Article

Determiner Asymmetry in Mixed Nominal Constructions: The Role of Grammatical Factors in Data from Miami and Nicaragua

Jeffrey Blokzijl [1],*, Margaret Deuchar [2] and M. Carmen Parafita Couto [1] (iD)

[1] Leiden University Center for Linguistics, Leiden University,
Van Wijkplaats 3, 2311 BX Leiden, The Netherlands; m.parafita.couto@hum.leidenuniv.nl

[2] Department of Theoretical and Applied Linguistics, University of Cambridge,
9 West Road, CB3 9DP Cambridge, UK; md118@cam.ac.uk

* Correspondence: jefblok15@gmail.com; Tel.: +31-615-484-221

Abstract: This paper focuses on the factors influencing the language of determiners in nominal constructions in two sets of bilingual data: Spanish/English from Miami and Spanish/English creole from Nicaragua. Previous studies (Liceras et al. 2008; Moro Quintanilla 2014) have argued that Spanish determiners are preferred in mixed nominal constructions because of their grammaticised nature. However, those studies did not take the matrix language into account, even though Herring et al. (2010) found that the language of the determiner matched the matrix language. Therefore, we hypothesise that the matrix language is the main influence on the language of the determiner in both mixed and unmixed nominal constructions. The results are consistent with our hypothesis that the matrix language of the clause provides the language of the determiner in mixed and unmixed Determiner Phrases (DPs). Once the matrix language is controlled for, the Miami data show a greater tendency for Spanish determiners to appear in mixed DPs than English determiners. However, in the Nicaragua data, we found only mixed DPs with an English creole determiner. This suggests that bilingual communities do not always follow the same pattern, and that social rather than grammatical factors may be at play. We conclude that while the language of the determiner is influenced by clause-internal structure, that of its noun complement and the matrix language itself depends on extralinguistic considerations.

Keywords: code-switching; matrix language; determiner-phrases; Spanish; English; Nicaraguan Creole English

1. Introduction

Since the 1980s, code-switching, "an activity which may be observed in the speech (or writing) of bilinguals who go back and forth between their two languages in the same conversation" [1], has been the focus of intensive study and debate. This linguistic phenomenon is not uncommon and can be found in various bilingual contexts [2]. Previous data have shown that individual utterances can combine elements from more than one language [3,4]. To date, the Spanish/English language pair is one of the most frequently examined, possibly because of the large number of speakers of both languages and the availability of collected data, such as can be found at the BangorTalk website [5]. We shall use the Spanish/English language pair to illustrate the range of possible combinations involving English and Spanish determiners and nouns. Examples (1a) and (1b) show Determiner Phrases (DPs) where the determiner and noun come from the same language, while examples (2a) and

(2b) illustrate mixed DPs where the determiner and noun are in different languages. Spanish words are shown in italics below, and determiners in both languages are shown in bold font.

1.　　English unmixed DP
　　　a.　　**The**　　　　　house
　　　　　　DET^1.DEF　　　N

　　　Spanish unmixed DP
　　　b.　　*La*　　　　　　*casa*
　　　　　　DET.DEF.F.S　　N.F.S

2.　　Mixed DP
　　　a.　　*La*　　　　　　house
　　　　　　DET.DEF.F.S　　N

　　　b.　　**The**　　　　　casa
　　　　　　DET.DEF.　　　　N.F.S

It has been reported previously that among mixed DPs, type (2a) occurs more frequently than type (2b), or in other words, Spanish determiners occur more frequently in mixed DPs than English determiners. For example, Liceras et al. reported, from their review of research on mixed Spanish–English DPs in spontaneous adult speech and their own study of child speech, that mixed DPs with Spanish determiners are far more frequent than with English determiners [6]. In their own study of child speech, only about 5% of the mixed DPs had English determiners; in adult speech, Jake et al. found 161 instances of Spanish determiners followed by English nouns, but no examples of English determiners followed by Spanish nouns [7]. However, Liceras et al. [6] do not provide information about the morphosyntactic frame in which the mixed DPs appeared, which Herring et al. [8] found to be relevant, as will be described below. Liceras et al. also do not consider the proportion of mixed vs. unmixed DPs with a given determiner, in case unmixed Spanish DPs should be more common than unmixed English DPs [6]. Instead, they explain the apparently greater frequency of Spanish determiners in mixed DPs in terms of the "intrinsic Gender feature of the Spanish Noun and the intrinsic Gender Agreement feature of the Spanish Determiner" [6] (p. 828), both of which features are absent in English. Moro Quintanilla also reports that Spanish determiners in mixed DPs are far more frequent in the Gibraltar data collected by Moyer than English determiners (only 2/243), and, like Liceras et al. [6], explains the distribution in terms of the "presence of an uninterpretable gender feature on the Spanish determiner, as opposed to its absence on the English determiner" [9] (p. 222). However, Moro Quintanilla also does not consider the morphosyntactic frame of the mixed DP or compare them with unmixed DPs [9]. Myers-Scotton and Jake also appear to concur with Liceras et al. [6], and Moro Quintanilla [9] on the assumption that the gender feature on Spanish determiners requires them to be 'elected' earlier in the language production process and that early election is related to greater frequency [10]. However, their earlier work had drawn attention to the importance of the morphosyntactic frame of the clause 'or matrix language' in influencing the language of the determiner.

