two separate processes, it is an avenue meriting further investigation if we are to fully understand the linguistic strategies bilingual speakers engage in. Such items suggest that there may be more linguistic variability in bilingual speech than has been suggested by variationist analyses which contrast borrowing and code-switching. As a consequence, two-word items present a challenge to the binary approach of the variationist method and compel a further re-examination of the model to ascertain how the model can best capture variability in bilingual discourse. If two-word items are truly borrowings, then perhaps the definitions of code-switching and borrowing need to be re-visited and expanded. If they are not borrowings, then the model needs to account for them and even consider that bilingual behaviour includes more than loanwords and code-switches. The categories in Muysken's (2000) typology, for example, despite the criticisms detailed in Chapter 2, would account for these two-word items under constituent insertion given that this category includes both single words and larger constituents. More recent expansions of his model with componential analysis also presents a non-binary approach and to some extent takes into account the overlap in Muysken's categories (Deuchar et al. 2007).

While these two-word items present a challenge to the variationist model, and even though the tools used in the current study could not fully account for them, these items were only revealed by adherence to the accountable methods of the variationist approach. These two-word items occurred ay high rates and may be unique to Belizean Spanish speakers, though, until other studies also analyze two-word items separately on a consistent basis it will not be known if this is the case. Furthermore, if this feature is unique to speakers in Belize, it may not be captured as such by the componential analyses cited above. By staying close to the *a priori* definitions of the

binary approach of the variationist method, this data emerged and now feeds back into the model elucidating both what the model can account for well and what it may not account for as well.

With regard to constraints on code-switching, as predicted by the equivalence constraint and the free morpheme constraint, intrasentential code-switches were found at sites of equivalence between English and Spanish and but not at sites where the grammars conflict. This was evidenced not only by the absence of switches at sites of negation, between clitic and verb or at adjective-noun boundaries where Spanish and English have different word order, but also in the fact that speakers appeared to avoid the adjective-noun boundary as a switch site by instead borrowing English [ADJ+ N] constituents. Finally, speakers were not found to use any discourse-pragmatic features to flag or otherwise introduce code-switches as has been suggested for other Spanish-English data (Aaron 2004). This is further evidence of the smooth, skilled switching that speakers engage in.

Despite the fact that Spanish and English give rise to numerous permissible switch sites and despite the fact that the same language pair is employed in the bilingual speech of the majority of the northern and western speakers, the patterns of language mixing in the north and west are not identical. This finding lends to support to other findings that community norms play an essential role in patterns of language mixing though more data needs to be analyzed from these two regions (Muysken 2000; Poplack 2004,1987). By assessing language mixing patterns according to self-reports of bilingual ability and using normalized rates, the analysis was also able to provide evidence in support of reported findings that code-switching is a discourse mode requiring bilingual ability (Bullock and Toribio 2012; Poplack 2004,1980). The low-bilinguals were all low code-switchers and all the high code-switchers were bilingual or trilingual. The global language mixing rate assigned to each of the 51 speakers in the corpus was useful for selecting the 12 interviews to

be analyzed here but this rate masked the details of the speakers' bilingual behaviour, thus, highlighting the need for and benefit of systematic quantitative analysis of other-language material.

In general, Spanish speakers in Belize language mix in similar ways as other bilingual communities, in particular, Spanish-English speakers, at least as far as the current sample goes. First, Belizean Spanish speakers in the sample analyzed are all high borrowers and borrow mostly nouns. Second, they also engage in smooth code-switching without flagging their switches, a pattern that is similar to Puerto Rican Spanish speakers (Poplack 1980) but different from French-English speakers (Poplack 1987) despite high bilingualism and typological similarity of French to Spanish and extended contact with English. Given that flagged switching, regardless of type of flag, is found to correlate with a level of social stigma attached to code-switching (Poplack 1987) or with the absence of code-switching as a community norm (Poplack et al. 1989), the pattern found in Belize suggests that code-switching is a discourse mode as in the case of Spanish-English PR speakers (Poplack 1980).

Spanish speakers in Belize do differ from speakers in other Spanish-English bilingual communities in that speakers in both regions appear to borrow a high proportion of verbs as indicated in Chapter 4, and they do so via BCV constructions. This is a pattern that is either unreported for other Spanish-English data (Poplack 1980), reported to be a minor pattern (Jenkins 2003; Pfaff 1979) or reported at lower rates than in the current data (Wilson and Dumont 2014). In addition, there seem to be some differences in how these speakers treat two-word English-origin fragments in Spanish discourse and this merits further exploration. Finally, these speakers also evidence the use of Kriol in their speech and though the Kriol component could not be subjected to quantitative analysis, its role in language mixing in Belize also warrants further research and this study sets the stage for such research.

CHAPTER 7 SUMMARY AND CONCLUSIONS

The current study set out to ascertain both overall and regional patterns of language mixing in the discourse speakers of Belizean varieties of Spanish. In order to uncover these patterns, the second objective was to establish the status of all English-origin material, which meant analyzing both single-word and multiword fragments. English-origin nouns and verbs were analyzed with special attention paid to the status of bilingual compound verbs (BCV) as these were found to constitute a regular pattern in the language mixing data and have been understudied in the Spanish-English literature. Additionally, intrasentential code-switches were examined for grammaticality as per the equivalence constraint and the free morpheme constraint (Poplack 1980). These objectives are now discussed in light of the results from preceding chapters.

With regard to the first objective, the global patterns of language mixing were found to consist of smooth or skilled intrasentential code-switching and robust borrowing of nouns and verbs. The skilled code-switching used by the cohort of speakers analyzed is consonant with other reports for Spanish-English bilinguals for whom code-switching is a "discourse mode" (Torres Cacoullos and Travis 2014; Poplack 1980). Given that these switches were not introduced by any type of flagging, the proposal that discourse markers are pivots for code-switches (Aaron 2004) did not find support in the current data.

Northerners showed higher rates of code-switching according to overall rates. This was mirrored in the analysis of normalized rates which showed that northerners tended to be higher code-switchers than the westerners, though both communities were high borrowers with relatively robust rates. When bilingual ability and language dominance were examined for individual speakers, however, the differences in code-switching seemed to be tied to the differing (self-reported) linguistic abilities among speakers. All the northern speakers who were identified as bilingual and trilingual, for example, were also the "high-code-switchers." In the west, however, only

one of three speakers identified as bilingual or trilingual engaged in high rates of code-switching. None of the Spanish-dominant or English-dominant speakers engage in high rates of code-switching. Despite the small sample size, this finding corresponds to the assertion prevalent in the literature that given the right circumstances, such as the community's tacit acceptance of code-switching as a discourse mode, those speakers who engage in code-switching will be those who are proficient in their respective languages (Poplack 2004). Thus, the current study provides corroboration that it is the interplay of bilingual ability, typological similarity between the languages in contact, and community norms that permit or foster this linguistic behavior, and are consequently responsible for the regular patterns of mixing that emerge within communities rather than one factor alone (Muysken 2000; Poplack 1987:71-72, Poplack 1980:612).

Another difference found between northern and western speakers is the use of Belize Kriol. The use of uncontroversially Kriol grammatical features (future temporal marker: *a*, plural marker: *dem*, negation expressed by preverbal *no*, completive *don*, etc.) could only be observed in the speech of northern speakers. In contrast, in the west, only a few isolated instances of possible Kriol use were found. These consisted of possible Kriol reduplication (*healthy, healthy*) and the use of the lexical item *gyal*, 'girl,' both by the same speaker (13WJ). Northerners were also found to use BCVs more than westerners, corroborating previous findings on BCVs in northern and western Belize (Fuller Medina 2005a,b) and pointing to higher levels of contact with English and concomitant bilingualism as a possible source. Recall that speakers of Northern Belizean Spanish have historically had more contact with English and Kriol, a factor which may have also facilitated more BCV use. Though further analysis is required with more speakers and less overlap of social factors. At this time, both communities appear to have the same linguistic patterns of language mixing, with regional differences limited to differential rates of use of intrasentential code-switching, BCV use, and use of Kriol in mixed discourse. Whether the differential rates of use are a true

reflection of differences between the two regions would need to be further assessed with a larger sample.

The second objective was to identify the status of the English-origin material in Spanish discourse. To this end, single and multiword English-origin fragments were analyzed. Single English-origin items were subjected to quantitative comparative analysis by making use of conflict sites between Spanish and English. The diagnostics of gender and number agreement as well as subject position were used. The results show that Spanish number agreement rules were consistently applied to the English-origin nouns and that these nouns also occurred postposed to the verb at rates parallel to the unmixed Spanish. These patterns are consistent with Spanish grammar but not English. In addition, the masculine default was applied at near categorial rates to English-origin nouns, which was taken as evidence of integration into Spanish based on the use of the masculine default in Spanish (Harris 1991). These results are in contrast to other data where various competing factors come into play in gender assignment, including the analogical criterion and phonological form (Dubord 2004; Smead 2000; Poplack et al. 1982) suggesting, once again, that the linguistic mechanisms employed by bilinguals are determined by the community norms (Poplack 2015:923; Poplack et al. 1982). This is the case for both borrowing and code-switching as previously noted in the discussion of the literature in Chapter 2.

In terms of English-origin verbs, BCVs constituted the major pattern and were found to be highly productive irrespective of region. By applying the diagnostics of variable clitic placement and verbal morphology within the comparative variationist framework, it was found that the English-origin items in BCVs were integrated into Spanish. Therefore, BCVs constitute a borrowing strategy. In fact, this construction turned out to be the preferred means for adapting English-origin verbs among the speakers in the cohort analyzed, which appears to be a distinctive feature of Spanish in Belize. The analysis effectively clears up the question concerning the status of BCVs by

indicating that, as borrowings, they are neither code-switches, evidence of a third grammar (Romaine 1986), or creolization (Balam et al. 2014; Gardner-Chloros 2009; Pfaff 1979).

The categorial status of the English-origin items was also confirmed. This was done by the systematic application of a series of discourse-based tests to determine the categorial status of the English-origin items. The English-origin items can be nominal or verbal, but the majority consisted of English-origin verbs. While not the first study to employ discourse-based tests to gain insight into the nature of the complement of *do*-verbs in BCV constructions (Jenkins 2003; Muysken 2000), this study is unique in that these tests were applied in a systematic fashion to all potential BCVs in order to define the sample for analysis. In so doing, the methodological tools for circumscribing BCVs in the current study are made transparent and therefore replicable.

Finally, the question of aspectual constraints on BCVs was also addressed by making use of the conceptual and methodological tools of the variationist method. One way of independently operationalizing lexical aspect was proposed, and the distributional analysis of BCVs revealed that stative verbs, as defined in this study, did not appear frequently in BCVs. Those statives that did occur in BCVs appear to belong to a subset of more "dynamic" statives. Thus, *hacer* seems to bear an aspectual sensitivity potentially limiting the class of verbs which may appear as its complement in a BCV construction. Such a finding supports the assertions that *hacer* may be grammaticalizing or grammaticalized given that statives are generally the last verbs to participate in constructions with verbs that are grammaticalizing. At the same time, because these conclusions are based on synchronic data, the analysis would benefit from comparative analyses with monolingual data as well as data from older speakers. An apparent-time analysis, utilizing data from older speakers may elucidate whether the use of BCVs has undergone change over time with respect to lexical aspect. Further examination from monolingual data could be used to determine the overall occurrence of stative verbs in the corpus as statives may simply be infrequent overall. In addition to

overall rates, the distribution of types of stative verbs would also need to be examined in the monolingual data given that statives appear to be on continuum of more or less dynamic states. If a variety of stative verbs occur in the monolingual data but only more dynamic statives occur in BCVs then this would bolster the analysis of *hacer* bearing an aspectual sensitivity to its verbal complement. If the reverse is true, and there is no difference in the types of stative verbs in the monolingual data and BCVs, then this would suggest that all verb classes can occur in BCVs. Expanding the current definition of statives would also be another avenue to pursue. The criteria set for stative verbs in the current study was relatively conservative, likely excluding some classes of verbs that should be considered, and more importantly, the criteria did not address any potential differences between Spanish, English, and Kriol with respect to lexical aspect.

