

Linguistic Aspects of Spanish-English  
Language Switching

by

John M. Lipski

special studies no. 25



Published by

Center for Latin American Studies

ARIZONA STATE UNIVERSITY

## Acknowledgements

*This work would not have been possible without the constant support and patience of my wife, Beverly, and of my daughter, Ursula; this book is gratefully dedicated to them. Publication of this book was facilitated by a typing grant from the University of Houston.*

## Introduction

### SCOPE OF THIS STUDY

Language switching among Spanish-English bilinguals living in the United States is one of the most salient characteristics of this speech community, and has been the subject of comment and research by educators, psychologists, anthropologists, linguists and literary scholars. Originally, before critical focus was directed at this phenomenon, language switching--now known as code switching--was taken as evidence for mental confusion, the inability to separate the two languages sufficiently to warrant the designation of true bilingualism. With the advent of interest in sociolinguistic and ethnolinguistic investigations of non-prestige groups, code switching became the object of scientific scrutiny, giving rise to the not surprising discovery that such language mixture is governed by a complicated and as yet not fully delimited set of constraints, which indicate a complex interaction between the two languages. This is a far cry from the anarchical confusion postulated previously. At present, scholarly interest in bilingualism is high, and code switching, particularly between Spanish and English, is a constant topic of discussion. Many state school systems are under a mandate to provide bilingual education for non-English speaking minorities, and in the concomitant process of determining language standards for classroom use, the topic of language mixture continues to appear, often encumbered by a clutter of anecdotal, misleading and incomplete descriptions, reflecting ignorance, prejudice or disinterest. Despite this almost overwhelming diversity of motivations for directing attention to bilingual language switching, certain common denominators may be extracted, which promise both immediate and long-range dividends for research scholars and educators alike, and as such are worth comparing. It is the purpose of this monograph to offer an appraisal of the major research strategies affecting the linguistic study of Spanish-English code switching, which in turn have implications for the determination of the linguistic competence of bilinguals and the manner in which the two languages are represented in the cognitive apparatus. Particular emphasis will be placed on the potential benefits of work recently completed or currently in progress, both by the present writer and by

Concha  
Ampli

other investigators who have dedicated themselves to the study of bilingual language switching. The ultimate goal will be the construction of models which not only have generative capacity, that is which will produce acceptable language-switched utterances similar to those produced in spontaneous speech, but which also have significance for the actual representation of two languages in the cognitive mechanism of bilingual speakers.

Bilingual code shifting is one of the most striking bits of evidence which may be adduced in favor of the hypothesis that bilingual speakers, at least those who are sufficiently proficient in both languages to engage in spontaneous switching have, in addition to two possibly distinct grammars, a mechanism which fully integrates the two, to the point where it becomes more useful to speak in terms of a bilingual grammar. Precisely the recognition of an internal structure for bilingual language shifting militates in favor of such a systematic integration, since without a structured basis for language shifting, one is free to simply postulate random interference resulting from imperfectly learned systems. Although it is now nearly universally accepted by researchers that a principled and formalizable basis exists for code switching in most contexts, a demonstration of the exact nature of these constraints will be postponed until the following chapters. This procedure can be justified by the current state of research on language shifting, which departs from the premise of a structured bilingual interaction; moreover, it will be possible in this introduction to describe the major theoretical questions raised by bilingual language shifting, so that the remaining chapters may be devoted to a concrete exploration of the implications of Spanish-English bilingualism for the principal theoretical issues in the study of bilingual behavior.

#### TYPES OF CODE SWITCHING

A preliminary distinction which must be effected is between intersentential and intrasentential language switching, since the latter is responsible for raising more fundamental theoretical questions. Intersentential switching consists of shifting languages at sentence boundaries which are frequently principal discourse boundaries. Thus, for example, a bilingual speaker may shift languages as he directs himself successively to several other speakers, each of whom is perceived to have different linguistic preferences. Similarly, a language shift may be embodied in a direct quotation reflecting remarks originally made in a language other than that of the ongoing conversation. Although truly fluent and 'balanced' bilingual speakers also engage in intersentential switching, it is not necessary to be fluently bilingual to engage in this type of linguistic behavior; indeed, the most elementary student of a foreign language customarily utilizes this discourse mode, as do the teachers who impart foreign language instruction. A speaker who has learned a second language in the postadolescence period, although he may speak this language fluently, will rarely engage in spontaneous intrasentential shifting, but may shift languages intersententially the same as a native bilingual individual, thus highlighting an important qualitative difference between the two forms of bilingual language alternation.