The matrix language framework (MLF) was developed by Myers-Scotton [11] in order to account for common patterns found in intraclausal code-switching. Its main contribution is to capture a common asymmetry between the two languages involved, such that one provides the morphosyntactic frame or matrix language, and the other (the "embedded language"). The matrix language can be identified by the word order of the clause (the Morpheme Order Principle) and by the language source of particular "system morphemes" (the System Morpheme Principle). System morphemes are

---

1　DET = Determiner, DEF = Definite, F = Feminine, N = Noun, S = Singular.

categorized as either "early" or "late". Early system morphemes are "conceptually activated to express a part of speakers' meanings that they wish to communicate" [10] (p. 344) and include plural marking on nouns as well as determiners. Early system morphemes in a clause with code-switching can come from either the matrix language or the embedded language, but they are more likely to come from the matrix language. Late system morphemes have less semantic content than early system morphemes and a particular subcategory of late system morphemes, "outsider late system morphemes", can only come from the matrix language and are thus important in determining the matrix language of a given clause. Examples of outsider late system morphemes are case markers or verb inflections which encode subject–verb agreement.

We can illustrate the identification of the matrix language in examples (3) and (4)[2] below:

3.   my mom got   the *manguera*,
                            hosepipe
       'My mom got the hosepipe.'   [herring9: CLA][3]

4.   *eso fue*   *en el* front desk   *en el* reception
       that was   at the   at the
       'That was at the front desk, at the reception.'   [zeledon1: CAR]

Example (3) has an English matrix language or morphosyntactic frame on the basis of the finite verb *got* being English, whereas example (4) has a Spanish matrix language because the finite verb *fue* 'was' is Spanish (word order is not relevant here to distinguish between an English and a Spanish matrix language).

Returning to the issue of whether or not Spanish determiners occur more frequently in mixed DP constructions, Myers-Scotton and Jake argue for the influence of the matrix language (ML) [10] (p. 356) even though they had appeared to support the viewpoints of Liceras et al. [6], and Moro Quintanilla [9]. They state that "If Spanish is the ML in any CS corpus, then it is likely Spanish determiners will dominate for this reason alone under an analysis based on the MLF model" [10] (p. 356). This prediction had already been captured in the 'Bilingual NP Hypothesis' proposed by Jake et al. [7] and was motivated by the Uniform Structure Principle according to which the "structures of the matrix language are always preferred" [11] (p. 8).

Herring et al. attempted a preliminary evaluation of the influence of the matrix language on the determiner by using Welsh–English and Spanish–English data to assess the extent to which the matrix language matched the source language of the determiner in mixed DP constructions [8]. If we look again at examples (3) and (4) above, we can see that the language of the determiner in (3) is English, and thus matches the English matrix language of (3), while the language of the determiner in (4) is Spanish and thus matches the Spanish matrix language of (4). So, in both these two examples, the language of the determiner and the finite verb match.

In the small amount of the data analysed by Herring et al. [8], there was only one example out of 89 of a determiner (Spanish) and matrix language (English) mismatching. The matrix language of the clauses was Spanish in 90% of the cases, and the proportion of mixed DPs with a Spanish determiner found in those clauses was 91%, supporting the idea of a close relation between the language of the determiner and the matrix language of the clause. The distribution of the data also provides a possible explanation for the quantitative results reported by Liceras et al. [6] and Moro Quintanilla [9], i.e., the reason why the majority of mixed DPs appeared in clauses with Spanish as matrix language was that Spanish was the matrix language in the majority of cases. In other words, Spanish determiners

---

[2]   In examples (3) and (4), English words appear in normal type while Spanish words appear in italics. The mixed DPs
       are underlined.
[3]   [filename: speaker's pseudonym]

could have been preferred to English determiners in mixed nominal constructions simply because speakers selected a Spanish morphosyntactic frame, or matrix language in which they inserted their mixed DPs.