With regard to multiword fragments, these consisted of various types of code-switches (intersentential and intrasentential) and a large category of ambiguous multiword fragments. As noted above, the intrasentential code-switching for these speakers was determined to be smooth and unencumbered by functional flagging. In Chapter 2 it was outlined that since the EC and FMC were first proposed, methodological tools had been developed for analyzing bilingual data, and based on subsequent studies, the definitions of borrowing and code-switching had been refined. Taking these advances into account, the current data, based on previously undescribed varieties, was assessed to determine if the EC and FMC were upheld and, in fact, the intrasentential code-switches produced by the speakers in the current study adhered to the predictions of the EC and FMC. The switches all occurred at sites of equivalence. Further support for these constraints was found in the apparent avoidance of switching at ADJ-N boundaries which is a conflict site for Spanish and English for most attributive adjectives.

The results have shown that single words are borrowings in their majority and multiword fragments tend to be code-switches, with the exception of two-word ADJ+N sequences. At the

252

same time, the analysis also revealed that single words can indeed be code-switches and can be identified as such in the data (e.g. **signed** *al label* 'signed to the label'), though they are rare in the current data as is the case in other bilingual corpora (Poplack 2015). On the other hand, while multiword fragments are generally code-switches, some multiword fragments are also present as borrowings, as in the case of the many ADJ+N combinations that speakers tend to treat as a single unit. The ambiguous multiword fragments, when subjected to the diagnostics of gender and number agreement, showed indices of integration in a similar fashion to single English-origin items; however, more research is needed to develop diagnostics for assessing such items, particularly since two-word items are not treated in a standard fashion in the literature and may not be as frequent as in the current corpora. These findings underscore the importance of assessing all data according to well-motivated criteria for code-switching and borrowing. It also suggests that borrowing may be a more sophisticated behaviour than the integration of a single non-native item. Despite the major tendencies in bilingual data for single items to constitute lexical borrowings and multiword fragments to be code-switches, this classification is not categorical and each community may produce unique configurations of types of code-switching and borrowing, or even other phenomena that fit neither category of code-switching or borrowing.

By employing the conceptual and methodological tools of the comparative variationist framework in the analyses in the preceding chapters, this study responds to unresolved questions regarding code-switching and borrowing from a strong empirical base. The questions include: (i) whether or not these two phenomena are distinct, (ii) whether they can be reliably identified or distinguished, (iii) and, consequently, which linguistic phenomena should be considered in each category, and finally, (iv) whether it matters for a theory of code-switching (Poplack 2015). The data analysis revealed an answer in the affirmative to these questions, though a portion of the data requires further study (ambiguous multiword fragments). By exploiting the conflict sites, the single

English-origin items were shown to be integrated into Spanish, and these differed markedly from those English-origin multiword fragments which were internally consistent with English (and in some cases Kriol). Further supporting the distinction is the fact that code-switches and borrowings are distributed differently across syntactic boundaries; code-switches occur in much more varied positions than do borrowings (Poplack 2015; Poplack and Dion 2012). Consequently, borrowing and code-switching do not appear to lie on a continuum as proposed by some authors (Myers-Scotton 2002; Thomason 2001:135-136), though bilingual behaviour may consist of more than these two categories. In addition, the fact that proficiency in both languages used in mixed discourse was correlated with code-switching, but not borrowing, challenges the assertion that both the respective languages in mixed discourse are activated for both borrowing and code-switching (Myers Scotton 2002). Many of these findings, repeated in numerous other studies[231], have direct implications for modelling code-switching and therefore confirm the relevance of recognizing a distinction between code-switching and borrowing. The majority of the data in bilingual discourse consists primarily of lexical borrowings (Poplack 2015), for example, and because this is a distinct process from code-switching, one of adaptation rather than alternation between grammatical systems, a theory of code-switching would not adequately account for the majority of the data that constitutes bilingual discourse.

In addition to responding to questions in the literature on language mixing, the data and results furnished by the current study contribute to theoretical approaches regarding light verbs and the analogical criterion. BCVs, for example, have served as data in theoretical analyses regarding the nature of little *v* and light verbs (Rao and Den Dikken 2014; Fuller Medina 2013, 2010; González Vilbazo and López 2012, 2011); thus, the current study makes available additional data for these discussions. For instance, the distribution of stative English-origin verbs in BCVs, as

---

[231] See Poplack (2015, 2004) for an overview of the literature on code-switching.

defined in this study, pointed to the possibility that only a restricted set of stative verbs may be borrowed via BCVs which would provide evidence that *hacer* retains some lexical meaning and may not be compatible with all Vs (in the sense of Folli and Harley 2007; Cuervo 2003). These results were also taken as possible synchronic evidence of an earlier diachronic process of grammaticalization also consistent with the literature on light verbs (Butt 2003). Consequently, while the use of the construction may be innovative on the part of bilinguals, it is unlikely a "new structure" which is exactly what the theoretical analyses of the internal structure show.

In terms of gender assignment on English-origin nouns, where gender appears to be assigned based on analogy with Spanish equivalents of English-origin nouns, the analysis is that a Spanish gender feature of the corresponding Spanish equivalent is assigned to the English-origin noun, which normally carries no gender feature, in order for agreement to take place (Liceras et al. 2008). The near categorical use of the masculine default in the current data suggests that no such feature assignment is taking place when English-origin nouns are used by the speakers in Spanish discourse. Thus, an alternative account of feature checking must be available (see Liceras et al. 2008). Although the variationist method holds the knowledge of variable use of linguistic structures in the speech community as the central aspect of competence and aims to tap this through naturalistic data, the results presented here provide clues as to other aspects of speaker competence which are more in keeping with generativist approaches and that merit further exploration. While the earlier discussion in Chapter 2 suggested that judgment tasks are limited in capturing the constraints on intrasentential code-switching, English-origin nouns which appear with Spanish determiners (or mixed DPs) may be a site where such tasks can be instructive. Spontaneous data from the various studies have shown that bilingual behaviour is correlated with bilingual ability. The experimental data from Liceras et al. (2008: 847-849) also show such correlations where Spanish L1 speakers preferred mixed DPs that met the analogical criterion and

English L1 speakers preferred the masculine form of the Spanish determiner to mark English-origin nouns. The spontaneous data from adult simultaneous bilinguals, which they analyze, suggests that these speakers also prefer the masculine default. Together, these results suggest that the representations of gender may be different for each group or, alternatively, that the tasks place different cognitive demands on speakers and lead to disparate results. As they suggest, both spontaneous data and experimental data from the same speakers would make it possible to address this issue. It would be a relatively straightforward task to employ judgment tasks for mixed DPs with the speakers whose discourse have been analyzed in this thesis which would then provide data on different aspects of competence from the same speakers. Recall that I argued that borrowing behaviour appears to be below the level of consciousness, which may mean that it is less stigmatized than the use of multiword fragments, consequently, the limitations that factors such as social stigma place on assessing intrasentential code-switching may not be as problematic in the case of English-origin items in Spanish discourse.

While the focus here has been on language mixing, some aspects of monolingual Spanish spoken in Belize (e.g. gender agreement rules which apply in monolingual portions of the data) were also uncovered. Previously reported non-standard features were confirmed in the data, though a quantitative analysis is needed. Some of these features, such as the possessive pronoun use for inalienable possession, might appear to be due to contact with English; however, since such features are documented in Old Spanish, these could likely be retentions from older varieties of Spanish. The question remains open as to whether or not, or to what extent, these non-standard features are due to contact effects. Some morphosyntactic differences between the north and west were also noted where previously phonological differences had been the main differentiating factor. The use of the indefinite article together with the possessive pronoun (*un mi tío* 'my uncle'), for example, was only found in the west.

Furthermore, some important points about Spanish in Belize were noted as early as the data collection and corpus building stage. Interviews and corpus compilation revealed that a number of interviews were highly monolingual and, furthermore, these were not all restricted to older speakers who might be expected to be more monolingual. Consequently, despite local perceptions of Spanish in Belize as being mixed overall (Fuller Medina 2005) and any similar suggestion, inadvertent or otherwise, in previous descriptions of the language (Fuller Medina 2005a,b; Hagerty 1996), it is, in fact, imprecise to describe the Spanish in Belize as comprised of mixed varieties (Fuller Medina 2015). The mixed discourse of bilingual speakers has tended to draw more attention (Fuller Medina 2015, 2014; 2005a,b; Balam et al 2014; Balam 2013; Hagerty 1996) than monolingual Spanish (Cardona Ramírez 2010; Quilis 1990; Hagerty 1979), but language mixing is but one feature of the varieties of Spanish spoken in Belize. Moreover, it became clear that it is a specific cohort of speakers who engage in robust language mix and that these are generally younger speakers. Thus, the focus on language mixing in this thesis is by no means intended as a description of varieties of Spanish in Belize as mixed varieties and should not be read in this way. The patterns described here may extend to a larger group of high language mixers, but not necessarily to all Belizean Spanish speakers**.**

At the same time, in addition to informing our understanding of the linguistic nature of bilingual behaviors, the study of bilingual data has the potential to bring insight to each language on its own at different levels (Myers-Scotton 2009:326). There is much to learn about monolingual language use from bilingual data. Perhaps because code-switching was initially seen as random and indicative of poor linguistic ability (Weinreich 1953/1968 as cited in Poplack 2004:598), there has been an emphasis on highlighting it as skilled behaviour as well as highlighting the correlation to proficient bilingual ability. Borrowing, on the other hand, has been framed, as has been done in this thesis, as language mixing behavior which requires less bilingual ability than code-switching

257

(Bullock and Toribio 2012; Poplack 1980). Additionally, borrowing is often viewed as a strategy to compensate for a failure to recall monolingual items, nonetheless, it can also be viewed as a clear indicator of linguistic competence. The consistency in using all morphological forms in the verbal paradigm appropriately in the conjugation of *hacer* in BCVs as shown in Chapter 4, for example, shows that speakers have a full understanding of the Spanish system of verbal morphology. Likewise, nominal integration shows comprehension of number agreement rules and how to apply them, as well as the knowledge that in order to integrate English-origin nouns with respect to gender, it is possible to appeal to the default masculine.[232] Speakers must have knowledge of both what is permissible in the recipient grammar and the right application of the rules in order to integrate non-native items; consequently, borrowing is a skilled form of language mixing. Furthermore, because areas of grammar, such as the use of the default masculine with nouns not marked for gender and the prohibition against bare stems in Spanish, are not explicitly taught, coupled with the limited formal instruction in Spanish that these speakers have, the results point to borrowing as a process that is below the level of consciousness and not one that can be acquired through instruction.

