American Language Journal (3)2, pp. 8-15
www.americanlanguagejournal.com
Revised version received 6 September 2019

Mixing Spanish and English in the Noun Phrase

Dr. Daniel Smith

Spanish and English contact data from 56 speakers in an Hispanic community in Northeast Georgia are examined to see how speakers manage the two languages they use daily to structure their noun phrases. Noun phrases in utterances and sentences often consist of a single noun or a noun with a determiner (such as the articles, 'the, a, some, el, la, los, las, un, una, unas, unos'; possessives such as 'my, your, mi, mis, tu, tus, su, sus'; or demonstratives such as 'this, that, este, ese, esta, esa'; etc.). In Spanish and English code-switching, in which words from both Spanish and English are used in the same utterance, there are various possible combinations between the two languages. This study, based in part in research on the noun phrase in other bilingual contexts in Myers-Scotton and Jake (2017) and more specifically for Spanish and English in Blokzijl, Deuchar and Parafita Couto (2017), categorizes the code-switching patterns in noun phrases in the Hispanic community in Northeast Georgia. We compare our data with noun phrase code-switching in other bilingual contexts to observe universal grammar tendencies in noun phrase structure.

Keywords: noun phrase, determiner phrase, codeswitching, bilingualism, language contact

Introduction

Analysis of bilingual codeswitching, the use of morphemes from two languages in the same utterance, can help us understand universal grammar patterns with greater specificity than only the analysis of monolingual sentences. This calls for the investigation of the morphosyntax of specific kinds of bilingual phrases, such as the noun phrase, verb phrase, and prepositional phrase. The aim of this investigation is to note the patterns of the possible combination of elements in the noun phrases (NPs) which appear in Spanish English codeswitching. The patterns noted will be compared and analyzed from the standpoint of some other researchers' claims regarding universal tendencies in the bilingual NP. The data for this study are from spoken transcribed utterances of the community of Hispanics in Northeast Georgia, U.S.A.

Spanish English Contact in Northeast Georgia

A large community of Hispanics, who speak mostly Spanish, live and work in the English dominant society of Northeast Georgia. Children come into contact with large amounts of English in schools, but Spanish is dominant at home. Adults also use English to varying extents mostly in the workplace. Spoken language data from taped conversations between or with Hispanics was transcribed and categorized into utterances which were all Spanish, all English, codeswitching (CS) between Spanish and English between and within sentences, and instances of sentences with all morphemes from either Spanish or English with grammatical influence from the other language. Data from 56 Hispanic speakers were transcribed. Smith (2006, 2007, 2008, 2009) presents other aspects of Spanish / English contact with the same data corpus. The data for this study are all the noun phrases in which codeswitching occurs or at codeswitching boundaries in the speech of those 56 speakers.

Codeswitching and Borrowing

Some early researchers investigating CS, such as Reyes (1976; 1982) and Poplack (1981; 1982), distinguish and exclude borrowed (BR) forms from what they term CS. Thus, they largely exclude BR in a discussion of the structural constraints behind CS. BR is, however, considered along with CS in the discussion of structural constraints but is excluded from CS proper because of the different terminology used for each. Poplack and Sankoff (1988) partially distinguish between CS and BR. They characterize most borrowed items to be "at least morphologically and syntactically integrated into the host language" and in many cases "phonologically" as well (p. 1177). Poplack and Sankoff still leave room for possible confusion between CS and BR, however. Hence, 'nonce BR' is distinguished from complete BR by its use only by bilinguals and not monolinguals of the host language (p. 1176).

Myers-Scotton (1993a) suggests that BR forms may be a result of words introduced into a host language through CS after an indefinite period of time and frequency of use. She claims that CS forms may be less integrated into the host language than are BR forms but that this is "a difference in degree [of integration], not in kind." (pp. 182-183). Since this study supports Myers-Scotton's characterization of BR and its relation to CS, no attempt is made to distinguish BR from CS, especially since our study is not concerned with the differences between them, but rather with how they may be related in a dynamic "continuum" (pp. 176, 182) of language contact and change.

BR and CS are not distinguished in this study for the following reasons: 1. Myers-Scotton (1993a) sees BR and CS as only different in degree, not in kind, because BR enters the language first as CS, including English verbs which take on Spanish endings; 2. the data had no cultural BR forms, that is, those forms that did not have substitutes in Spanish; 3. the community is relatively new and any BR forms have only entered the language of the informants within the past two to three decades, either from English or from the Spanish of already partly assimilated Mexican-Americans.