Recent experimental evidence provides support from two types of acceptability judgments for Herring et al.'s conclusion. To experimentally test these two sets of predictions regarding the language of the determiner in nominal constructions, Parafita Couto and Stadthagen-González tested two separate groups of 40 early Spanish–English bilinguals [12]. Their task was to evaluate the acceptability of sentences with code-switches between the determiner and the noun that reflected the predictions of the Minimalism Program, the MLF, both or none. The first group rated them on a Likert scale, while the second group performed a two-alternative forced-choice acceptability task (2AFC). Both experiments yielded converging evidence supporting Herring et al.'s [8] suggested preference for a match between the language of the determiner and the matrix language.

In the present study, we attempted to build on Herring et al.'s [8] work by investigating the link between the language of the determiner and the matrix language in a larger dataset than used previously. We focus on both mixed and unmixed nominal constructions in order to try to come closer to an empirically supported account of the regularities involved. Controlling for the matrix language, we measure the proportion of mixed DPs with each determiner as a proportion of the total number of DPs with the same determiner. Thus, we take into account the possibility that Spanish determiners might precede nouns more frequently than English determiners for internal linguistic reasons [13–19].[4]

Our data will come from two language pairs: Spanish–English from Miami, USA, and Spanish–English creole from the south Atlantic coast of Nicaragua. Although the language pairs in the two communities are similar, the differing distribution of matrix languages and determiners will allow us to consider the relative influence of linguistic and social factors on the code-switching patterns found.

2. Data for This Study

For our study, we used two bilingual corpora, one collected from conversations between Spanish–English speakers in Miami (FL, USA) [5], and the other from sociolinguistic interviews with Spanish–English creole speakers in various cities of the South Caribbean Coast Autonomous Region of Nicaragua. These two corpora have been chosen for comparative analysis because of the fact that English creole, also called Nicaraguan Creole English or Miskito Creole English [20,21], shares with English the absence of gender or number marking on its determiner, unlike Spanish, which the two corpora have in common. This means that if Liceras et al. [6] and Moro Quintanilla [9] are correct in assuming the overriding importance of grammatical features in influencing the appearance of Spanish vs. English determiners in mixed nominal constructions, then we would expect to find a significantly higher proportion of Spanish determiners in both corpora, regardless of the matrix language of the clause.

2.1. Miami Corpus

The Miami corpus [5] was collected in 2008 by Jon Herring and local assistants [22]. From the 1960s onwards, Miami has undergone an influx of Spanish speakers, resulting in intensive language contact between English and Spanish [23,24]. The first movement of Spanish-speaking immigrants were Cubans that sought to escape the Cuban revolution. The younger generation of Cuban immigrants became bilingual in English and Spanish. In the 1980s, there was a second influx of young immigrants from Central American countries that were suffering from civil wars. Nowadays, the Spanish speakers

---

4   Frequency counts also suggest that Spanish determiners are produced more often than English determiners (cf. for example, the rate of Spanish vs. English definite determiners per million words: Spanish 49,820.26 per million words vs. English definite determiners: 9999.99 per million words) [18,19].

in Miami are not only from Cuba or Central America but from a wide range of Latin American countries, and immigration continues. The corpus has 84 bilingual speakers of Spanish and English and provides a total time of 35 h of natural speech conversation. The data have been transcribed, glossed and coded. We analysed the entire dataset, yielding 8586 nominal constructions in 7115 clauses, with some clauses having more than one DP. However, because the Miami data are relatively large, we used an automatic analysis to codify the matrix language of the clauses and identify the nominal constructions as mixed or unmixed [25]. In order to test the automatic analyses, we took a sample of the data (10%) that we checked manually. From this sample, only 7% of clauses had a wrong matrix language assigned. In other words, we can safely conclude that the automatic analysis is reliable.