As a consequence, this study also has direct relevance for Belize because local varieties of Spanish are still highly stigmatized (Fuller Medina 2005a) and Spanish continues to be taught as a foreign language despite the native language status for many speakers and the levels of competence evinced by the current data. The data and results presented in this thesis can inform pedagogical practices if educators are open to viewing borrowing as an indicator of competence in Spanish and to viewing the Spanish spoken in Belize as local varieties and not "bad Spanish," "kitchen Spanish," or "Spanglish" (a newer term being used in Belize). The empirical account of the

---

[232] Potentially speakers could leave the nouns completely unintegrated by leaving them determinerless and instead inserting repair phenomena as Finnish-English speakers do (Poplack et al. 1989) or bracketing or flagging the nouns (e.g. with *as they say in English*) as did the Ottawa-Hull French speakers.

systematicity of how speakers use the resources of the languages available to them (without violating the respective grammars) may help to mitigate stigma since it brings to light the discrepancies between perceptions of language use and actual language use. Therefore, to refer to Spanish in contact situations only by the salient feature of language mixing is in essence a disservice to speakers as it reinforces stigmatization of their Spanish (mixed or unmixed) as well as the notion that they are not native speakers and do not have command of the grammatical systems of Spanish (Otheguy and Stern 2011). This takes on additional importance in Belize where the lingua franca, Kriol, is also stigmatized, and many speakers do not have full access to English. The insight offered by Otheguy and Stern (2011:86) that Spanish speakers "would be better served in their efforts to learn formal varieties of Spanish, as well as English, by recognizing that they are already speakers of Spanish," while made with reference to the U.S. context, is highly pertinent to Belize.

This thesis, and the corpus on which it is based, also makes a contribution to the commitment outlined by the Government of Belize to "Identify, research and digitally record language patterns within Belize and to examine models which allow for the use, retrieval and preservation of Belize's languages" (NICH 2014:24). While Spanish may not be considered a minority language or an endangered one, it is currently unclear if the local varieties are being eroded in some communities and how the influx of newer Central American varieties may impact maintenance or shift in local varieties. As noted earlier (fn 85 Chapter 3), in at least one community, speakers under the age of 60 who met the criteria of Belizean Spanish speakers could not be found. This raises concern that the local variety may be undergoing displacement and is an area where further research is needed. Coupled with the growing recognition of *varieties* of languages as endangered (Michnowicz 2011 on Yucatecan Spanish; Travis and Torres Cacoullos 2013 on

New Mexican Spanish; Riegelhaupt et al. 2003), documenting Belizean varieties of Spanish may take on additional importance for the goal of language documentation and preservation.

It has been noted several times that the analyses presented here could benefit from more data. For example, a larger stratified sample would be needed to determine the role, if any, of the extralinguistic factors of age, level of education, and social class, as these were left unexplored due to the relatively homogenous nature of the cohort with respect to these factors. The same can be said for regional differences: the north-west opposition overlapped with a rural/urban divide and differing degrees of contact with English and Kriol in the current study. Teasing out these factors is a necessary endeavour to fully understand any differences between the north and west, and whether such differences may be attributable to community norms, dialectal difference, contact with English, and/or rural/urban distinctions.

In addition, though the results showed robust rates for borrowing as well as a pattern of equivalence based on intrasentential code-switching, it would be important to assess these phenomena in a larger sample of high language mixers to confirm if patterns of mixing reported here extend at the community level. Likewise, notwithstanding the finding in the current data that code-switching and bilingual ability appear correlated (a finding widely reported in the literature), the cohort of speakers analyzed was relatively small, and it would be of relevance to assess such correlations with more speakers to determine if such correlations achieve statistical significance.

Finally, more Kriol data is needed to facilitate quantitative analysis and to uncover its role, if any, in the interplay of language mixing employed by speakers as well as any contact effects. To facilitate such an analysis, diagnostics must be developed that target the underlying grammar of Kriol and which do not rely solely on overt forms. Some Kriol forms, for example, may be indistinguishable from English in their surface manifestations but might be constrained by Kriol grammar in ways that differ from English; bare verbs being a case in point, as they are found in

both English and Kriol but their use in discourse is patterned differently (Poplack and Tagliamonte 2001:5). Recall that the complement of *hacer* in BCVs was deemed *English-origin* for this very reason; they were considered neither English nor Kriol in the absence of the appropriate analysis to confirm system membership, and were simply considered to have an English etymological source. Given the appeal of analyzing bilingual forms as instances of creolization, in particular in the case of Belize where Spanish is in contact with Kriol, it is crucial for clearly formulated hypotheses, diagnostics, and analysis to be developed to ascertain the role of Kriol in contact with Spanish.

In summary, the current study was able to draw clear conclusions regarding some aspects of language mixing in Belize as detailed above (e.g. the status of single English-origin items) and to draw tentative ones with regard to others (e.g. two-word fragments and aspectual constraints on BCVs). The analysis also brought to light areas where conclusions could not be drawn, either because the data and diagnostics were not available or because new lines of inquiry were generated (e.g. the role of language dominance in nominal gender assignment). All the above findings were revealed by employing the tools of the variationist comparative method which dictates clearly identifying the object of study at each stage of the analysis, employing appropriate benchmarks within the application of the comparative framework, and carrying out systematic quantitative analysis of the data. These tools were applied to areas that had not been studied in detail in Spanish-English data such as BCVs and the aspectual nature of the English-origin items in BCVs. Consequently, this thesis makes methodological contributions for circumscribing BCVs, operationalizing lexical aspect, and defining stative verbs. In addition, by analyzing data from both dialect regions of Belize, this study shows that speakers demonstrate homogeneity in choice of borrowing mechanisms (e.g. BCVs for verbs, masculine default gender assignment for nouns) and type of code-switching (smooth and skilled) but also evidence some potential disparity in regional

rates of use of Kriol, BCVs, and code-switching. While these findings are in need of confirmation from a larger sample, these insights would not have been gained without analyzing data from both regions.

At the start of the thesis it was noted that the "meager documentation available on Belizean Spanish demonstrates a void in Hispanic Dialectology which must be filled" (Hagerty 1979:2). While this void remains in large part, it has begun to be filled bit by bit in the decades since Hagerty first made that observation (Fuller Medina 2015, 2013, 2010, 2005a,b; Balam et al. 2014; Balam 2013; Cardona Ramírez 2010; Hagerty 1996; Quilis 1990). It is the hope that with the current thesis, a major step has been taken by bringing to light details of the bilingual discourse of those speakers who engage in language mixing as well as insights into the features and underlying grammatical systems of Northern and Western Belizean Spanish. In so doing, the contributions that the study of language in Belize can make to our understanding of bilingualism and language in general are highlighted and the stage is set for further research on Spanish and Spanish in contact in Belize.

REFERENCES CITED

Aaron, J.E. (2014). Lone English-origin nouns in Spanish: The precedence of community norms. *International Journal of Bilingualism*, Advance online publication.

Aaron, J. E. (2006). *Variation and change in Spanish future temporal expression: Rates, constraints, and grammaticization*. Unpublished Doctoral Dissertation. University of New Mexico.

Aaron, J.E. (2004). So respetamos un tradición del uno al otro. *Spanish in Context, 1*(2), 161-179.

Alonso Ramos, M. (2004). Las colocaciones con verbo de apoyo. *Madrid: VisorLibros*.

Amastae, J., & Elías-Olivares, L. (Eds.). (1982). *Spanish in the United States: Sociolinguistic aspects*. New York: Cambridge University Press.

Arad, M. (2003). Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language & Linguistic Theory, 21*(4), 737-778.

Backus, A. (1992). *Patterns of language mixing: A study in Turkish-Dutch bilingualism*. Wiesbaden: Otto Harrassowitz Verlag.

Bailey, G., Wikle, T. Tillery, J. & Sand, L. (1991). The apparent time construct. *Language Variation and Change, 3* (3), 241-264.

Balam, O., de Prada Pérez, A. & Mayans, D. (2014). A congruence approach to the study of bilingual compound verbs in Northern Belize contact Spanish. *Spanish in Context, 11*, 243–265.

Balam, O. (2013). Variable neutralization of the intervocalic rhotic contrast in northern Belizean Spanish. *Borealis: An International Journal of Hispanic Linguistics, 2*(2), 285-315.

Bandi-Rao, S., & den Dikken, M. (2014). Light switches: On *v* as a pivot in codeswitching, and the nature of the ban on word-internal switches. In MacSwan, J. (Ed.), *Grammatical theory and bilingual codeswitching*, (pp. 161-183). New York: Routledge.

Bhatia, T. K., & W. C. Ritchie. (1996). Light verbs in code -switched utterances: Derivational economy in I-language or incongruence in production? In Stringfellow, D., Cahana-Amitay, D., Hughes, E. &

Zukowski, A. (Eds.), *Proceedings of the twentieth annual Boston University Conference on Language Development*, (Vol. 1), (pp. 52–62). Somerville, MA: Cascadilla Press.

Becker, K. (2013) The sociolinguistic interview. In Mallinson, C., Childs, B., & Van Herk, G. (Eds.). *Data collection in sociolinguistics: Methods and applications*. (pp 91-100) New York: Routledge, Taylor & Francis Group.

Belazi, H.M., Rubin, E.J. & Toribio, A.J. (1994). Code switching and X-bar theory: The functional head constraint. *Linguistic Inquiry*, *25*(02), 221-237.

Bentahila, A., & Davies, E. E. (1995). Patterns of code-switching and patterns of language contact. *Lingua*, *96*(2), 75-93.

Bentahila, A. & Davies, E. E. (1991). Constraints on code-switching: A look beyond grammar. In *Papers for the symposium on code-switching in bilingual studies: Theory, significance and perspectives*. (pp. 369–404). Strasbourg: European Science Foundation.

Berk-Seligson, S. (1986). Linguistic constraints on intrasentential code-switching: A study of Spanish/Hebrew bilingualism. *Language in Society*, *15*(03), 313-348.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of spoken and written English* (Vol. 2). Massachusetts: MIT Press.

Bickerton, D. (1973). On the nature of a creole continuum. *Language 49*, 640-669.

Bills, G.D., & Vigil, N.A. (2008). *The Spanish language of New Mexico and Southern Colorado: A linguistic atlas*. Albuquerque: UNM Press.

Bokamba, E. G. (1987). Are there syntactic constraints on code-switching? In Denning, K. M., Inkelas. S., McNair-Knox, S. C. & Rickford, J. R. (Eds.), *Variation in language: Proceedings of the fifteenth annual conference on New Ways of Analyzing Variation* (pp. 35-51). Stanford: Department of Linguistics, Stanford University.

Bonner, D. M. (2001). Garifuna children's language shame: Ethnic stereotypes, national affiliation, and transnational immigration as factors in language choice in southern Belize. *Language in Society, 30*(01), 81-96.

Borer, H. (1984). *Parametric syntax*. Dordrecht: Foris Publ.

Bosque, I. (2001). On the weight of light predicates. In Herschensohn, J., Mallén, E., & Zagona, K. (Eds.), *Features and interfaces in Romance: Essay's in honor of Heles Contreras.* (pp. 23-38). Amsterdam; Philadelphia: John Benjamins.

Bruhn de Garavito, J., Heap, D., & Lamarche, J. (2002). French and Spanish se: Underspecified, non-reflexive. In Burelle, I.S. & Somesfalean, S. (Eds.), *Actes de l'ACL 2002/ 2002 CLA Proceedings* (pp. 42-54).