Noun Phrase (NP) Codeswitching

NPs in utterances and sentences often consist of a single noun or a noun with a determiner (such as the articles, 'the, a, some, el, la, los, las, un, una, unas, unos'; possessives such as 'my, your, mi, mis, tu, tus, su, sus'; or demonstratives such as 'this, that, este, ese, esta, esa'; etc.). Noun phrases may also include adjectives of quantity or quality; the position of nouns and adjectives are often different in the two languages. In Spanish and English code-switching, in which words from both Spanish and English are used in the same utterance, the possible combinations between the two languages are multiplied. This study, based in part in research on the noun phrase in other bilingual

contexts in Myers-Scotton and Jake (2017) and more specifically for Spanish and English in Blokzijl, Deuchar and Parafita Couto (2017), will categorize the code-switching patterns in noun phrases in the Hispanic community in Northeast Georgia, specifically the language of the determiner of the NP as compared to the language of the noun (N). Comparing our data with NP code-switching in other bilingual contexts will highlight universal grammar rules or at least tendencies for how noun phrases can be structured.

In her 4-Morpheme (4-M) model Myers-Scotton (2002) categorizes noun phrase determiners as early system morphemes, with the claim that they are acquired earlier in language acquisition, unlike late system morphemes such as person and number markers on verbs which are acquired later in language acquisition. In the 4-M model the lexical portion of words are termed content such as com- in comer, minus the non-content system morphemses e-r. Early system morphemes only relate to the one content morpheme in its immediate vicinity. Late system morphemes relate more than one content morpheme, such as a subject and a verb, with implications for the entire clause. In bilingualism, according to her Matrix Language Frame (MLF) model, Myers-Scotton (2002) claims that words or phrases from one language can be embedded into a clause of the other language, the matrix language. The matrix language will contribute the person and number markers on verbs, late system morphemes. Early system morphemes can accompany the embedded content morphemes (which are typically the lexical or content morphemes of nouns, verbs, and adjectives), but the embedded language cannot contribute late system morphemes since those must come from the matrix language.

Blokzijl, Deuchar and Parafita Couto (2017) concluded that the matrix language of the clause in which the noun phrase is located is the most important indicator of the language of the determiner of the noun phrase. But they also comment that the matrix language of any given mixed clause and the language of the content noun in the noun phrase is due somewhat to social factors impacting the speech event.

Myers-Scotton (1993a, 2002) makes a very good case for the asymmetry of codeswitched utterances within clauses, with a matrix language controlling the grammatical structure of the sentence and an embedded language contributing mostly content morphemes but also some grammatical or function morphemes such as determiners, and gender, person, and number morphemes on nouns adjective, and verbs.

Myers-Scotton and Jake (2017: 356) explain that in Spanish, gender and number checking of determiners with their nouns must occur early in speech production and that this checking does not occur in English because English determiners do not agree in gender and number with their nouns. From this they conclude that when English is the matrix language of clauses in which Spanish nouns appear, that instead of English determiners elected to match the matrix language, instead, Spanish determiners are more likely to occur with Spanish nouns. While Myers-Scotton and Jake (2017) present data showing the strong tendency for Spanish matrix language clauses to favor Spanish determiners with embedded English nouns, no data is presented to show clauses in which English is the matrix language with embedded Spanish nouns.

The Data

The spoken utterances form the Northeast Georgia Hispanic community, as shown in Table 1, include mostly Spanish with no codeswitching or other influence from English. The second most common type utterance is English with no codeswitching or other influence from Spanish. After that, the next most common type of utterance is a single English word inserted into a Spanish matrix language sentence. Two other types of codeswitching, both intersentential (between sentences) and

multi-word intrasentential (inside sentences; more than one word of one language used in the sentence of the other language), were the next most common type of utterance. The least common type of utterance is a single Spanish word inserted into an English matrix language sentence.

Table 1. Types of utterances in the Northeast Georgia Hispanic community data

Type of utterance listed from most frequent (1) to least frequent (6)
1. all Spanish
2. all English
3. Spanish matrix language with single English word insertion
4. Intrasentential CS with more than one word inserted into a sentence of the other language
5. Intersentential CS (between sentences in the same utterance)
6. English matrix language with single Spanish word insertion

The data to be analyzed required isolation of all NPs in which the noun is a single word or morpheme inserted into a matrix language of the other language, as shown in Examples 1-5.

(1) ¿Quién quiere ser el baby? 'Who wants to be the baby?'
(2) No, es tape. 'No, it is tape.'
(3) Ese, y es un la única movie que tienen del barrendero o tienen más? 'That one, and it is a the only movie that they have of the sweeper or do they have more?'
(4) Presentamos a los dos babies que van a cantar. 'We are introducing the two babies that are going to sing.'
(5) My favorito i, is a snake. 'My favorite is a snake.'

Also isolated were NPs in utterances of intrasentential CS in which the noun or the determiner was uttered immediately before or after a morpheme of the other language, even if the noun was in a phrase with another word of the same language such as an adjective. Examples 6-9 show instances of NPs in this type of intrasentential CS.