2.2. Nicaragua Corpus

The Nicaragua data contain sociolinguistic interviews from the South Caribbean Coast Autonomous Region of Nicaragua collected in 2006 by A. Koskinen [26]. This area was first colonised by the British. In fact, by 1630, the British dominated the total Atlantic area of Central America [21]. The British allowed the indigenous populations, the African slaves, and the refugees from Jamaica, to create their own state [20]. The result was creolisation of the English language that was also influenced by indigenous languages (Miskito, Rama, etc.) of the area. This English creole variety is now known as Nicaraguan Creole English (NCE). However, in 1860, this area became part of Nicaragua due to intervention from the United States [21,27]. From that moment onwards, the area became populated by Nicaraguans from other regions who brought the Spanish language with them. Spanish also became the official language [28]. Nowadays, all citizens of the cities in the South Caribbean Coast Autonomous Region (RACCS) are bilingual in both the creole language and Spanish, with both languages being taught at school. The Nicaragua corpus consists of a total of 16 h of recordings of 42 bilingual speakers being interviewed in creole at home, work or in school. The data used for this study consist of 3222 clauses and 3506 determiner phrases that were manually extracted and coded. Data from clauses to which a matrix language could not be assigned were excluded.

The main characteristics of the two corpora are summarised in Table 1 below.

Table 1. Summary of the main characteristics of the two corpora.

	Miami	Nicaragua
Languages	Spanish–English	Spanish–Nicaraguan Creole English
Type of data	Natural speech conversation	Sociolinguistic interviews
Number of speakers	84	42
Time	35 h	12 h
Clauses	7115	3222
Determiner Phrases	8586	3506

3. Analysis of Data

All clauses containing a determiner phrase were extracted and coded according to the matrix language of the clause in which they appeared. The automatic analysis of the Miami data included DPs with both definite and indefinite articles, while the manual analysis of the Nicaragua data also included demonstratives and possessives to compensate for the difference in corpus size. Because the word order of Spanish, English, and NCE are similar (Subject-Verb-Object), we used the language of the finite verb to determine the matrix language of the clause in our automatic analysis. Non-finite clauses were excluded. The nominal constructions were coded according to the language of the determiner, the language of the noun and whether or not the determiner and noun were in the same language ('unmixed') or not ('mixed'). This allowed us to study the proportion of mixed and unmixed DPs in clauses of each matrix language and to determine the extent of match between the language of the determiners and the finite verb. Examples (5) and (6) provide examples of an extracted mixed and unmixed DP respectively.

5.　　She was trying to be　　<u>a</u>　　　　　　<u>turista</u>
　　　　　　　　　　　　　　　DET.INDEF　　N.　　　　　[maria31: MAR]

In example (5), the underlined mixed DP consists of an English determiner and a Spanish noun, and the matrix language is English. In example (6), below the matrix language is also English but the DP is unmixed since the determiner and noun are both in English.

6.　　or are you still　　<u>a</u>　　　　<u>turist</u>　　　in your own city
　　　　　　　　　　　　DET.INDEF　　N.　　　　　　　　　　　　　　[maria31: MAR]

In example (7), the underlined mixed DP consists of an English creole determiner followed by a Spanish noun, and the matrix language is English creole. After the mixed DP, the speaker produces another DP that is unmixed.

7.　　an　　he　　uz　　to fight for　　ur　　　　luna,　　di　　　　moon
　　　　　　　　　　　　　　　　　　DET.POSS　　N.　　DET.DEF　　N.
　　　'and he used to fight for our moon, the moon.'　　　　　　　　[F-BLU-9-07]

Table 2 provides an example of our data coding.

Table 2. Example of data coding.

Clause	DP	Language of Determiner	Language of Noun	Type	Finite Verb	Language of Verb
she was trying to be *a turista*!	a *turista*	English	Spanish	mixed	was	English
or are you still a tourist in your own city state?	a tourist	English	English	unmixed	*are*	English
an he uz to fight for ur *luna*, di moon	ur luna	NCE	Spanish	mixed	uz	NCE
an he uz to fight for ur *luna*, di moon	di moon	NCE	NCE	unmixed	uz	NCE
she put it in the *boca*	the *boca*	English	Spanish	Mixed	put	English
no pero tal vez consigue un roommate	*un* roommate	Spanish	English	Mixed	*consigue*	Spanish
di king wife went to waz to diz *pila*	di king wife	NCE	NCE	unmixed	went	NCE
di king wife went to waz to diz *pila*	diz *pila*	NCE	Spanish	Mixed	went	NCE

DP: Determiner Phrase; NCE: Nicaraguan Creole English.

