This paper reports on a preliminary quantitative study of syntactic constraints on code-switching within discourses in which no change in participants, setting or topic is evident. The goals of the study are to provide a syntactic description of the points at which switches from Spanish to English and English to Spanish are possible and to assess the extent to which actual behavior conforms to stereotypes represented by acceptability judgements. Use of the quantificational technique of accountable reporting of relative frequencies permits investigation of a wide range of questions, including (1) Do implicational relations exist among syntactic environments such that it can be predicted that a speaker for whom a switch in environment A is possible may also switch in environment B, but not vice-versa? and (2) Are the constraints on code-switching uniform throughout all populations and in all social situations or are there differences in frequencies and/or implicational hierarchies? The data consist of taped conversations of adults and adolescents collected by bilingual participant-observers in the San Joaquin Valley of California and in South-central Texas. Preliminary findings indicate that differences between speakers are variable rather than categorical but that implicational relationships do exist and that these are related to social phenomena. (Author/CLK)
SYNTACTIC CONSTRAINTS ON CODE-SWITCHING: A QUANTITATIVE STUDY OF SPANISH/ENGLISH

Carol W. Pfaff
University of Texas at Austin

Paper presented at Linguistic Society of America Annual Meeting, December, 1975
Introducion

In recent years general linguists as well as students of bilingualism have increasingly recognized the importance of studying code-switching. Such mixture of two or more languages within a single discourse, or commonly within a single sentence, has implications both for the theory of language change and for the general theory of communicative competence.

Previous literature on code-switching has focused primarily on the social function, however, little has been said about the syntactic constraints. An early investigation by Espinosa 1917 claimed that the use of regular English words and phrases in the Spanish of New Mexico had no fixed limits and could not follow regular laws. Although Lance 1975:143 gives a fairly short list of 14 switch points, he also suggests that there are no syntactic restrictions on where switching can occur.

More recently, it has been claimed that code-switching is not random. Gumperz and Hernandez, for instance, proposed specific examples of ungrammatical mixed Spanish/English sequences, implying that there is an underlying grammar of code-switching. Gingras 1974 and Timm 1975 have attempted to get at this grammar, using acceptability judgements of native speakers as data.

My purpose in this paper is to pursue the subject of syntactic constraints on code-switching by means of the direct study of code-switching performance. My remarks are based on preliminary results
of an ongoing attempt to apply the technique of accountable reporting to the analysis of code-switching. I will discuss several methodological and theoretical issues:

First, why and how should we study code-switching performance? Second, what can we expect to find? That is, what is the nature of the syntactic constraints?

Third, and finally, I will examine a particular constraint on verb switching and will suggest there may be functional explanations which account for how code-switching interacts with other rules of Spanish and English grammar.

All examples cited are from conversations collected by bilingual participant observers in California and Texas unless otherwise noted.

Before turning to the main issues it should be noted that there is some disagreement in the literature about which types of language mixing are to be classified as code-switching, which as borrowing and which as interference. Gumperz and Hernandez 1972 for example, exclude use of Spanish sentence introducers, connectors or tags as in (1) and (2) and certain Spanish nouns as in (3) and (4), claiming these are simply part of the speakers' English style which function as indicators of ethnic identification.

(1) Si, pero we didn't get together yesterday. SJV f2 'yes, but...'
(2) Y the town is quiet, y todo. SA m 'And...and all that'
(3) Boy did she like them, hombre. SJV m '...man'
(4) They're going to have a big pachanga. A f 23 '...party'

A similar argument could be made that the parallel use of English in Spanish as in (5) is not code-switching.

(5) Yeah, mi papá nos explicaba, you know. A f '...my father explained it to us...'

1
Gumperz and Hernandez otherwise identify code-switching as a type of borrowing. Other investigators, notably Gingras 1974, exclude instances of English nouns in otherwise Spanish sentences, claiming that these are not code-switching, but simply borrowing.