Budzhak-Jones, S. (1998a). Against word-internal codeswitching: Evidence from Ukrainian-English bilingualism. *International Journal of Bilingualism*, *2*(2), 161-182.

Budzhak-Jones, S. (1998b). Developing diagnostics: Word-internal code-switching versus borrowing. *Cahiers Linguistiques d'Ottawa*, 1-14.

Budzhak-Jones, S. (1998c). *Single-word incorporations in Ukrainian-English bilingual discourse: Little things mean a lot*. Unpublished Doctoral Dissertation. University of Ottawa (Canada).

Budzhak-Jones, S. (1996). Social conditions and code switching patterns. In Pasquini, A., Roswell,L. & Smith, L.C. (Eds.), *CLA annual proceedings/actes de l'ACL congres annuel* (pp. 23-34). Calgary: Calgary Working Papers in Linguistics, University of Calgary.

Bullock, B.E & Toribio, A.J. (2006). Intra-system variability and change in nominal and verbal morphology. In Gess, R., & Arteaga, E. (Eds.), *Historical romance linguistics.* (pp. 305-325). Amsterdam/Philadelphia: John Benjamins.

Butt, M. (2010). The light verb jungle: Still hacking away. In Amberber, M., Baker, B., & Harvey, M. (Eds.), *Complex predicates in cross-linguistic perspective*. (pp. 48-78). Cambridge: Cambridge University Press.

Butt, M. (2003). The light verb jungle. *Harvard Working Papers in Linguistics, 9*, 1–49.

Butt, M., & Lahiri, A. (2013). Diachronic pertinacity of light verbs. *Lingua, 135*, 7-29.

Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, J., & Thompson, S. (1997). Three frequency effects in syntax. In Juge, M.L. & Moxley, J.L. (Eds.), *Proceedings of the twenty-third annual meeting of the Berkeley Linguistics Society: General Session and parasession on pragmatics and grammatical structure* (pp. 378-388). Ann Arbor, Michigan: Sheridan Books.

Callahan, L. (2004). *Spanish/English codeswitching in a written corpus*. Amsterdam; Philadelphia: John Benjamins Publishing.

Campbell, L. (2013). *Historical linguistics*. Edinburgh: Edinburgh University Press.

Cantone, K.F. & MacSwan, J. (2009). The syntax of DP-internal codeswitching. In Isurin, L., Winford, D. & De Bot, K. (Eds.), *Multidisciplinary approaches to code switching* (pp. 243-78). Amsterdam/Philadelphia: John Benjamins Publishing.

Cardona Ramírez, M.A. (2010). La fonética del español en Belice, In Quesada, M. Á. (Ed.), *El español hablado en América Central. Nivel fonético,* (pp. 21-48). Fráncfort: Vervuert.

Cassidy, F. G., & LePage, R. B. (1967). *Dictionary of Jamaican English*. Cambridge: Cambridge University Press.

Chatterjee, T. (2012). Bilingual complex verbs: So what's new about them? In *Proceedings of the annual meeting of the Berkeley Linguistics Society 38,* (pp. 47-62). Berkeley, CA.

Childs, B., Van Herk, G. & Thorburn, J. (2011). Safe harbour: Ethics and accessibility in sociolinguistic corpus building. *Corpus Linguistics and Linguistic Theory*, 7(1), 163-180.

Chomsky, N. (2001). Derivation by phase. In Kenstowicz, M. (Ed.)*, Ken Hale: A life in language* (pp. 1-52). Cambridge: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Clegg, J.H. (2010). On becoming a borrowing: Integration, diffusion and attestation of English-origin nouns in New Mexico Spanish. In Rivera-Mills, S. & Trujillo, J.A. (Eds.), *Building communities and making connections* (pp. 184-201). New Castle upon Tyne: Cambridge Scholars Publishing.

Company Company, C. (1994). Semántica y sintaxis de los posesivos duplicados en el español de los siglos XV y XVI. *Romance Philology*, 48(3), 111-135.

Comrie, B. (2012). Some argument-structure properties of "give" in the languages of Europe and northern and Central Asia. In Suihkonen, P. B. Comrie, B. & Solovyev, V. (Eds.), *Argument structure and grammatical relations: A cross-linguistic typology* (pp. 17–35). Amsterdam: Benjamins.

Cuervo, C. (2003). *Datives at large*. Doctoral Dissertation. Massachusetts Institute of Technology.

D'Arcy, A. (2007). Like and language ideology: Disentangling fact from fiction. *American Speech*, *82*(4), 386-419.

Davies, M. (2002) *Corpus del Español: 100 million words, 1200s-1900s*. Available online at http://www.corpusdelespanol.org.

Davies, M. (1995). Analyzing syntactic variation with computer-based corpora: The case of modern Spanish clitic climbing. *Hispania*, 78, 370–380.

Decker, K. (2005) *The song of Kriol*. Belize City: The Belize Kriol Project.

Deuchar, M. (2011). Review of B. E. Bullock & A. J. Toribio (Eds.), The Cambridge handbook of linguistic code-switching. *Studies in Second Language Acquisition*, 33, 128-129.

Deuchar, M., Muysken, P., & Wang, S. L. (2007). Structured variation in codeswitching: towards an empirically based typology of bilingual speech patterns. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 298-340.

Dowty, D. (1979). Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ. Dordrecht: Kluwer.

DuBord, E.M. (2004). Gender assignment to English words in the Spanish of Southern Arizona. *Divergencias : Revista de Estudios Lingüísticos y Literarios*, *2*(2), 27-39.

Eastman, C.M. (1992). Codeswitching as an urban and language-contact phenomenon. In Eastman, C. (Ed.), *Codeswitching* (pp. 1-18). Clevedon: Multilingual Matters.

Eckert, P. (2003). Elephants in the room. *Journal of Sociolinguistics*, 7(3), 392-397.

Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. New York: Teachers College Press.

Eckert, P. & Rickford, J.R. (2001). *Style and sociolinguistic variation*. Cambridge University Press.

Edwards, M. & Gardner-Chloros, P. (2007). Compound verbs in codeswitching: Bilinguals making do? *International Journal of Bilingualism*, *11*(1), 73-91.

Eliasson, S. (1991). Models and constraints in code-switching theory. In *Papers from the Workshop on Constraints, Conditions and Models* (pp. 17-50). Strasbourg: European Science Foundation.

Eliasson, S. (1989). English-Maori language contact: Code-switching and the free morpheme constraint. *Reports from Uppsala University.* Department of Linguistics, (pp. 18:1–28). Uppsala University, Sweden.

Eppler, E. (1999). Word order in German-English mixed discourse. *UCL Working Papers in Linguistics*, *11*, 285-308.

Escure. G. (2013a). Belizean Creole structure dataset. In Michaelis, S. M., Maurer, P., Haspelmath, M. & Huber, M. (Eds.), *Atlas of pidgin and creole language structures online*. Leipzig: Max Planck

Institute for Evolutionary Anthropology. (Available online at http://apics-online.info/contributions/9, Accessed on 2015-03-11.)

Escure, G. (2013b). Belizean Creole. In Michaelis, S.M., Maurer, P., Haspelmath, M. & Huber, M. (Eds.), *The survey of pidgin and creole languages* (Vol. 1) (pp. 92-100). Oxford: Oxford University Press.

Escure, G. (2004). Garifuna in Belize and Honduras. In Escure, G. & Schwegler, A. (Eds.), *Creole Language Library*, *27*, (pp. 35-66). Amsterdam/Philadelphia: John Benjamins

Eze, E. (1998). Lending credence to a borrowing analysis: Lone English-origin incorporations in Igbo discourse. *International Journal of Bilingualism*, *2*(2), 83-201.

Eze, E. (1997). *Aspects of language contact: A variationist perspective on codeswitching and borrowing in Igbo-English bilingual discourse*. Unpublished Doctoral Dissertation. University of Ottawa (Canada*).*

Ferguson, C.A., (1959). Diglossia. *Word*, *15*, 325-40.

Fernández-Fuertes, R., J. M. Liceras and A. Alba de la Fuente. (2016) Beyond the subject DP versus the subject pronoun divide in agreement switches. In Tortora, Ch. (Ed.), *Selected proceedings of the 43th Linguistic Symposium on Romance Languages (LSRL-43).* (pp. 79-98). Amsterdam: John Benjamins.

Fernández Soriano, M. O. (1999). El pronombre personal : Formas y distribuciones. Pronombres átonos y tónicos. In Bosque, I & Demonte, V. (Eds.), *Gramática descriptiva de la lengua española (Vol 1)* (pp. 1204-1274). Madrid: Espasa Calpe.

Flores-Ferrán, N. (2014). So pues entonces: An examination of bilingual discourse markers in Spanish oral narratives of personal experience of New York City-born Puerto Ricans. *Sociolinguistic Studies*, *8*(1), 57–83.

Folli, R. & Harley, H. (2007). Causation, obligation, and argument Structure: On the nature of little *v*. *Linguistic Inquiry, 38*(2), 197-238.

269

Fontanella de Weinberg, M. B. (1999). Sistemas pronominales de tratamiento usados en el mundo

    hispánico. In Bosque, I. & Demonte, V. (Eds.), *Gramática descriptiva de la lengua española*, (vol.

    1), (pp. 1399-1425). Madrid: Espasa.

Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of

    some current proposals. *Second Language Research, 17*(3), 213-247.

Franco, J. (2000). Agreement as a continuum. In Beukema, F., & Den Dikken, M. (Eds.). *Clitic*

    phenomena in European languages (Vol. 30), (pp. 147-190). Amsterdam/Philadelphia: John

    Benjamins Publishing.

Fuller Medina, N. (2015). Let's talk about Spanish: A sociolinguistic study of Spanish in Belize. In

    Encalada, N., Cocom, R., Pelayo, P., & Pinelo, G. (Eds.), *Research reports in Belizean history and*

    anthropology, Volume 3: Proceedings of the XII annual Belize Archaeology and Anthropology

    Symposium (pp. 166-176). Benque Viejo Town, Belize: BRC Printing.

Fuller Medina, N. (2014). Let's talk about Spanish: A sociolinguistic study of Spanish in Belize. Paper

    presented at *The XII Annual Belize Archaeology and Anthropology Symposium*. ISCR/IA. San

    Ignacio, Belize. July 1-4.

Fuller Medina, N. (2013). The Syntax of [*Hacer* +V]: *Haciendo* borrow in Belizean Spanish. *Paper*

    presented at the Hispanic Linguistics Symposium, University of Ottawa, October 19.

Fuller Medina, N. (2010). Haciendo borrow: Bilingual compound verbs in Belizean Spanish. Poster

    presented at the *18th Biennial conference of the Society for Caribbean Linguistics- Caribbean*

    Languages and Popular Culture. Barbados, August 9-13.

Fuller Medina, N. (2005a). *El uso de* hacer *como estrategia de préstamo: La perífrasis 'hacer +V' en el*

    español beliceño. (Master's Thesis). University of Western Ontario, London, Ontario.

Fuller Medina, N. (2005b). Spanish-English contact in Belize: the case of '*Hacer* +V'". In Gurski, C. (Ed),

   Proceedings of the 2005 Canadian Linguistics Association annual conference.

   http://westernlinguistics.ca/Publications/CLA-ACL/FullerMedina.pdf

Fuller Medina, N. & Roy, J. (2010) A variationist approach to sex: Statistics and the construction of

   social meaning" paper presented at the *New Ways of Analyzing Variation 39,* San Antonio, Texas

   November 4-6.