4. Results

The results of the Miami data analysis can be found in Table 3. The rows show mixed and unmixed DPs and the total number of DPs, while the middle columns indicate the frequency of the determiners matching vs. not matching the matrix language, with the results for Spanish and English as matrix languages given separately. As the Table shows, there is a match of 98.1% between the language of the determiner and the matrix language. Thus, the overwhelming majority of both unmixed and mixed DPs have a determiner with the same language as the finite verb of the clause.

Table 3. Results of Miami data analysis: mixed and unmixed DPs with matching and non-matching matrix language (ML).

	Language of Determiner	Matching ML	Not Matching ML	Total
Unmixed DP	Spanish	3579	94	8310 (96.8%)
	English	4574	63	
Mixed DP	Spanish	240	8	276 (3.2%)
	English	28	0	
Total		8421 (98.1%)	165 (1.9%)	8586 (100%)

On the other hand, 1.9% of all DPs have a determiner that does not match the matrix language. Still, of this group, 95.15% (157/165) are embedded language islands. These are all of the unmixed DPs which do not match the ML, as shown in the above table. An example of such an island is given in example (8). The clause in example (8) has an English matrix language, yet the determiner phrase *una pareja* 'a couple' has both the determiner and the noun in Spanish.

8.　　I hope mom doesn't think they're　　*una pareja*　　you know
　　　　　　　　　　　　　　　　　　　　　　a couple　　　　　　　　　　　　[sastre12: MAD]

In embedded language islands, the grammar of the Embedded Language temporarily prevails and so we expect its internal constituents to appear unaffected by the matrix language [10] (p. 139).

Of the mixed constructions, only 2.9% (8/276 DPs) did not match the matrix language. This is a very small number but we can note some similarities between those eight cases, of which three examples are given below.

9.　　*pero aquí* [*en el* north side]$_{AdvP}$ we don't ever get direct sun.
　　　　but here on the north side　　　　　　　　　　　　　　　　　　　[María1: MAR]

10.　　[*en los* dorms]$_{AdvP}$ they have a laundry room
　　　　in the dorms　　　　　　　　　　　　　　　　　　　　　　　　　　[Herring14: CON]

11.　　they did a sonogram blah blah　　*tumor en*　　[*en el* spleen]$_{AdvP}$.
　　　　blah

　　　　　　　　　　　　　　　　　　　tumor in　　　in the spleen　　　　[Zeledon8: MAR]

Examples (9)–(11) contain mixed DPs that appear in Spanish adverbial phrases introduced by the Spanish preposition *en*, 'on' in example (9) and 'in' in example (10). In the case of (9) and (10), the switch from a Spanish determiner to an English noun may have been anticipating the change of matrix language to English which occurs in the following clause (*we don't ever get direct sun* in (9) and *they have a laundry room* in (10)). All three of these examples could be characterised by what Muysken has called "alternational" switching [29], in which the switch occurs at a peripheral place in the clause. Adverbial phrases can be considered peripheral since they are not involved in the argument structure of the verb.

In addition to investigating the link between the language of the determiner and the matrix language, a second aim of our study was to measure the proportion of mixed DPs with each determiner as a fraction of the total number of DPs with the same determiner. We conducted this analysis on a subset of the data represented in Table 3, in particular the data shown in the column headed "Matching ML", where the determiner matched the ML. This was the case for 98.1% of the data as shown above. The results of this second analysis are shown in Table 4. As the Table shows, there is

indeed a higher proportion[5] (6.3%) of Spanish determiners followed by an English noun than English determiners followed by a Spanish noun (0.6%). Given the tendency of determiners to match the matrix language, this means that bilingual speakers are more likely to switch language after Spanish determiners than after English determiners.

Table 4. Results of Miami data analysis: proportion of mixed DPs.

	Unmixed DP	Mixed DP	Proportion of Mixed DPs
English determiner and English matrix language	4574	28	0.6%
Spanish determiner and Spanish matrix language	3579	240	6.3%
Total	8153	268	3.1%

Nicaragua Data

The results of the analysis of the Nicaragua data can be found in Table 5. As in Table 3, the rows show mixed and unmixed DPs and the total number of DPs, while the middle columns indicate the frequency of the determiners matching vs. not matching the matrix language. Next to each figure, we provide the percentage out of the total number of DPs. Table 5 shows that there is a match of 99.7% between the language of the determiner and the matrix language.

Table 5. Results of Nicaragua data analysis: mixed and unmixed DPs with matching and non-matching ML.