Examples (6)-(10) show that such nouns are used in a wide range of functions: as subjects, objects, predicate nominals, objects of prepositions; and can be modified by Spanish adverbs and adjectives.

(6) El flight EP f
(7) Deme toast EP m 'give me...'
(8) El fue accountant toda su vida. SA 'He was an accountant all his life'
(9) Saben que vivo en trailer. EP m 'They know that I live in a...'
(10) Allá tienen unos traditions muy bonitos. EP f 'they have some very pretty traditions there'

English adjectives which occur after Spanish adverbs much as muy as in (11) or as predicates in in (12) might also reasonably be considered as borrowings.

(11) Mi papá es muy protective. SA f 18 'my father is very...
(12) No estan free en muchas palabras. SA m 17 'they're not free in so many words'

Adjectives which occur as part of noun phrases which also contain a switched noun, as in (13), might more likely be considered as instances of true switching.

(13) va a estar el same price y todo. SA m 17 'it's going to be the same price and all that'

Switches illustrated in (14-19) at verb phrases, adverbial and prepositional phrases and clauses, and switches of entire sentences as in (20) are generally considered "true switching" by all investigators.
(14) El perro chewed him up. SA m 'the dog...

(15) Y todo los Mexicanos were riled up. SA m 28 'and all of the Mexicans...

(16) ¿Antes viene aqui o on the way back? EP f 'is she coming here before or...

(17) We're joining together para tener más poder A f '...to have more power'

(18) Me dijeron que estaba upstairs in the new store. A m 'they told me he was...

(19) No voy tanto como iba pero I still believe in it you know. SA m 'I don't go as much as I used to but...

(20) ¿Oye que está haciendo Jason? Is he walking around? A m 'listen, what is Jason doing...

These examples are by no means an exhaustive list of all the switch types noted in the literature nor attested in the present study, and a much finer classification is both possible and profitable. Table 1, however, gives an overview of the relative frequency of the types of switches illustrated in (1-20).

<table>
<thead>
<tr>
<th>NP</th>
<th>Adj</th>
<th>V</th>
<th>PP</th>
<th>AdvP</th>
<th>V+O</th>
<th>be+pred</th>
<th>misc</th>
<th>cls</th>
<th>S</th>
<th>filler</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-E(SA)</td>
<td>70</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>3-E(EP&amp;A)</td>
<td>55</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>0</td>
<td>3</td>
<td>27</td>
<td>4</td>
</tr>
<tr>
<td>3-S(EP&amp;A)</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>24</td>
<td>48</td>
<td>9</td>
</tr>
</tbody>
</table>

Table 1: Percent of Switch Types
"-" = less than one percent n = instances of code-switches

The first line of Table 1 represents switches which occurred in a series of interviews conducted primarily in Spanish. The second two lines represent Spanish to English and English to Spanish switches from more casual conversation amongs friends and family in El Paso and Austin.

These data are reported separately because they occur in distinct conversational contexts. Although I do not have the
time to go into this interesting topic in detail here, I want to mention that there appear to be several distinct styles of code-switching associated with different social situations and having somewhat different syntactic characteristics.

In the present corpus, three types have been identified.

**Type I**

The first type is typical of casual interaction between peers, close friends or regular associates when conversation centers around everyday topics.

Syntactically, Type I is characterized by high frequency of what can be called 'deep S' switches—that is switches which occur either at surface sentence breaks or at independent or dependent clause breaks. Conjunctions introducing clause switches may occur in either language, not necessarily in the same language as the clause they introduce. Switches of conjunction only, especially into Spanish, also occur in this discourse style. In addition, switches occur at adverbial phrases, often propositional phrases, primarily those indicating time and manner. One or two word lexical switches, principally of nouns, occur with much lower frequency than in Type II.

**Type II**

The second variety is typical of casual or more formal interactions which, for one of several reasons, seem to be mainly in Spanish. In the present corpus, this included conversation of members of a bilingual family, a meeting of a Chicano student organization, and, of course, the primarily Spanish interviews conducted in San Antonio.