Intrasentential language switching involves the shift from  $L_1$  to  $L_2$  in the middle of a sentence, often with no interruptions, hesitations, pauses, or other indications of a major categorical shift. An

*con bis automática*

example is provided by the title of Poplack's (1980) study: "sometimes I'll start a sentence in English y termino en español." This type of language shifting, which is characterized by a smooth flow between Spanish and English, is common in most United States Hispanic communities, and has even become an identifying characteristic of much of Hispanic-American literature in the United States. Intrasentential language shifting immediately impresses the observer by its smooth transitions between the languages, which are largely responsible for the early impressions of unprincipled confusion. It is also noteworthy that speakers who engage in intrasentential shifting are frequently unaware of having shifted until well after the fact and may not be able to recall the exact moment of the shift, or to give justification for having changed languages in midstream. As a consequence, the study of intrasentential language shifting will yield the greatest fruits in the way of characterizing the linguistic organization of the bilingual cognitive apparatus. In keeping with this tradition, the majority of the remarks in the succeeding chapters will be confined to intrasentential language shifting and language shifting at sentence boundaries in which the  $L_1-L_2$  transition is effected without pauses or other major discourse shifters.

#### TYPES OF BILINGUALISM

A second distinction which has become implanted in the literature on bilingual behavior is that between compound and coordinate bilinguals. Originally proposed by Ervin and Osgood (1954), this distinction has been the subject of much controversy, but still continues to reappear in studies of bilingualism. Initially, this concept was designed to account for the manner in which languages are learned and, by extrapolation, of the way in which they are organized in the mind of the bilingual. A coordinate bilingual, it was claimed, is one who has learned each of the two languages in completely different environments, such as home and school, and who maintains each language separate from the other in association with a single domain of usage. The compound bilingual has learned both languages simultaneously in similar or identical environments and makes no exclusive association of one language or the other with particular contexts. By extension from this distinction, it has been assumed that the coordinate bilingual has two noninterfering language systems, in which each surface stimulus has two separable linguistic associations. The compound bilingual is assumed to possess a single undifferentiated system in which each referent immediately triggers a dual linguistic response. In other words, in the compound bilingual, two sets of linguistic signs are associated with a single set of cognitive associations, whereas the coordinate bilingual maintains two separate systems of association. Kolors (1968) used the metaphor of a water tank, in which coordinate bilinguals possess two separate tanks, one for each language, while the compound bilingual has a single tank (the 'undifferentiated' lexicon) with two separate faucets. Diller (1974) has categorically rejected these notions, in view of the largely inconclusive experimental evidence in favor of the differentiation and the impossibility of definitively placing most individuals in one of the two categories, based either on developmental histories or on synchronic performance. Other investigators have suggested modifying the framework rather than rejecting it out of hand. Lambert (1969) reviewed the literature

up to that time and concluded that the distinction had some validity if each language was learned in an entirely separate cultural setting, as in the case of Anglophone and Francophone Canadian environments. Lambert also suggested that different types of language teaching methods could result in different types of (quasi-) bilingualism in successful students. Kirstein and de Vicenz (1974) utilized a transformational-generative model, in which it was postulated that a coordinate bilingual has completely unrelated grammars which are connected only at the deepest level of linguistic universals, whereas with compound bilinguals there is interconnection of the two grammars at all levels of reference. A concomitant prediction of the model is that spontaneous translation in the coordinate bilingual in some sense traverses a 'longer path' since it must reach through to the level of language universals. Experimental evidence makes such a claim difficult to sustain. Beardsmore (1974) and Genesee et al. (1978) chose to interpret the distinction so that coordinate bilinguals are those who have learned a second language considerably later than the first, perhaps after puberty and/or out of the family environment; whereas compound bilinguals are those brought up more or less bilingually. The latter authors went so far as to suggest that individuals who had become bilingual in infancy or childhood use primarily left hemisphere, possibly semantic linguistic strategies, and that individuals who became bilingual in adolescence or later tend to use right hemisphere, possibly 'gestalt-like' or melodic processing strategies. This may be a restatement of Lambert's (1969) remarks concerning language teaching and learning strategies, as well as being related to overall linguistic proficiency, which is frequently diminished when the second language is acquired in the postadolescence period. Herras and Nelson (1972) report experimental evidence that Spanish-English bilingual children retain semantic content of stimulus sentences but rapidly forget in which language the material occurred or what syntactic form it assumed. Penfield and Roberts (1959) give some evidence derived from their famous cortical stimulation experiments which suggests a justification for the compound-coordinate distinction, at least for French Canadians. A related model was proposed by Riegel (1968), who noted the influencing action of  $L_1$  when  $L_2$  is learned considerably later. Albert and Obler (1978), after considering most previous studies, concluded that in usual circumstances the distinction compound-coordinate is not empirically sustainable, although general tendencies may be grouped around these two theoretically polar opposites.