	Language of Determiner	Matching ML	Not Matching ML	Total
Unmixed DP	Spanish	9	9	3364 (96%)
	NCE	3346	0	
Mixed DP	Spanish	0	0	142 (4.0%)
	NCE	142	0	
Total		3497 (99.7%)	9 (0.3%)	3506 (100%)

The results of the Nicaragua data support the predictions of the MLF: only 0.3% of the DPs do not have a match between the language of the determiner and the matrix language of the clause. As in the case of the Miami data, the mismatched cases involve embedded language islands. An example of such an island is given in example (12). The clause in example (12) has an English creole matrix language, yet the DP *la escuela* 'the school' is entirely in Spanish. All the islands found were Spanish determiner phrases in a NCE matrix language clause.

12. di refreshment,   hav di *celebración*   *de la escuela*
       the   have the celebration   of the school
       'the refreshment, have the celebration in the school.'   [F-BLU-1-06]

All mixed constructions matched the matrix language.

Table 6 shows the numbers of unmixed and mixed DPs for each determiner and matrix language. As is clear, use of a Spanish matrix language is very rare in Nicaragua. However, a Fisher test ($p = 0.63$) suggests no significant difference between the proportion of mixed DPs with a Spanish determiner and with an NCE determiner.

---

5   The results of a chi square test showed that the difference is significant: $p < 0.01$.

Table 6. Results of Nicaragua data analysis: proportion of mixed DPs.

	Unmixed DP	Mixed DP	Proportion of Mixed DPs
NCE determiner and NCE matrix language	3346	142	4%
Spanish determiner and Spanish matrix language	9	0	0%
Total	3355	142	4%

5. Discussion

Our results suggest that speakers do not appear to have much choice regarding the language of the determiner: instead, this is influenced by the language of the morpho-syntactic frame or matrix language, and it is in selecting the matrix language that speakers do appear to have some choice. Once they have done this and have selected a matching determiner, the next option is whether or not to switch to a different language when selecting the noun following the determiner. We have noted that this happens more often where the matrix language (and determiner) is Spanish in the Miami data. In the Nicaragua data, however, we have only a small number of clauses with Spanish matrix language, and no statistical indication of a difference in the proportion of switched nouns following Spanish as opposed to NCE determiners. However, in trying to account for the asymmetry that we find in the Miami data, we may note that previous work by Bhatt on Indian data has suggested that the directionality of switches tends to be towards the language of power, or the language with superior social status [30]. Our findings seem consistent with this suggestion in that English has been the official language of Florida, the state where Miami is located, since 1988 [31]. So the more numerous[6] switches from Spanish determiners to English nouns than the reverse are in the direction of the official language. In Nicaragua, we can see that even though there is no significant difference between the proportion of mixed DPs with a Spanish determiner and with an NCE determiner, all the switches observed are from creole to Spanish. If this trend is confirmed in further studies, it would once again indicate switching in the direction of the language of higher prestige [28,30]. Koskinen reports that although the regional languages of the Caribbean coast including English creole were made official in 1993, creole was not used officially in education until 2007 [28]. Koskinen also reports that although the other regional languages have gained in status, creole "continues to be considered a form of 'broken English' or 'bad English'" [28] (p. 143). Spanish, on the other hand, is described as the "national language" [28] (p. 153) and is clearly superior in prestige.

Other explanations for the asymmetrical pattern of switching following determiners in the Miami data would require more exploration, but Fricke and Kootstra's work on the Miami data has established the importance of priming by material in the previous discourse, and this could be investigated in our data [32]. This account would be supported by the exposure-driven account posited by Valdés-Kroff [33], whereby bilingual speakers converge upon conventional production patterns. Such an emergent approach would offer an alternative as to how to account for asymmetrical structural distributions such as the ones we observed in our Miami and Nicaragua data. Another avenue to pursue would be the idea that code-switching tends to mark high information content as proposed by Myslin and Levy [34]. They consider words with high information content to be less predictable than those of lower information content, and to signal to the listener that special attention is needed. In relation to our data, we would need to examine whether there is evidence of the switches to nouns in the minority language having higher information content than those in the official language. Another variable that could be considered would be the language proficiency or dominance of the speaker. For example, Liceras et al. argued that it is possible to gain insights from the code-switching patterns and preferences which differentiate child and adult native speakers, simultaneous bilingual

---

6　Although we focused specifically on switches between the determiner and the noun, it is interesting to note that Fricke and Kootstra [32] (p. 11), using the same Miami corpus, also found fewer switches of any kind in bilingual clauses with English matrix language than with Spanish matrix language.

speakers and L2 speakers [35]. This, they say, could account for the conflicting evidence observed in the spontaneous switches produced in different communities of code-switchers.