Syntactically, Type II is characterized by a high frequency
of loan words, primarily nouns and noun compounds, but to a lesser extent also adjectives and verbs. Of the "deep S" switches, whole sentence switches are much more frequent than switches at dependent or independent clauses, and switching of conjunction only occurred only once in the present corpus of approximately 1,000 switches. All "deep S" switching, including whole sentence switching was infrequent in the basically monolingual Spanish interviews. Adverbial phrase switching occurs with even lower frequency than deep S switching in Type II, but sentence introducers and tags occur relatively freely.

**Type III**

Type III is street talk, the jargon of the bato loco. In the present data, this conversation is basically in Spanish, but switches frequently into English for single nouns, verbs and adjectives and set phrases.

Undoubtedly, future research will identify other code-switching varieties.

**Accountable reporting of Code-Switching**

As Table 1 indicates, the overall frequency of switching to English at noun phrases is very high, not surprisingly particularly so in the basically Spanish interviews. Whole sentence switches and switches to Spanish clauses occur quite frequently in casual conversation. Many individuals switched only at such junctures, and the overall frequency of the "interesting" types of code-switching was quite low.

Nevertheless, I believe the study of code-switching performance should not yet be abandoned, and I will now briefly discuss the
applicability of the technique of accountable reporting to this problem.

The technique of accountable reporting, first applied to sociolinguistic data by Labov, has yielded considerable insight into the constraints on variation in Black and Anglo English and has since been applied to the description and analysis of variation in a number of other monolingual and multilingual communities by Fishman, Wolfram, Fasold, Cedergren, Sankoff and Bickerton, to cite only a few.

Accountable reporting involves counting instances of realizations of variable linguistic features in relevant linguistic and social contexts. The essential task is to determine what and how to count. Labov 1972:82-3 has discussed three crucial issues in applying this technique. First, he notes, one must identify the total population of utterances in which the feature under consideration varies. Second, one must decide on the number of variants which can be reliably identified, and third, one must identify all the subcategories which would reasonably be relevant in determining the frequency with which the rule in question applies.

The first point, excluding invariant environments may not be applicable to the study of code-switching. Since code-switching is never obligatory, it may prove to be impossible to distinguish environments where there is no variation from categorical constraints against switching.

The answer to the second question of realizations, however, is fairly clear. There are two end points: no switch and complete switch to the other language, with intermediate realizations representing various degrees of syntactic, morphological and phonological adaptation of the switched sequence to the rules of
the other language. A simple example, switching from Spanish to English at a single noun is illustrated by the four possibilities in (21):

(21)a. el camion  b. la troca/el troque  c. el truck  d. the truck

The first possibility is no switch (21a), a complete switch is represented by (21d), intermediate values are the phonological adaptations in (21b) and the typical unadapted borrowing (21c) which switches between the determiner and the noun.

The possibilities become more numerous as more complex syntactic switches such as phrases and clauses are considered. For instance, for switches in more complex noun phrases as in (10)

(10) Alla tienen unos traditions muy bonitos EP f 'they have some very pretty traditions there'

there are 16 possible combinations of Spanish and English in the Spanish word order determiner, noun, adverb, adjective as in (10) and 16 more in the English word order determiner, adverb, adjective, noun, as in the gloss of (10). One would not expect word orders which are ungrammatical in both Spanish and English to occur, but these possibilities must not be rejected out of hand.

The answer to the third question, identifying the conditioning environments must emerge from the ongoing analysis as Labov points out. With respect to the syntactic environment, at least three factors are of interest. First, the constituency or syntactic structure of the switch point; second, whether the switched sequence is an idiom or set phrase, and third, whether the switch involves a syntactic difference between the two languages—that is, whether the word order, co-occurrence restrictions or optionally or obligatorily
expressed relations, etc. differ.