(6) No copias this balloon. 'You are not copying this balloon.'
(7) Y este es rocks here. 'And this is rocks here.'
(8) Escoge todos los yellow ones. 'Choose/Pick all the yellow ones.'
(9) Está feo el new one. 'The new one is ugly.'

Table 2 shows the different phrase structures of the NPs isolated in the data according to which language is used for the N, the determiner (Det), and rarely but occasionally other associated elements, including Adjective, Numeral, Adverb, and Verb Participle, used within the NP. The numbers in Table 2 are the number of instances of that type of NP found in the Northeast Georgia Hispanic community data.

Table 2. Noun Phrases inside sentences with Spanish/English CS

NPs with Spanish N	number of instances	
SpanN	4	
SpanN+EngAjd	1	
EngDet+**SpanN**	2	
SpanDet+**SpanN**	1	in CS sentences in which the N or Det is next to a morpheme of the other language
EngDet+SpanVPart+SpanDet+**SpanN**	1	
SpanDet+**SpanN**+EngAdv+EngAdj	1	
	Total-10	

NPs with English N	number of instances	
EngN	53	
SpanDet+**EngN**	74	
EngDet+**EngN**	8	in CS sentences in which the N or Det is next to a morpheme of the other language
SpanDet+SpanAdj+**EngN**	2	
SpanDet+SpanNumber+**EngN**	1	
SpanDet+EngAdj+**EngN**	4	
	Total-142	

Key: Eng=English, Span=Spanish, NP=noun phrase, N=noun, Det=determiner, Adj=adjective, Adv=adverb, CS=codeswitching

Analysis and Discussion

Single word English insertions into Spanish matrix language sentences are the most prevalent of all the types of CS in the data (Table 1), which is explained by the fact that Spanish is the predominant language used by the community, given that it is the first language (L1) of most of them. According to Myers-Scotton (1993b: 119), the single morpheme insertion CS type should require only a minimal amount of competence in a second language. Based on this claim, this kind of CS is exactly the one that one would expect to be the most prevalent in a speech community in which Spanish is by far the dominant language. The fact that Spanish is the ML for the majority of the data also explains why English word insertions into a Spanish ML are much more frequent than Spanish word insertions into an English ML. Single word insertions require less skill in one of the languages than other types of CS. Single Spanish words inserted into English MLs are very infrequent as compared to single English words inserted into Spanish MLs. Gal (1979) found a similar pattern in Hungarian/German bilingualism in which German, the language of greater prestige, was freely inserted, one word or morpheme at a time, into Hungarian MLs, but Hungarian words or morphemes were not found in similar proportions in German MLs. Gal's finding as well as ours indicate different incursion patterns depending on which language is the ML, whether it be the language of the dominant culture or that of the non-dominant culture. Our findings may be due to the social status that Spanish has as a home language and thus matrix or 'mother' form. English has not achieved such a status of frequency in the community. When it is used, it is less associated with the home culture and not perceived as a 'fit' base or matrix for sentences.

The very frequent appearance in the CS data of English NPs as compared to Spanish NPs is very noticeable in Table 2, showing a total of 142 English NPs to only 10 Spanish NPs. This of course is due to the prevalence of Spanish as the ML in which English is more frequently embedded. We also note that even though English NPs without determiners are very frequent, even more frequent are English NPs with Spanish determiners. Of all the inserted nouns in an ML of the other language, the most frequent pattern is a Spanish determiner with an English noun (SpanDet+EngN), with 79 instances of a Spanish determiner out of a total of 152 instances, those instances including NPs with no determiner from either language. There are only 8 instances of an English determiner used with an English noun (N) in the data in which an English NP appears immediately next to morphemes from Spanish. In the data there are only two other instances of an English determiner used at all, this time with a Spanish N, giving a total of only 11 uses of the English determiner out of a total of 152 NPs in the data. This is largely due to the fact that Spanish is the ML of the majority of the data.

In the 10 instances in which a Spanish N is inserted into a sentence also containing English, there are 3 instances of an English determiner used. Even though the total number of 10 is rather small to be making very strong claims about the frequency of the appearance of English determiners with Spanish nouns, 3 out of 10 indicates that it is not unusual at all. Given more Spanish NPs in CS with English, we could expect to see many more instances of EngDet+SpanN. Considering the claims of Myers-Scotton and Jake (2017), there is not sufficient evidence in our data to confirm that English determiners are not frequently used because they do not agree in gender and number with Spanish nouns. In fact, in the 10 CS utterances in which there is a Spanish NP, there are 3 uses of an English determiner with Spanish nouns, and only 2 instances of a Spanish determiner used. The other 5 instances out of the 10 have no determiner at all. Still taking into consideration that we are only talking about 10 instances total, if anything there is at least some evidence that gender and number agreement of the determiner with the noun is not as important as Myers-Scotton and Jake have claimed; and they did not present evidence to show that clauses in which English is the matrix language with embedded Spanish nouns show a tendency not to favor the use of English determiners.