One question that remains to be addressed is that of what determines the selection of the matrix language, since we have argued that the language of the determiner follows from this choice. We expect extralinguistic factors such as age of acquisition, language proficiency and the language of social networks to be all relevant, and hope to explore this question in the future.

6. Conclusions

The first objective of this study was to build on previous research that suggested that the language of determiners in mixed nominal constructions depends on the matrix language of the clause. The results confirm our hypothesis that the language of the determiner in mixed and unmixed nominal constructions generally does match the matrix language. The match between the language of the determiner and the matrix language seems to be unaffected by any grammaticised features in the language of the determiner.

The second objective was to compare the occurrence of mixed and unmixed DPs with English and Spanish determiners. We found that the frequency of switching from the determiner to the noun was asymmetric in the Miami data, being more frequent from Spanish to English in the Miami data. In the Nicaragua data, we only observed switches from the NCE to Spanish. We considered some explanations for our findings, and provisionally suggested that the relative prestige of the two languages may help to account for the asymmetry in the Miami data.

To summarise, we found that the matrix language was the most influential factor affecting the language of the determiner in mixed nominal constructions. However, extralinguistic factors seem to influence whether or not there is a switch after the determiner.

References

1. Deuchar, M. Code-switching. In *Encyclopedia of Applied Linguistics*; Chapelle, C., Ed.; Wiley: New York, NY, USA, 2012; pp. 657–664.
2. Inuwa, Y.N.; Althea, C.A.; Haryati, B.H. Factors Motivating Code Switching Within the Social Contact of Hausa Bilinguals. *IOSR J. Humanit. Soc. Sci. (IOSR-JHSS)* **2014**, *19*, 43–49. [CrossRef]
3. Poplack, S. Sometimes I'll start a sentence in Spanish y termino en español: Toward a typology of code-switching. *Linguistics* **1980**, *18*, 581–618. [CrossRef]
4. Timm, L.A. Spanish-English code-switching: El porqué y how-not-to. *Roman. Philol.* **1975**, *28*, 473–482.
5. Canolfan ESRC Centre for Research on Bilingualism. Corpus: MIAMI. Available online: www.bangortalk.org.uk/speakers.php?c=miami (accessed on 4 October 2017).
6. Liceras, J.; Fernández Fuertes, R.; Perales, S.; Pérez-Tattam, R.; Spradlin, K.T. Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional-lexical mixings. *Lingua* **2008**, *118*, 827–851. [CrossRef]
7. Jake, J.L.; Myers-Scotton, C.; Gross, S. Making a minimalist to code-switching work: Adding the matrix language. *Biling. Lang. Cognit.* **2002**, *5*, 69–91. [CrossRef]
8. Herring, J.; Deuchar, M.; Parafita Couto, M.C.; Moro Quintanilla, M. "I saw the madre": Evaluating predictions about codeswitched determiner-noun sequences using Spanish-English and Welsh-English data. *Int. J. Biling. Educ. Biling.* **2010**, *13*, 553–573. [CrossRef]

9.   Moro Quintanilla, M. The semantic interpretation and syntactic distribution of determiner phrases in Spanish-English Codeswitching. In *Grammatical Theory and Bilingual Codeswitching*; Macswan, J., Ed.; MIT Press: Cambridge, MA, USA, 2014; pp. 213–226.

10. Myers-Scotton, C.M.; Jake, J.L. Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level. *Int. J. Biling.* **2016**, *21*, 340–366. [CrossRef]

11. Myers-Scotton, C. *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*; Oxford University Press: Oxford, UK, 2002.

12. Parafita Couto, M.C.; Stadthagen-González, H. El book or the libro? Insights from acceptability judgments into determiner/noun code-switches. *Int. J. Biling.* **2017**. [CrossRef]

13. Chierchia, G. Reference to kinds across languages. *Nat. Lang. Semant.* **1998**, *6*, 339–405. [CrossRef]

14. Longobardi, G. Reference and proper names. *Linguist. Inq.* **1994**, *25*, 609–665.

15. Longobardi, G. How comparative is semantics? A unified parametric theory of bare nouns and proper names. *Nat. Lang. Semant.* **2001**, *9*, 335–369. [CrossRef]

16. Pérez-Leroux, A.T.; Schmitt, C.; Munn, A. Syntactic features and discourse factors in children's interpretation of definite determiners in inalienable possession. *Actes de l'ACL 2002/2002 Can. Linguist. Assoc. Proc.* **2003**, 245–255.