Gardner-Chloros, P. (2009). *Code-switching.* Cambridge/New York: Cambridge University Press.

Gardner-Chloros, P. (1995). Code-switching in community, regional and national repertoires: The myth

   of the discreteness of linguistic systems. In Milroy, L. & Muysken, P. (Eds.)*, One speaker, two*

   languages: Cross-disciplinary perspectives on code switching (pp. 68-89). Cambridge: Cambridge

   University Press.

Gardner-Chloros, P. (1987). Code-switching in relation to language contact and convergence. In Ludi, G.

   (Ed.), *Devenir bilingue*/parler bilingue* (pp. 33-113). Tubingen: Max Niemeyer Verlag.

Gardner-Chloros, P. & Edwards, M. (2002). Touching base: The relevance of grammatical models to

   code-switching.In *Actas/ Proceedings: II Simposio Internacional bilinguismo*, 1433-1452.

   http://ssl.webs.uvigo.es/actas2002/07/05.%20Penelope%20Gardner.pdf

Gast, V. (2007). I gave it him—on the motivation of the 'alternative double object construction' in

   varieties of British English. *Functions of Language*, *14*(1), 31-56.

Gómez Molina, J.R. & Gómez Devís, M.B. (2004). *La disponibilidad léxica de los estudiantes*

   preuniversitarios valencianos: Estudio de estratificación sociolingüística. Valencia: Universitat de

   València.

González López, V. (2008). *Spanish clitic climbing*. Unpublished Doctoral Dissertation. Pennsylvania

   State University. State College.

González-Vilbazo, K., Bartlett, L., Downey, S., Ebert, S., Heil, J., Hoot, B., Koronkiewicz, B. & Ramos, S. (2013). Methodological considerations in code-switching research. *Studies in Hispanic and Lusophone Linguistics, 6*(1), 118-138.

González-Vilbazo, K. & López, L. (2012). Little *v* and parametric variation. *Natural Language & Linguistic Theory, 30* (1), 33-77.

González-Vilbazo, K. & López, L. (2011). Some properties of light verbs in code-switching. *Lingua, 121*(5), 832-850.

Gooden, A., & Hackett, V. R. (2013). 'Outsider Within': Reflexivity and working with African Caribbean immigrants in qualitative research. *Southern Journal of Canadian Studies, 5*(1).

Greene, L. A. (1999). *A grammar of Belizean Creole: compilations from two existing United States dialects*. New York: Peter Lang.

Gumperz, J.J. (1982). *Discourse strategies*. Cambridge: Cambridge University Press.

Gumperz, J. (1971). *Language in social groups*. Stanford, CA: Stanford University Press.

Gutiérrez, M.J. & Silva Corvalán, C. (1993). Spanish in a contact situation. In Roca, A. & Lipski, J.M. (Eds.), *Spanish in the United States: Linguistic contact and diversity*. Berlin; New York: Mouton de Gruyter.

Guy, G. (1988). Advanced VARBRUL analysis. In Ferrara, K., Brown, B., Walters, K. & Baugh, J . (Eds.), *Linguistic change and contact*, (pp. 124-136). Austin, TX: Department of Linguistics, University of Texas at Austin.

Hagerty, T.W. (1996). The influence of English on the Spanish of Belize. In Phillips, M.D. (Ed.), *Belize: Selected proceedings from the second Interdisciplinary Conference.* (pp. 131-142). Madison.: University Press of America.

Hagerty, T.W. (1979). *A Phonological analysis of the Spanish of Belize.* Unpublished Doctoral Dissertation. University of California, Los Angeles, California.

Halmari, H. (1997). *Government and codeswitching: explaining American Finnish*. Amsterdam/Philadelphia: John Benjamins Publishing.

Halmari, H. & Regetz, T. (2012). Syntactic aspects of code-switching. In Schendl, H. & Wright, L. (Eds.), *Code-switching in early English* (pp. 115-153). Berlin: Walter de Gruyter.

Harley, H. (2009). The morphology of nominalizations and the syntax of *vP*. In Monika, R. & Anastasia, G., (Ed.), *Quantification, definiteness and nominalization*. (pp 321-343). Oxford: OUP.

Harley, H. (2005). How do verbs get their names? Denominal verbs, manner incorporation, and the ontology of verb roots in English. In Erteschik-Shir, N., & Rapoport, T. (Eds.), *The Syntax of aspect: Deriving thematic and aspectual interpretation* (Vol. 10), (pp 42-64). Oxford: OUP.

Harley, H. & Noyer, R. 1999. State-of-the-article: Distributed morphology. *GLOT International*, 4,3-9.

Harris, J. (1991). The exponence of gender in Spanish. *Linguistic Inquiry, 22*(1), 27-62.

Harris, J. W. (1969). Sound change in Spanish and the theory of markedness. *Language,* 45, 538-552.

Hasselmo, N. (1970). Code-switching and modes of speaking. In Gilbert, G. (Ed.), *Texas studies in bilingualism* (pp. 179-210). Berlin: Walter de Gruyter.

Haugen, E.I. (1956). *Bilingualism in the Americas: A bibliography and research guide American Dialect Society Monograph 26*. Alabama: University of Alabama Press.

Hazen, K. (2000). The role of researcher identity in conducting sociolinguistic research: A reflective case study. *Southern Journal of Linguistics*, *24*, 103-120.

Heine, B., & Kuteva, T. (2005). *Language contact and grammatical change*. Cambridge; New York: Cambridge University Press.

Herrera, Y., Manzanares, M., Woods, S., Crosbie, C., Decker, K., & Crosbie.P. (2007). *Kriol–Inglish Dikshineri, English–Kriol Dictionary*. Belmopan: Print Belize.

Herring, J.R., Deuchar, M., Couto, M.C.P., & Quintanilla, M.M. (2010). 'I saw the madre': evaluating predictions about codeswitched determiner-noun sequences using Spanish–English and Welsh–English data. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 553-573.

Himmelman, N. (1998). Documentary and descriptive linguistics. *Linguistics* 36: 161-195.

Hoebens, E.L. (2000). Bileez Kriol : Usos, desusos y abusos. *DiversCité ILangues*. Retrieved from http://www.teluq.uquebec.ca/diverscite/SecArtic/Arts/2000/hoebens/hoebens_txt.htm.

Hoi Ying Chen, K. (2014). Styling bilinguals: Analyzing structurally distinctive code-switching styles in Hong Kong. In Stell, G., & Yakpo, K. (Eds.), *Codeswitching between structural and sociolinguistic perspectives* (pp. 163-185). Berlin: Walter de Gruyter.

Holm, J. (1982). *Central American English*. Heidelberg: Verlag.

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge: Cambridge University Press.

Hualde, J.I. (2013). *Los sonidos del español: Spanish language edition*. Cambridge: Cambridge University Press.

Jake, J.L., Myers-Scotton, C., & Gross, S. (2002). Making a minimalist approach to codeswitching work: Adding the matrix language. *Bilingualism: Language and Cognition*, *5*(1), 69-91.

Jackendoff, R. (1996). The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language & Linguistic Theory*, *14*(2), 305-354.

Jackendoff, R. (1990). On Larson's treatment of the double object construction. *Linguistic Inquiry*, *21*(3), 427-456.

Jenkins, D.L. (2003). Bilingual verb constructions in southwestern Spanish. *Bilingual Review/La Revista Bilingüe*, *27*(3), 195-204.

Joseph, V. L. (2003). *Differential performance in English tests among different language groups: Sociolinguistic study of Belizean speech communities*. Unpublished Doctoral Dissertation. University of Kansas.

Kayne, R. S. (1975). *French syntax: The transformational cycle*. Cambridge: MIT Press.

Krifka, M. (1989). Nominal reference, temporal constitution and quantification in event semantics: *Semantics and contextual expression*, *75*, 115.

Labov, W. (2006). *The social stratification of English in New York City*. (2nd Ed). Washington: Centre for Applied Linguistics.

Labov, W. (2001). The anatomy of style shifting. In Eckert, P. & Rickford, J.R., (Eds.), *Style and sociolinguistic variation* (pp. 85-108). Cambridge: Cambridge University Press.

Labov, W. (1996). When intuitions fail. In McNair, L., Singer, K., Dolbrin L., & Aucon, M. (Eds.), *32nd regional meeting of the Chicago Linguistics Society 32* (pp. 76-106).

Labov, W. (1984). Field methods of the project on linguistic change and variation. In Baugh, J., & Sherzer, J., (Eds.), *Language in use: Readings in sociolinguistics* (pp. 28-53). Englewood Cliffs: Prentice Hall.

Labov, W. (1972). *Sociolinguistic patterns.* Philadelphia: University of Pennsylvania Press.

Larson, R. K. (1988). On the double object construction. *Linguistic inquiry*, *19*(3), 335-391.

Larson, R. K. (1990). Double objects revisited: Reply to Jackendoff. *Linguistic inquiry*, *21*(4), 589-632.

Latimer, J. (2009). *Buccaneers of the Caribbean: How piracy forged an empire.* Harvard University Press.

LeBlanc, C. (2010). Coding compositional aspect in French. In Walker, J.A., (Ed.), *Aspect in grammatical variation* (pp. 65-80). Amsterdam/Philadelphia: John Benjamins Publishing.

Lehman, M. (1982). Introduction: Diachronic linguistics. In Lehmann, W. & Malkiel, Y. (Eds.) *Perspectives on historical linguistics*. (Amsterdam Studies in the Theory and History of Linguistic Science, IV. Current Issues in Linguistic Theory, 24.) (pp. 1-16) Amsterdam: Benjamins.

Le Page, R. B., & Tabouret-Keller, A., (1985). *Acts of identity: Creole-based approaches to language and ethnicity.* Reprinted (2006). (2. éd. augmentée). Brussels : Éd. Modulaires Européennes.

Levey, S. (2007). *The next generation: Aspects of sociolinguistic variation in preadolescence.* Doctoral Dissertation. Queen Mary, University of London, London.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago press.

Levin, M. (2013). The progressive verb in modern American English. In Aarts, B., Close, J., Leech, G., & Wallis, S. (Eds.). (2013). *The verb phrase in English: Investigating recent language change with corpora*. (pp 187-216). Cambridge: Cambridge University Press.

Levon, E. (2013). Ethnographic fieldwork. In Mallinson, C., Childs, B., & Van Herk, G. (Eds.), *Data collection in sociolinguistics: methods and applications*. (pp 69-79). New York: Routledge, Taylor & Francis Group*.*

Lewis, E. G. (1978). Bilingual education and social change in the Soviet Union. In Spolsky, B. D., & Cooper, R. L. (Eds.)*, Case Studies in Bilingual Education.* (pp. 203-248) Rowley, MA.: Newbury House.

Lewis, E.G. (1972). *Multilingualism in the Soviet Union*. Hague : Walter De Gruyter Mouton.

Liceras, J.M., Fernández Fuertes, R. & Klassen, R. (2015). Language dominance and language nativeness: the view from English-Spanish codeswitching. In Guzzardo Tamargo, R., Mazak, C., Parafita Couto, M.C. (Eds.), *Spanish-English codeswitching in the Caribbean and the U.S. Issues in Hispanic and Lusophone Linguistics series*. Amsterdam/Philadelphia: John Benjamins.