The identification of the syntactic constituent switched or the syntactic environment is not always immediately obvious. As is often the case in empirical research, the theoretical assumptions of the investigator play a major role in the interpretation of specific cases.

For example, one question which arises is whether deep or surface structure is the proper level of analysis. For example, in (22)

(22) los...los...uh...your muscles a veces react SA f 18 'the...the...sometimes'

how many switches are involved? and what is their constituency? Are there three switches: from Spanish to English between los and your, back to Spanish between muscles and a veces and again Spanish to English between veces and react? It seems perhaps more plausible to regard your muscles react as just one switch at some stage of the derivation and for the adverb a veces to be moved in by a later transformation. If this analysis is correct, then constraints on code-switching are not defined exclusively at the level of surface structure.

(22) bears on another important issue. A frequently asked question is whether speakers who code-switch have one grammar containing elements of both English and Spanish, or whether they have two separate grammars with rules for switching between them. The study of performance of code-switches such as (22) where the grammars of English and Spanish differ provides empirical evidence which permits us to decide between these alternatives.

Note that in (22), the co-occurrence restrictions between
the determiner and noun are not identical in Spanish and English. Spanish requires the definite article with body parts, while English requires a pronominal determiner, either your as in (22) or, more formally, one's. Although switching is normally fluent and frequently occurs between determiner and noun with no necessity for repeating the determiner, the fact that the speaker has repeated the determiner and changed the syntax to conform to English requirements, indicates that she has control of both grammatical systems, at least as far as this aspect of the systems is concerned.

Constraints on Verb Switching

A more obvious reason for studying code-switching performance is simply to seek empirical confirmation of proposed constraints. Timm 1975, using acceptability judgements of three speakers, has recently put forward five quite specific claims about constraints on code-switching within major constituents. Three of these, which pertain to verb switching are illustrated in (23-30), Timm’s examples:

First, Timm claims, switching cannot occur between pronominal subjects and objects and finite verbs as in (23) and (24):

(23) *yo went *I fui
(24) *mira him *him mira *she sees lo *lo she sees

Second, Timm claims that switching is prohibited between finite verbs and their infinitive complements, as in (25) and (26):

(25) *(they) want a venir *(they) want a come *quieren (to) come
(26) *(I'm) going a decidir *(I'm) going a decide
   *voy to decide *voy a decide

12
Third, she claims that switching between auxiliary and verb is "aberrant" unless the principal verbal element is a phonologically adapted English loan word. Thus, according to Timm, (27-29) are unacceptable and (30) is questionable:

(27) *I must esperar *debo wait
(28) *(he) has visto *ha seen
(29) *(I) was caminando *estaba walking
(30) *(he) was wachando = watching

Acceptability judgement tests such as Timm's seem to provide a manageable way of getting at constraints on code-switching since they permit the linguist maximal control of the syntactic environments involved. It must not be forgotten, however, that at best, acceptability judgements represent stereotypes of code-switching behavior and do not always correspond to actual use, as Rickford's 1974 study of the remote perfective been in Black English has shown.

In the present study, several counterexamples to Timm's constraints on verb phrase switching were found to occur, although with quite low frequency.

Spanish pronominal subject and object occur with an English verb in (31):

(31) tú lo underestima a Chito EP m 'you underestimate Chito'

Spanish finite verbs plus English infinitive complements are found in (32-35):

(32) y fui a cash su cheque SA 'I went to cash his check'
(33) no van a bring it up in the meeting A f 'they're not going to.'
(34) cuando lo comenza train. SA m 15 'when he started to train him
and that he wanted to jump her, that he wanted to rape.

Sentences (36-43) illustrate switches between various auxiliaries and main verbs. (41) and (42) are progressives with estar plus English participle:

(36) estaba *training* para pelear. SA m 15 'he was training to fight'

(37) ¿dónde estás *teaching*? SJV m 'where are you teaching?'