17. Vergnaud, J.-R.; Zubizarreta, M.L. The definite determiner in French and English. *Linguist. Inq.* **1992**, *23*, 595–652.

18. Brysbaert, M.; New, B. Moving beyond Kucera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. *Behav. Res. Methods* **2009**, *41*, 977–990. [CrossRef] [PubMed]

19. Cuetos, F.; Glez-Nosti, M.; Barbon, A.; Brysbaert, M. SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicologica* **2011**, *32*, 133–143.

20. Bartens, A. Nicaraguan Creole English. In *The Survey of Pidgin and Creole Languages*; Michaelis, S.M., Maurer, P., Haspelmath, M., Huber, M., Eds.; Oxford University Press: Oxford, UK, 2013; pp. 115–126.

21. Holm, J.A. *The Creole English of Nicaragua's Miskito Coast: Its Sociolinguistic History and a Comparative Study of Its Syntax and Lexicon*; University of London: London, UK, 1978.

22. Deuchar, M.; Davies, P.; Herring, J.; Parafita Couto, M.C.; Carter, D. Building Bilingual Corpora. In *Advances in the Study of Bilingualism*; Thomas, E., Mennen, I., Eds.; Multilingual Matters: Bristol, UK, 2014; pp. 93–110.

23. Carter, D.; Deuchar, M.; Davies, P.; Parafita Couto, M.C. A systematic comparison of factors affecting the choice of matrix language in three bilingual communities. *J. Lang. Contact* **2011**, *4*, 153–183. [CrossRef]

24. Gathercole, V. Miami and North Wales, so far and yet so near: Constructivist account of morpho-syntactic development in bilingual children. *Int. J. Biling. Educ. Biling.* **2007**, *10*, 224–247. [CrossRef]

25. Carter, D.; Broersma, M.; Donnelly, K.; Konopka, A. Presenting the Bangor Autoglosser and the Bangor Automated Clause-Splitter. *Digit. Scholarsh. Humanit.* **2017**. [CrossRef]

26. Koskinen, A. (Universidad de las Regiones Autónomas de la Costa Caribe Nicaraguense). Personal communication, 2006.

27. Zapata Webb, Y.H. Módulo 2: Una historia diferente. In *Manual de Educación Ciudadana Intercultural Y Autonómica, Universidad de las Regiones Autónomas de la Costa Caribe Nicaragüense*; Ford Foundation URACCAN: Managua, Nicaragua, 2012.

28. Koskinen, A. Kriol in Caribbean Nicaragua schools. In *Creoles in Education: An appraisal of Current Programs and Projects*; Bettina Léglise, M.I., Bartens, A., Eds.; John Benjamins: Amsterdam, The Netherlands, 2010; pp. 133–166.

29. Muysken, P. *Bilingual Speech: A Typology of Code-Mixing*; Cambridge University Press: Cambridge, UK, 2000.

30. Bhatt, R. Bilingual Situations in India: Power Relations between Languages Analysed through Code-Switching. Unpublished work, 2013.

31. Tatalovich, R. *Nativism Reborn? The Official English Language Movement and the American States*; The University of Kentucky Press: Lexington, KY, USA, 1995.

32. Fricke, M.; Kootstra, G.J. Primed codeswitching in spontaneous bilingual dialogue. *J. Mem. Lang.* **2016**, *91*, 181–201. [CrossRef]

33. Valdés-Kroff, J.R. Mixed NPs in Spanish-English bilingual speech: Using a corpus-based approach to inform models of sentence processing. In *Spanish-English Codeswitching in the Caribbean and the US*; Guzzardo Tamargo, R., Mazak, C., Parafita Couto, M.C., Eds.; John Benjamins: Amsterdam, The Netherlands, 2016; pp. 281–300.
34. Myslín, M.; Levy, R. Code-switching and predictability of meaning in discourse. *Language* **2015**, *91*, 871–905. [CrossRef]
35. Liceras, J.M.; Fernández Fuertes, R.; Klassen, R. Language dominance and language nativeness: The view from English-Spanish codeswitching. In *Spanish-English Codeswitching in the Caribbean and the US*; Guzzardo Tamargo, R., Mazak, C., Parafita Couto, M.C., Eds.; John Benjamins: Amsterdam, The Netherlands, 2016; pp. 107–138.