Liceras, J.M. (2013). Gender agreement patterns in mixed concord and agreement structures: does 'code-switching' matter? Paper presented at the *International Conference "Code-switching in the bilingual child: Within and across the clause"*. Bergische Universität Wuppertal (North Rhine-Westphalia, Germany).

Liceras, J. M., Fuertes, R. F., Perales, S., Pérez-Tattam, R., & Spradlin, K. T. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua*, *118*(6), 827-851.

Liceras, J. M., Fernández-Fuertes, R. Perales, S., Rocío Pérez-Tattam, R. & Martínez, C. (2006). L2A and the pidgin/creole *continuum*: The pidginization/nativization hypothesis revisited. Paper presented at *EUROSLA*, Boğaziçi University, Antalya, Turkey.

Lipski, J. (2015). Is Spanglish the third language of the south? In Picone, M. & Evans Davies, C. (Eds.), *New perspectives on language variety in the south: Historical and contemporary approaches* (pp. 657 -677). Tuscaloosa: University of Alabama Press.

Lipski, J.M. (2012). Geographical and social varieties of Spanish: An overview. In Hualde, J.I., Olarrea, A. & O'Rourke, E. (Eds.), *The handbook of Hispanic linguistics* (pp. 1-26). Malden, MA: Wiley-Blackwell.

Lipski, J.M. (2009). "Fluent dysfluency" as congruent lexicalization: A Special case of radical code-mixing. *Journal of Language Contact*, *2*(2), 1-39.

Lipski, J.M. (2008). *Varieties of Spanish in the United States.* Washington, DC: Georgetown University Press.

Lipski, J. (2005). Code-switching or borrowing? No sé so no puedo decir, you know. In Sayahi, L. & Westmoreland, M. (Eds.), *Selected proceedings of the second Workshop on Spanish Sociolinguistics* (pp. 1-15). Somerville, MA: Cascadilla Publishing.

Lipski, J.M. (1985). *Linguistic aspects of Spanish-English language switching Special Studies 25*.

    Tempe, Arizona: Arizona State University Centre for Latin American Studies.

Lipski, J. (1978). Code-switching and the problem of bilingual competence. In Paradis, M. (Ed.), *Aspects*

    of bilingualism (pp. 250-264). Columbia SC: Hornbeam Press.

MacSwan, J. (2014). *A minimalist approach to intrasentential code switching*. New York: Routledge.

MacSwan, J. (2010). Plenary address: Unconstraining codeswitching theories. In *Proceedings from the*

    annual meeting of the Chicago Linguistic Society 44 (2) (pp. 151-168). Chicago: University of

    Chicago Press.

MacSwan, J. (2005). Codeswitching and generative grammar: A critique of the MLF model and some

    remarks on "modified minimalism". *Bilingualism: Language and Cognition*, *8*(01), 1-22.

MacSwan, J. (2000). The architecture of the bilingual language faculty: Evidence from intrasentential

    code switching. *Bilingualism: Language and Cognition*, *3*(01), 37-54.

MacSwan, J. (1999). *A Minimalist approach to intrasentential code switching*. New York: Garland Press.

Mallinson, C., Childs, B., & Van Herk, G. (Eds.). (2013). *Data collection in sociolinguistics: methods and*

    applications. New York: Routledge, Taylor & Francis Group.

Marantz, A. (1997). No escape from syntax: Don't try morphological analysis in the privacy of your own

    lexicon. *University of Pennsylvania working papers in linguistics*, *4*(2), 14.

Marín, R., & McNally, L. (2011). Inchoativity, change of state, and telicity: Evidence from Spanish

    reflexive psychological verbs. *Natural Language & Linguistic Theory*, *29*(2), 467-502.

Matras, Y. (2009). *Language contact*. Cambridge New York: Cambridge University Press.

Matras, Y. (2007). The borrowability of structural categories. In Matras, Y., & Sakel, J. (Eds.),

    Grammatical borrowing in cross-linguistic perspective. Empirical approaches to language typology,

    (Vol. 38), (pp 31-73). Berlin/New York: Walter de Gruyter.

Mayr, R.J. (2014). *Belize: tracking the path of its history: From the heart of the Mayan empire to a retreat for buccaneers, a safe-haven for ex-pirates and pioneers, a crown colony and a modern nation* (Vol. 13). Zurich/Berlin: LIT.

McClure, E. (2001). Oral and written Assyrian-English codeswitching. In Jacobson, R. (Ed.), *Codeswitching worldwide. Trends in linguistics studies and monographs*, (126), (pp. 157-192). Berlin/New York: Mouton de Gruyter.

McClure, E. (1977). Aspects of code-switching in the discourse of bilingual Mexican-American children. *Linguistics and anthropology: Georgetown University round table on languages and linguistics*, 93-115.

Meakins, F. & O'Shannessy, C. (2012). Typological constraints on verb integration in two Australian mixed languages. *Journal of Language Contact, 5*(2), 216-246.

Meechan, M. & Poplack, S. (1995). Orphan categories in bilingual discourse: Adjectivization strategies in Wolof-French and Fongbe-French. *Language Variation and Change*, 7(2), 69-194.

Mendieta, E. (1999). *El préstamo en el español de los Estados Unidos*. New York: Peter Lang.

Meyerhoff, M., Chie, A., Nanbakhsh, G. & Strycharz, A. (2012). Sociolinguistic fieldwork. In Thieberger, N. (Ed.), *The Oxford handbook of linguistic fieldwork* (pp. 121–146). Oxford: Oxford University Press.

Meyerhoff, M. (2009). Replication, transfer, and calquing: Using variation as a tool in the study of language contact. *Language Variation and Change*, *21*(03), 297-317.

Michnowicz, J. (2008). Final nasal variation in Merida, Yucatan. *Spanish in Context*, *5*(2), 278-303.

Milian Hita, S. (2014). Codeswitching and the syntax-semantics interface: The role of aspectual features in constraining intrasentential codeswitching involving the verb. In MacSwan, J. (Ed.), *Grammatical theory and bilingual codeswitching*, (pp. 227-251). Cambridge, Massachusetts: MIT Press.

Milroy, L., & Gordon, M. (2003). *Sociolinguistics: method and interpretation*. Oxford: Blackwell.

Milroy, J., & Milroy, L. (1978). Belfast: Change and variation in an urban vernacular. In Trudgill, P. (Ed.)
Sociolinguistic patterns in British English (pp. 19-36). London: Edward Arnold.

Moravcsik, E. (1975). Verb borrowing. *Wiener Linguistische Gazette*, (8), 3-31.

Morin, R. (2006). Spanish gender assignment in computer and Internet related loanwords. *Journal Devoted to the Teaching of Spanish and Portuguese*, *61*(4), 865-876.

Morin, R. (2006). Evidence in the Spanish language press of linguistic borrowings of computer and Internet-related terms. *Spanish in Context*, *3*(2), 161-179.

Moro, M. (2015). The universality of syntactic constraints on Spanish–English codeswitching in the USA. *Language and Intercultural Communication, 15*(3), 1-16.

Moseley, C. (Ed.). (2010). Atlas of the world's languages in danger, 3rd edn. Paris: UNESCO Publishing. Online version: http://www.unesco.org/culture/en/endangeredlanguages/atlas

Moyer, M. (1993). *Analysis of code-switching in Gibraltar.* Unpublished Doctoral Dissertation. Universidad Autónoma de Barcelona, Spain.

Muysken, P. (2015). Deja voodoo or new trails ahead? In Torres Cacoullos, R., Dion, N. & Lapierre, A. (Eds.) *Linguistic variation: Confronting fact and theory* (pp. 242-262)*.* New York; London: Routledge, Taylor & Francis Group.

Muysken, P. (2013). Language contact outcomes as the result of bilingual optimization strategies. *Bilingualism: Language and Cognition*, *16*(04), 709-730.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing.* Cambridge: Cambridge University Press.

Myers-Scotton, C. (2009). Supporting a differential access hypothesis. In Kroll, J. & De Groot, A. (Eds.) *Handbook of bilingualism: psycholinguistic approaches*, (pp. 326-348). New York: Oxford University Press.

Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes.* Oxford; New York: Oxford University Press.

Myers-Scotton, C. (1993). *Dueling languages: Grammatical structure in codeswitching.* Oxford: Oxford University Press.

Myers-Scotton, C. & Jake, J.L., (2001). Explaining aspects of codeswitching and their implications. In Nicol, J.L. (Ed.), *One mind, two languages: Bilingual language processing* (pp. 84-116). Malden, MA: Blackwell.

Myhill, J. (1988). Variation in Spanish clitic climbing. In Walsh, T. J. (Ed.), *Synchronic and diachronic approaches to linguistic variation and change* (pp. 227–250). Washington, DC: Georgetown University Press.

Nakajima, T. (2008). Loan word syntax: A case in the light verb construction. *Toronto Working Papers in Linguistics*, *28, 260-278.*

National Institute of Culture and History (NICH). (2014). Our cultures, our values, our identity, our prosperity: National culture policy, NICH: Belize.

Navarro Tomás, N. (1918). *Manual de pronunciación española.* Madrid: Centro de Estudios Históricos.

Nishimura, M. (1995). Varietal conditioning in Japanese/English code-switching. *Language Sciences*, *17*(2), 123-145.

Olsen, M. B. (1997). *A semantic and pragmatic model of lexical and grammatical aspect.* New York: Garland Publishing.

Olsen, M.A. (1994). *A Semantic pragmatic model of aspect.* Doctoral Dissertation. Northwestern University.

Ordóñez, F. (2012). Clitics in Spanish. In Hualde, J. I., Olarrera, A & O'Rourke, E. (Eds.), *The Handbook of Hispanic Linguistics* (Vol 69), (pp. 423-452). Oxford: Wiley- Blackwell

Otheguy, R. & Lapidus, N. (2003). An adaptive approach to noun gender in New York contact Spanish. In Cameron, R., López, L. & Núñez-Cedeño, R. (Eds.), *A Romance perspective on language knowledge and use* (pp. 209-229). Amsterdam: John Benjamins.

Otheguy, R., & Stern, N. (2011). On so-called Spanglish. *International Journal of Bilingualism*, *15*(1), 85-100.

Palacio, J. O. (1990). *Socioeconomic integration of Central American immigrants in Belize. SPEAReports 2*. Belize City: Cubola Productions.

Palmer, F.R. (1987). The typology of subordination: results, actual and potential. *Transactions of the Philological Society, 85*(1), 90-109*.*

Parafita Couto, M.C., Deuchar, M. & Fusser, M. (2014). How do Welsh-English bilinguals deal with conflict? Adjective-noun order resolution. In Stell, G. & Yakpo, K. (Eds.), *Code-switching between structural and sociolinguistic perspectives* (pp. 65-84). Berlin : Walter de Gruyter.

Parafita Couto, M.D.C., Carter, D., Davies, P. & Deuchar, M. (2011). Factors determining code-switching patterns in Spanish-English and Welsh-English communities. *Revista Española de Lingüística Aplicada*.

Pato, E. (2002). La estructura posesiva *una mi amiga* en el español de Guatemala, In Palacios, A. & García, A. I. (Eds.), *El indigenismo americano III. Actas de las terceras jornadas sobre indigenismo americano* (pp. 121-154). Valéncia : Universitat de Valéncia.