(38-40) contain ir or andar plus English present participle:

(38) si va a *going shopping*, vaya con Mickey EP f 'if you're going to go shopping, go with Mickey'

(39) porque vas *bicycle riding* SA 'because you're going...'

(40) y tambien, una vez allí *horseback riding* SA 'once we went horseback riding there'

Sentence (41) is in a sense the opposite of (36-40), a present perfect progressive with auxiliaries in English plus Spanish present participle:

(41) and ever since then I have been *speaking Spanish* SA '...speaking Spanish'

Incidentally, my native speaker informants, students at the University of Texas at Austin, find (41) possible, but untypical, suggesting it is a play on code-switching which would be in keeping with the content of the message. In general, it seems likely that constraints on code-switching from Spanish to English differ from those in the opposite direction.

Further examples of switching between Aux and Main Verb are found in (42) and (43). (42) has haber plus English past participle, (43) has the modal (or perhaps quasi-modal) deber plus English infinitive:
(42) yo creo que apenas se había washed out. SA
'I think it had just washed out'

(43) luego pos el companero debe de suck on that para que
salga la pozona y escupila. SA m 45
'and then, well, his partner should suck on that so
that the poison can come out and spit it out'

It is tempting to look for theoretical explanations for the
constraints on code-switching and, indeed, if we hope to carry
the study of code-switching beyond language particular observational
and descriptive adequacy, it is essential. I suggest that two previously
established theoretical approaches are particularly relevant to
code-switching. These are, first, contrastive analysis, and,
second, functionalism.

Contrastive analysis has long been familiar to students of
bilingualism and second language learning as a method of identifying
potential interference points, and has already figured in my analysis
of (22).

Functionalism basically suggests that nonsyntactic factors,
for example requirements of communication and sentence perception,
constrain syntax.

In the present case of verb switching, it is well known
that contrastive analysis shows one of the primary differences
between English and Spanish syntax is in the verb conjugation,
where Spanish requires agreement with the subject NP. Furthermore,
whereas the subject NP is required to be present in surface
structure of non-elliptical English sentences, surface expression
of the subject in Spanish is generally omitted unless emphatic,
as it is marked on the verb.
It seems plausible that switching to English verbs would be maximally acceptable if the verb is preceded by a conjugated auxiliary or if it is the infinitival complement of a conjugated Spanish verb, since the relevant agreement and tense information is given on the preceding verb. This accounts for the use of sentences such as (32-43), contrary to Timm's hypothesis.

If the sentence would require an inflected main verb, there is an alternative to complete switching, namely using an English verb stem with Spanish inflectional endings, as in (44) and (45):

(44) El venia y chequiaba a las siete. SA 'he came and checked at 7 o'clock'

(45) los que sainiaron. SA 'those who signed'

Such hispanicized English verb forms also occur in infinitival and participial forms as in (46) and (47):

(46) ya voy a cuitiar ya. SA 'I'm finally go. to quit'

(47) porque estaban chiriando SA 'because they were cheating'

Table 2 shows the incidence of morphologically adapted and unadapted English verb switching. Only those instances where the verb begins the switch are included.

<table>
<thead>
<tr>
<th></th>
<th>present participle</th>
<th>infinitive (verb stem)</th>
<th>inflected verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>unadapted English verb</td>
<td>15</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>morphologically and phonologically adapted English verb</td>
<td>2</td>
<td>11</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 2: English verbs in Spanish

Table 2 shows that the use of adapted English verbs is most frequent when the structure requires an inflected verb.
There is a second alternative which permits switching at English verbs. This construction, attested so far only in data from California, uses a conjugated form of Spanish hacer 'do' followed by an English infinitive as in (48) and (49):

(48) su hija hace teach allá en San Jose. SJV
   'his daughter teaches there in S.J.'

(49) bueno porque te hicieron beat up? SJV
   'well, why did they beat you up?'

University of Texas student informants report that sentences like (48) and (49) are not used in Texas code-switching dialects with which they are familiar. They may, of course, occur in the speech of other groups in Texas.