Considering the claims of Blokzijl, Deuchar, and Parafita Couto (2017), that the language of the ML in which the NP is located is the most important indicator of the language of the determiner, the large number of English NPs embedded into Spanish MLs in our data support this assertion. However we have not yet considered the social factors that Blokzijl, Deuchar, & Parafita Couto suggest. That topic is outside the scope of this analysis but will be important to consider in a follow-up investigation considering the social factors of the Northeast Georgia Hispanic community data in association with NPs in CS utterances.

Myers-Scotton (2002) claims, according to the MLF and 4-M models, that early system morphemes, such as determiners in NPs, may come from the embedded language but that late system morphemes, such as person and number markers on verbs can only come from the matrix language. This assertion is made regarding all languages in all bilingual or multilingual contact situations, thus a universal claim regarding language in general. This claim is corroborated with the data presented here. The instances of CS in the data show examples of embedded language determiners, as in example (6) above, but no instances of embedded language late system morphemes.

Conclusion

The Northeast Georgia Hispanic community embedded NP data show a strong tendency to use the Spanish determiner in English nouns embedded into Spanish MLs. This favors the claim that the language of the ML favors use of the determiner from that language. The few instances in which English determiners are used when Spanish nouns are embedded into English MLs also support this

theory. The data do not completely disconfirm the assertion that Spanish determiners are favored since English determiners do not agree with Spanish nouns, but neither do the data support it. Further research is needed to discover any social factor associations with the NP patterns we have noted in this investigation. The Northeast Georgia data support a larger more general claim regarding universal grammar, that early system morphemes like NP determiners frequently come from the embedded language in bilingual clauses, unlike late system morpheme person and number markers on verbs which do not.

References

Blokzijl, J.; Deuchar, M. & Parafita Couto, M. C. 2017. "Determiner Asymmetry in Mixed Nominal Constructions: The Role of Grammatical Factors in Data from Miami and Nicaragua", *Languages,* 2(4), 20, https://doi.org/10.3390/languages2040020

Gal, S. 1979. *Language shift: Social determinants of linguistic change in bilingual Austria.* New York: Academic Press.

Myers-Scotton, C. 1993a. *Dueling Languages: Grammatical structure in codeswitching.* Oxford: Clarendon Press.

Myers-Scotton, C. 1993b. *Social motivations for codeswitching: Evidence from Africa*. Oxford: Clarendon Press.

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

Myers-Scotton, C. & Jake, J. 2017. "Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level", *Interntional Journal of Bilingualism*. 21(3), 340–366.

Poplack, S. 1981. Syntactic structure and social function of code-switching. In R. Duran (Ed.), *Latino language and communicative behavior* (pp. 169-184). Norwood, NJ: Ablex.

Poplack, S. 1982. "Sometimes I'll start a sentence in Spanish y termino en español": Toward a typology of code-switching. In J. Amastae & L. Elías-Olivares (Eds.), *Spanish in the United States: Sociolinguistic aspects* (pp. 230-263). Cambridge: Cambridge University Press.

Poplack, S., & D. Sankoff. 1988. Code-switching. In U. Ammon, N. Dittmar, & K. Mattheier (Eds.), *Sociolinguistics: An international handbook of the science of language and society*, vol. 2 (pp. 1174-1180). Berlin: Walter de Gruyter.

Reyes, R. 1976. Language mixing in Chicano bilingual speech. In J. D. Bowen & J. Ornstein (Eds.), *Studies in Southwest Spanish* (pp. 182-188). Rowley, Mass.: Newbury House.

Reyes, R. 1982. Language mixing in Chicano Spanish. In J. Amastae & L. Elías-Olivares (Eds.), *Spanish in the United States: Sociolinguistic aspects* (pp. 154-165). Cambridge: Cambridge University Press.

Smith, D. 2009. "Asymmetrical mixing patterns of Spanish and English in a new immigrant community", *Bilingual Research Journal*, 32(2), 188-206.

Smith, D. 2008. "Family, friends, and language preference influence on Spanish and English use in the U.S.", *Southern Journal of Linguistics* 32(2), 26-77.

Smith, D. 2007. "Spanish/English bilingual children in the Southeastern USA: Convergence and codeswitching", *The Bilingual Review* 23(3), 99-108.

Smith, D. 2006. "Thresholds leading to shift: Spanish/English codeswitching and convergence in Georgia, USA", *The International Journal of Bilingualism* 10(2), 207-240.