Pérez-Leroux, A.T., O'Rourke, E. & Sunderman, G. (2014). Language dominance and codeswitching asymmetries. In MacSwan, J. (Ed.), *Grammatical theory and bilingual codeswitching* (pp. 283-312). New York: Routledge.

Pfaff, C.W. (1979). Constraints on language mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language,* 291-318*.*

Pinelo, G. (2015). Western Belize, 1916: In the crossfire of a counter-revolution. In Encalada, N., Cocom, R., Pelayo, P., & Pinelo, G. (Eds.), *Research reports in Belizean history and anthropology, Volume 3: Proceedings of the XII Annual Belize Archaeology and Anthropology Symposium* (pp. 177- 187). Benque Viejo Town, Belize: BRC Printing.

Poplack, S. (2015). Code-switching (Linguistic). In Wright, J.D. (editor-in-chief) *International encyclopedia of the social and behavioral sciences, 2nd edition*. (pp. 918-925). Oxford: Elsevier Science Ltd.

Poplack, S. (2004). Code-switching. In Ammon, U., Dittmar, N., Mattheier, K.J. & Trudgill, P. (Eds.), *Soziolinguistik. An international handbook of the science of language* (pp. 589-596). Berlin: Walter Degruyter.

Poplack, S. (2000). Introduction. In Poplack, S. (Ed.), *The English history of African American English* (pp. 1-32). Oxford: Wiley-Blackwell.

Poplack, S. (1993). Variation theory and language contact: Concepts, methods and data. In Preston. D. (Ed.), *American dialect research* (pp. 251-286). Amsterdam: John Benjamins.

Poplack, S. (1987). Contrasting patterns of code-switching in two communities. In Wande, E., Anward, J., Nordberg, B., Steensland, L. & Thelander, M., (Eds.), *Aspects of multilingualism* (pp. 51-77). Uppsala: Borgströms. Reprinted from Warkentyne, H. J. (Ed.) *Methods V: papers from the V international conference on Methods in Dialectology*. (pp. 363-385). Victoria, B.C.: University of Victoria.

Poplack, S. (1981). Syntactic structure and social function of code-switching. In Duran, R. (Ed.), *Latino language and communicative behavior* (pp. 169-184). New Jersey: Ablex Publishing Corp.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching. *Linguistics*, *18*(7-8), 581–618.

Poplack, S. & Torres Cacoullos, R. (2015). Linguistic emergence on the ground: A variationist paradigm. In MacWhinney, B. & O'Grady, W. (Eds.), *The handbook of language emergence* (pp. 267-291). West Sussex: Wiley Blackwell.

Poplack, S. & Dion, N. (2012). Myths and facts about loanword development. *Language Variation and Change*, *24* (3), 279-315.

Poplack, S. & Levey, S. (2010). Contact-induced grammatical change. In Auer, P. & Schmidt, J. E., (Eds.), *Language and Space – An international handbook of linguistic variation: Volume 1 – Theories and methods* (pp. 391-419). Berlin: Mouton de Gruyter.

Poplack, S., Walker, J. A., & Malcolmson, R. (2006). An English "like no other?": Language contact and change in Quebec. *The Canadian Journal of Linguistics/La revue canadienne de linguistique*, *51*(2), 185-213.

Poplack, S., & Tagliamonte, S. (2001). *African American English in the diaspora*. Malden, MA: Blackwell.

Poplack, S. & Meechan, M. (1998a). How Languages fit together in codemixing. *International Journal of Bilingualism*, *2*(2), 127-38.

Poplack, S. & Meechan, M. (1998b). (Eds.), Instant loans, easy conditions: the productivity of bilingual borrowing. [Special issue]. *International Journal of Bilingualism*, *2*(2).

Poplack, S. & Meechan, M. (1995). Patterns of language mixture: Nominal structure in Wolof-French and Fongbe-French bilingual discourse. In Muysken, P. & Milroy, L. (Eds.), *One speaker, two languages* (pp. 199-232). Cambridge: Cambridge University Press.

Poplack, S., Sankoff, D. & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, *26*(1), 47-104.

Poplack, S., Wheeler, S. & Westwood, A. (1987). Distinguishing language contact phenomena: evidence from Finnish-English bilingualism. In Lilius, P. & Saari, M. (Eds.), *The Nordic languages and modern linguistics 6* (pp. 35-56). Helsinki: University of Helsinki Press.

Poplack, S., Pousada, A. & Sankoff, D. (1982). Competing influences on gender assignment: Variable process, stable outcome. *Lingua*, *57*(1), 1-28.

Quilis, A. (1990). Notas sobre el español de Belice. *Voz y Letra: Revista de Filologia*, *1*, 139-148.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York: Longman.

Ramchand, G. (2007). Verb meaning and the lexicon: A first phase syntax. Draft Ms. University of Tromsø.

Ravindranath, M. (2009). *Language shift and the speech community: Sociolinguistic change in a Garifuna community in Belize*. Unpublished Doctoral Dissertation. University of Pennsylvania, Philadelphia.

Real Academia Espanola (RAE) (2009). *Nueva gramática de la lengua española. Fonética y fonología. morfología. (Sintaxis)*, (*3*). Madrid : Asociación de Academias de la Lengua Española.

Real Academia Española (RAE). (2010). *Manual de la nueva gramática de la lengua española*. Madrid : Asociación de Academias de la Lengua Española.

Real Academia Española y Asociación de Academias de la Lengua Española (RAE and AALE). (2010). *Nueva gramática de la lengua española. Manual.* Madrid: Espasa.

Reyes, R. (1976). Language mixing in Chicano bilingual speech. In Bowen, J. & Ornstein, J. (Eds.), *Studies in Southwest Spanish* (pp. 183-188). Rowley, Massachusetts: Newbury House.

Rickford, J.R. (1986). The need for new approaches to social class analysis in sociolinguistics. *Language & Communication*, *6*(3), 215-221.

Rickford, J. R., & McNair -Knox, F. (1994). Addressee and topic influenced style shift: A quantitative sociolinguistic study. In Biber, D. & Finegan, E. (Eds.), *Perspectives on register: Situating register variation within sociolinguistics*, (pp 235-276). Oxford: Oxford University Press.

Riegelhaupt, F., Carrasco, R. L., & Brandt, E. (2003). Spanish: A language of indigenous peoples of the Americas. In Reyhner J., Trujillo, O.V., Carrasco, R.L., and Lockard, L. (Eds.) *Nurturing native languages*, (pp. 129-140). Flagstaff: University of Arizona Press.

Ritchie, W.C. & Bhatia, T.K. (1996). Codeswitching, grammar, and sentence production: The problem of dummy verbs. Retrieved from http://files.eric.ed.gov/fulltext/ED403776.pdf.

Rivas, A. M. (1977). *A theory of clitics*. Unpublished Doctoral dissertation, Massachusetts Institute of Technology.

Rivas, J. (2013). Variable subject position in main and subordinate clauses in Spanish: a usage-based approach. *Moenia,* 19, 97-113.

Roberts, I. (1993). Verbs and diachronic syntax: a comparative history of English and French. Dordrecht: Kluwer.

Roca, I. (1989). The organization of grammatical gender. *Transactions of the Philological Society, 87,* 1-32.

Roca, A., & Jensen, J. B. (Eds.). (1996). *Spanish in contact: Issues in bilingualism*. Somerville, MA: Cascadilla Press.

Roca, A., & Lipski, J. M. (Eds.). (1993). *Spanish in the United States: Linguistic contact and diversity*. Berlin; New York: Mouton de Gruyter.

Romaine, S. (1989). *Bilingualism*. Oxford: Blackwell.

Romaine, S. (1986). The syntax and semantics of the code mixed compound verb in Panjabi/English bilingual discourse. In Deborah, T. & Alatis, J. E. (Eds.)*, Languages and linguistics: The interdependence of theory, data, and application* (pp. 35-50). Washington, D.C.: Georgetown University Press.

Rothmayr, A. (2009). *The structure of stative verbs* (Vol. 143). Amsterdam/Philadelphia: John Benjamins Publishing.

Rothstein, S. (in press)**.**   Aspect. In Aloni, M. & Dekker, P. (Eds.), *Cambridge handbook of semantics* (pp. 384-420)*.* Cambridge: Cambridge University Press.

Rothstein, S. (2008). Telicity, atomicity and the Vendler classification of events. In Rothstein, S. (Ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect* (pp 43-78). John Benjamins: Amsterdam.

Rothstein, S. (2004). *Structuring events: A study in the semantics of aspect*. Vol. 5. Oxford : John Wiley & Sons.

Sánchez, M. (1995). *Clasificación y análisis de préstamos del inglés en la prensa de España y México*. Lewiston/Quennston/Lampete: Mellen University Press.

Sankoff, D. & Poplack, S. (1981). A formal grammar for code-switching. *Research on Language & Social Interaction*, *14*(1). 3-45.

Sankoff, G., Thibault, P., Nagy, N., Blondeau, H., Fonollosa, M.O. & Gagnon, L. (1997). Variation in the use of discourse markers in a language contact situation. *Language Variation and Change*, *9(*2), 191-217.

Sankoff, D., Poplack, S. & Vanniarajan, S. (1990). The case of the nonce loan in Tamil. *Language Variation and Change* 2(1), 71-101.

Schiffrin, D. (1987). *Discourse markers*. Cambridge [Cambridgeshire]; New York: Cambridge University Press.

Schiffrin, D. (2003). Discourse markers: language, meaning and context. *The Handbook of Discourse Analysis. Oxford: Blackwell*. 54-75.

Schiffrin, D. (2006). *In other words: variation in reference and narrative*. Cambridge: Cambridge University Press.

Schilling-Estes, N. (2013). *Sociolinguistic fieldwork*. New York: Cambridge University Press.

Schilling-Estes, N. (2008). Stylistic variation and the sociolinguistic interview: A reconsideration. In

   Monroy Casas, R. & Sánchez Pérez, A. (Coords.), *25 años de lingüística en España [Recurso*

   electrónico]: hitos y retos/ 25 years of applied linguistics in Spain: Milestones and challenges (pp.

   971-986). Universidad de Murcia.

Schilling-Estes, N. (2007). Sociolinguistic variation. In Bayley, R., & Lucas, C. (Eds.), *Theories, methods,*

   *and applications. (*pp. 163-189*).* Cambridge: Cambridge University Press.

Schmidt, A. (2014). *Between the languages: Code-switching in bilingual communication.* Hamburg:

   Anchor Academic Publishing.

Schreier, D. (2013). Collecting ethnographic and sociolinguistic data. In Manfred. K. & Schlüter, J.

   (Eds.), *Research methods in language variation and change (Studies in English Language Series)*

   (pp. 1-23). Cambridge: Cambridge University Press.

Schwenter, S. A., & Torres Cacoullos, R. (2014). Competing constraints on the variable placement of

   direct object clitics in Mexico City Spanish. *Revista Española de Lingüística Aplicada*, *27*(2), 514-

   536.

Schwenter, S., & Torres Cacoullos, R. (2009). Variation in Spanish clitic placement: Constructional and

   pragmatic effects. Paper presented at *The 39th Linguistic Symposium on Romance Languages*

   (LSRL). University of Arizona.

Sharma, D., & Deo, A. (2010). A new methodology for the study of aspect in contact. In Walker, J.A.