The hacer plus infinitive construction appears to differ from both Spanish and English structures. Parallel constructions all in English or all in Spanish are ungrammatical, as in (50) and (51):

(50) *why did they do beat you up  *why they did beat you up?
(51) *porque te hicieron catiar?

There is, however, a causitive construction with hacer plus infinitive in Spanish as in (52):

(52) me hizo estudiar 'he made me study'

Apparent counterexamples to my claim that unconjugated English verbs can occur only when preceded by a conjugated Spanish verb are found in my examples (31), cited earlier, and in (53) and (54):

(31) tú lo underestimate a Chito. EP f 'you underestimate Chito'

In (31), note that the subject pronoun tú occurs, although the subject
is not especially emphasized in this conversation and would not need to be expressed overtly according to the rules of Spanish. We might speculate that the surface occurrence of this subject is conditioned by the use of the unadapted and unconjugated English verb. That is, in functionalist terms, there is a surface structure constraint conditioned by the switching of the verb, such that the syntax adapts to express the subject overtly if it cannot for some reason be expressed on the verb. Perhaps the length and morphological complexity of underestimate make it unsuitable for adaptation to the Spanish pattern. The constraint on subject pronouns, incidentally, is already known to operate in exclusively Spanish sentences, where the subject is more frequently expressed if the conjugated verb form is ambiguous as to subject.

In sentence (53), tense and subject information is provided in the conjugated verbs of the conjoined clauses, caminaba and llegaba, so it would be redundant on the English verb transfer in any case.

(53) So caminaba de mi casa...al bus stop y luego transfer y luego de ahí al otro bus hasta Kelly Field y llegaba alla como a los once. SA f 50

'so I went out of the house to the bus stop and later transferred and later from there to the other bus to KF and arrived there about 11 o'clock'

In (54) I suggest that the verb improve is actually an infinitive, that in an earlier stage of the derivation it followed para (or perhaps to) parallel to comprar in the first clause.

(54) Me da más dinero pa comprar libros pa que no ande en strikes y todo eso, pos uh pos también improve la escuela. SAm17

'he gives me more money to buy books and so that I won't walk in strikes and all that, well also so I'll improve in school.'
Conclusions

To summarize, I have attempted to show that the study of code-switching performance is a fruitful approach to syntactic constraints on underlying competence. Not only does performance data provide essential empirical evidence to confirm or disconfirm hypotheses, I have suggested that even such phenomena as hesitations and repetitions can provide evidence of speakers' underlying competence. Further, attempting to deal with the naturally occurring complexities of performance suggests new hypotheses about the nature of syntactic constraints. I have suggested that functional considerations, possibly related to the perceptual strategies of hearers play a role in syntactic constraints within sentences.

As to the question of whether code-switching should be distinguished from borrowing and interference, I have suggested that it is advantageous at this point to study the interplay of all these language-contact phenomena simultaneously. While the distinctions should not be obscured, all these can be considered as alternative realizations of underlying meanings and can be tabulated and quantified as such, using the technique of accountable reporting.

To conclude, I believe that this approach is an extremely promising methodology for studying the stylistic, situational and possibly even geographical varieties of code-switching which appear to exist. Accountable reporting will yield not only information about general syntactic constraints on code-switching, but will provide syntactic correlates of social constraints.
FOOTNOTES

1 I am grateful to a number of colleagues and students at the University of Texas and the California State University at Fresno who have made their data and intuitions available to me. In particular, Professor R. Joe Campbell of UT San Antonio, Ana Huerta, José Galvan and Ernesto Zamora of UT Austin and Efrain Rodriguez and Teresa Trejo of CSUF.

2 In the examples, capital letters indicate the location of the interview or conversation: SJV = the San Joaquin Valley of California, SA = San Antonio, EP = El Paso and A = Austin, Texas. Small m and f are sex and following numbers represent speakers' ages.
REFERENCES