   (Ed.), *Aspect in grammatical variation*, (pp 111-130). Amsterdam/Philadelphia: John Benjamins

   Publishing.

Shoman, A. (1995). *Thirteen chapters in Belizean history*. Belize City: Angelus Press.

Silva-Corvalán, C. (2001) *Sociolingüística y pragmática del español*. Washington, D.C.: Georgetown

   University Press.

-- (1994) *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon.

-- (1989). *Sociolingüística: teoría y análisis*. Madrid: Editorial Alhambra.

-- (1983). On the interaction of word order and Intonation: Some OV constructions in Spanish. In Klein-
Andreu, F. (Ed.), *Discourse perspectives on syntax* (pp. 117-139). New York: Academic Press.

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

Smead, R.N. (2000). On the assignment of gender to Chicano Anglicisms: Processes and results.
Bilingual Review/La Revista Bilingüe, *25*(3), 277-297.

Smith, C. S. (1991). The parameter of aspect (Studies in Linguistics and Philosophy, 43). Dordrecht:
Kluwer.

Smith, H. L. (2013). Habitual aspect marking in Palenquero: Variation in present temporal reference. In
Carvahlo, E. and Beaudrie, S. (Eds.), *Selected proceedings of the 6th Workshop on Spanish
Sociolinguistics* (pp. 97-108). Somerville, MA: Cascadilla Proceedings Project.

Solé, Y. (1966). *Hacer: Verbo funcional y lexical*. Washington, DC: Georgetown University Press.

Sportiche, D. (1999). Subject clitics in French and romance complex inversion and clitic doubling. In
Johnson, K. & Roberts, I.G. (Eds.), Beyond *principles and parameters*: Essays in memory of
Osvaldo Jaeggli (pp. 189-221). Netherlands : Springer.

Stabler, E.P. & MacSwan, J. (2014). A minimalist-parsing model for codeswitching. In MacSwan, J.
(Ed.), *Grammatical theory and bilingual codeswitching* (pp. 257-282). New York: Routledge.

Stammers, J.R. & Deuchar, M. (2012). Testing the nonce borrowing hypothesis: Counter-evidence from
English-origin verbs in Welsh. *Bilingualism: Language and Cognition*, *15*(3), 630-643.

Whitley, M. S. (2002). *Spanish/English contrasts: A course in Spanish linguistics*. Washington:
Georgetown University Press.

Statistical Institute of Belize (SIB). (2013). *Belize population and housing: Census Country Report 2010*. Belmopan, Belize: Statistical Institute of Belize.

Stell, G. & Yakpo, K. (2015). Elusive or self-evident? Looking for common ground in approaches to code-switching. In Stell, G. & Yakpo, K. (Eds.), *Code-switching between structural and sociolinguistic perspectives* (pp. 1-18). Berlin: Walter de Gruyter.

Suárez Molina, V. (1996). Voces mayas y mayismos en el español de Yucatán. Retrieved from http://www.mayas.uady.mx/articulos/voces.html.

Tagliamonte, S. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Malden, MA: Wiley-Blackwell.

Tagliamonte, S. A. (2006). *Analysing sociolinguistic variation*. Cambridge University Press.

Tenny, C. (1992). The aspectual interface hypothesis. Lexical matters. In Sag, I. A. & Szabolcsi, A. (Eds.). Palo Alto, Stanford University.

Thomason, S.G. (2001). *Language contact*. Edinburgh: Edinburgh University Press.

Thomason, S. G. & Kaufman, T. (1988). *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.

Timm, L.A. (1975). Spanish-English code-switching-el-porque-y-how-not-to. *Romance Philology*, *28*(4), 473-482.

Toribio, J.A. (2001). On the emergence of bilingual codeswitching competence. *Bilingualism: Language and Cognition*, *49*(3), 203-231.

Toribio, A.J. & Bullock, B.E. (Eds.). (2012). Introduction. *The Cambridge handbook of linguistic code-switching*. New York: Cambridge University Press.

Toribio, J., Bullock, B. & Greaser, C. (2012). The bilingual compound verb [*hacer* + VE] in Texas

    Spanish: The value of corpus data. Paper presented at the *Hispanic Linguistics Symposium*,

    Gainesville, FL.

Torres, L. (2002). Bilingual discourse markers in Puerto Rican Spanish. *Language in Society*, *31*(01), 65-

    83.

Torres Cacoullos, R. (2011). Variation and grammaticalization. In Hualde, J. I., Olarrera, A. & O'Rourke,

    E. (Eds.), *The Handbook of Hispanic linguistics* (pp. 148-167). Oxford: Wiley- Blackwell

Torres Cacoullos, R. (1999). Construction frequency and reductive change: Diachronic and register

    variation in Spanish clitic climbing. *Language Variation and Change*, *11*(02), 143-170.

Torres Cacoullous, R. & Aaron, J. (2003). Bare English-origin nouns in Spanish: Rates, constraints, and

    discourse functions. *Language Variation and Change, 15*, 289-328.

Torres Cacoullos, R. & Travis, C.E. (2014). Introduction. Gauging convergence on the ground:

    Codeswitching in the community. *International Journal of Bilingualism*, Advance online publication.

Torres Cacoullos, R. T., & Walker, J. A. (2009a). The present of the English future: Grammatical

    variation and collocations in discourse. *Language*, *85*(2), 321-354.

Torres Cacoullos, R., & Walker, J. A. (2009b). On the persistence of grammar in discourse formulas: A

    variationist study of that. *Linguistics*, *47*(1), 1-43.

Travis, C.E. & Torres Cacoullos, R. (2013). Making voices count: Corpus compilation in bilingual

    communities. *Australian Journal of Linguistics*, *33*(2), 170-194.

Treffers-Daller, J. (2001). Contrastive sociolinguistics: Borrowed and code switched past participles in

    Romance-Germanic language contact. In Jacobson, R. (Ed.), (2001). *Codeswitching worldwide II.*

    Trends in linguistic studies and monographs, *126.* (pp. 75-90). Berlin: Walter de Gruyter.

Treffers-Daller, J. (1994). *Mixing two languages: French-Dutch contact in a comparative perspective.* Berlin/New York: Walter de Gruyter.

Treffers-Daller, J. (1991). Towards a uniform approach to codeswitching and borrowing. In ESF network on codeswitching and language contact (Ed.), *Papers for the workshop on constraints, conditions and model* (pp. 259-279). European Science Foundation: Strasbourg.

Turpin, D. (1998). 'Le français, c'est le last frontier': The status of English-origin nouns in Acadian French. *International Journal of Bilingualism*, 2(2), 221-233.

Udz, S. (2012). *Improving Kriol language attitude and English accuracy: An exploratory study* Unpublished Doctoral Dissertation. Nova South Eastern University.

Valenzuela, E., Bruhn de Garavito, J., Barski, E., De Luna Villalón, M.E., Faure, A., Pangtay, Y., Ramírez Trujillo, A. & Reis, S. (2009). What code-mixed DPs can tell us about gender. Poster presented at *Mind-Context Divide Workshop*, University of Iowa.

Vendler, Z. (1959). *Linguistics in philosophy.* Ithaca: Cornell University Press.

Verkuyl, H. (1972). *On the compositional nature of aspects.* (Vol. 15)*, Foundations of language supplementary series.* Dordrecht: Reidel.

Walker, J.A. (2010a). Affairs of state: Defining and coding stativity in English and English-based Creole. In Walker, J.A. (Ed.), *Aspect in grammatical variation* (pp. 95-110). Amsterdam/Philadelphia: John Benjamins Publishing.

Walker, J.A. (Ed.). (2010b). *Aspect in grammatical variation*. Amsterdam/Philadelphia: John Benjamins Publishing.

Wei, L. (1994). *Three generations, two languages, one family: Language choice and language shift in a Chinese community in Britain*. Clevedon: Multilingual Matters.

Weinreich, U., William L. & Herzog, M. (1968). Empirical foundations for a theory of language change. In

    Lehmann, W. P. & Malkiel, Y. (Eds.), *Directions for historical linguistics* (pp. 97-195). Austin:

    University of Texas Press.

Wentz, J. & McClure, E. (1976). Ellipsis in bilingual discourse. In *Papers from the regional meeting of the*

    Chicago Linguistic Society, 12, 656-665.

White, L., Valenzuela, E., Kozlowska-MacGregor, M. & Leung, Y. (2004). Gender and number

    agreement in nonnative Spanish. *Applied Psycholinguistics, 25,* 105-133.

Whitley, M. S. (2002). *Spanish/English contrasts: A course in Spanish linguistics*. Washington:

    Georgetown University Press.

Wichmann, S. & Wohlgemuth, J. (2008). Loan verbs in a typological perspective. In Stolz, T., Bakker,

    D., & Salas Palomo. R. (Eds.), *Aspects of language contact* (pp. 89-121). Berlin/New York: Mouton

    de Gruyter.

Wichmann, S. & Wohlgemuth, J. (2005). Loan verbs in a typological perspective. In Stolz, T., Salas

    Palomo, R. & Bakker, D. (Eds.), *Proceedings of "Romanisation worldwide"*. Bremen.

Wilk, R. & M. Chapin. 1990. Ethnic Minorities in Belize: Mopan, Kekchi, and Garifuna. Cambridge:

    SPEAR Reports 1. Belize City, Belize: SPEAR.

Wilson, D. V., & Dumont, J. (2014). The emergent grammar of bilinguals: The Spanish verb *hacer* 'do,'

    with a bare English infinitive. *International Journal of Bilingualism*, Advance online publication.

Wilson, D. (2013). One construction, two source languages: Hacer with an English infinitive in bilingual

    discourse. In Carvalho, A. M. & Beaudrie, S. (Eds.), *Selected proceedings of the 6th Workshop on*

    Spanish Sociolinguistics (pp. 123-134). Somerville, MA: Cascadilla Proceedings Project.

Winford, D. (2009). On the unity of contact phenomena and their underlying mechanisms. In Isurin, L.,

    Winford, D. & De Bot, K. (Eds.), *Multidisciplinary approaches to code switching* (pp. 279-305).

    Amsterdam/Philadelphia: John Benjamins Publishing.

Winford, D. (2003). *An introduction to contact linguistics*. Malden, Mass: Blackwell Publishing.

Woods, S. (1991). Belize Creole: Prevalent crippling attitudes. *SPEAReports 6: Third annual studies on Belize conference*. Belize City, Belize: Society for the Promotion of Education and Research [SPEAR]. 147–158.

Woolford, E. (1983). Bilingual code-switching and syntactic theory. *Linguistic Inquiry, 14*(3), 520-536.

Yakpo, K. (2009). Complexity revised: Pichi (Equatorial Guinea) and Spanish in contact. In Faraclas, N. G. & Thomas, K. (Eds.), *Simplicity and complexity in creoles and pidgins* (pp. 183-215). London: Battlebridge.

Young, C.N. (1995). *Language and education in Belize*. Belize City, Belize: National Printers.

Young, C.N. (1973). *Belize Creole: A study of the creolized English spoken in the city of Belize, in its cultural and social setting.* Unpublished Doctoral dissertation. University of York.

Young, R., & Bayley, R. (1996). VARBRUL analysis for second language acquisition research. In Bayley, R., & Preston, D. R. (Eds.). *Second language acquisition and linguistic variation* (pp *253-306*). Amsterdam: John Benjamins Publishing.