According to the compound-coordinate distinction as proposed and refined in the above-mentioned studies, intrasentential language shifting would be much more common in, if not exclusive to, compound bilinguals, since the two languages would be more closely integrated; whereas coordinate bilinguals would rarely engage in such behavior. It must be noted that many of the studies in which the compound-coordinate distinction has been proposed dealt with bilingual situations outside of the United States, and did not comment on the behavior of Spanish-English bilinguals in the United States. An examination of the latter situation, no matter how superficial, immediately calls into question both the validity of the compound-coordinate dichotomy and the consequences of this dichotomy for language switching, at least when restricted to this particular linguistic community. Bilingual language acquisition among Spanish-English bilinguals in the United States runs the entire gamut from completely mixed situations in which parents and other family members use both languages indiscriminately,

to clear separations of context, where Spanish is only spoken in relatively restricted domains. Despite this diversity, the majority of United States Hispanics engage in intrasentential language shifting, although it has been shown that the frequency of certain types of shifts is somewhat related to individual language proficiency and perhaps also to mode of acquisition.<sup>2</sup> At the same time, it is not really possible to speak of complete separation of contexts among United States Hispanics, since due to the nature of the bilingual communities, English penetrates, in one form or another, into nearly all aspects of family life; whereas the existence of numerous large and stable Spanish-speaking communities ensures that the bilingual speaker will be able to use Spanish in an increasingly large number of contexts outside the home.<sup>2</sup>

Observations of the linguistic behavior of United States Hispanics suggest that the compound-coordinate distinction may enjoy some validity if considered as a measure of long-term language usage, in the home and elsewhere, since subjects who habitually live in stable bilingual communities where Spanish and English are used and combined freely engage in more spontaneous intrasentential shifting than individuals who live outside such bilingual communities. To employ this definition of the dichotomy, however, is to come full circle, since language-switching ability is in effect being used as a definition of the type of bilingualism, and vice versa. Since there appears to be no way of avoiding total circularity, in the case of Spanish-English interfaces in the United States, it is more profitable to accept fluent language shifting as a *sui generis* phenomenon, which may ultimately yield dividends in explicating theoretical typologies of bilingual language acquisition and organization, but which cannot be extrapolated from previously given typologies. It is unlikely that any speaker who is part of a bilingual community (as opposed to isolated bilinguals, one of whose languages may only be spoken in the home or among a limited group of speakers), can truly maintain the two languages separated in a fashion suggested by the proposed coordinate bilingual category, regardless of the fashion in which the two languages were originally learned.<sup>3</sup> Consequently, it is more profitable to study language switching as a possible window of access to the organization of the two languages as part of an integrated system, whose degree of integration will vary according to numerous parameters, but whose very existence is postulated by the bilingual switching.

### ONE GRAMMAR OR TWO?

If bilingual language switching follows principalled guidelines, then the bilingual speaker must be able, consciously or unconsciously, to keep the languages apart. That is, there must exist a language tagging mechanism which unequivocally labels lexical elements, phonological structures and syntactic formations as part of one language or the other (or conceivably as part of both, in certain overlapping cases). At the same time, it is necessary to consider the existence of a 'switch' which will enable the shift from one language to the other, both in fluent intrasentential shifts and in more deliberate language shifts. Penfield and Roberts (1959) suggest the existence of such a switch and Macnamara and Kushnir (1971) utilize experimental evidence indicating a lag time in comprehension of bilingual utterances to postulate both an input and an output switch, operating sequentially

and independently. Similar response-time evidence from multilingual lists was given by Tulving and Colotla (1970). Albert and Obler (1978: 250) note, "We must assume that there is some way of identifying language units as 'belonging' to one or the other language . . . in some sense those syntactic structures that occur in only one of the two languages must be 'tagged' as pertaining to that language." These authors conclude that at present it is not possible to determine whether language tagging involves a special label associated with each particular item, by contiguous storing of similar items in separate storage registers, or by some other means. The very openness of the question suggests that the rapidly accumulating store of data on Spanish-English language switching may be able to shed light on these issues.

Evidence for the existence of separate grammars or more integrated and shared mechanisms has largely come from situations of forced code shifting in performance and of responses to code-switched stimuli. In the case of forced switching during reading passages, it has been found that production time is significantly increased in switched discourse, due presumably to the need to shift registers or production mechanisms, which may provide psycholinguistic evidence for a 'language switch' or channeling mechanism. Lenneberg (1967: 18) noted that the time for a regular language shift approaches the minimum threshold of speed for all language operations, and Haugen (1973: 550) observed that formulation in two languages increases the uncertainty of choice and the necessity of monitoring the language tags that identify the items and rules of each language. On the other hand, most studies have shown that perception of switched utterances (as opposed to lists of individual words) takes generally the same amount of time as for monolingual material, when subjects are bilingual. Thus there is evidence both for a language-switching mechanism, which channels productive forces, and of a shared mechanism at an abstract level, which does not differentiate between languages. Kagers (1966b) found that repeating a word in two languages had the same effect on perception as repeating it in the same language, for bilingual subjects. Interlingual word associations produced the same results, as a consequence of which the conclusion has been offered that the entire question of polar opposites in bilingual storage of language is probably invalid, and that bilingual competence must be characterized simultaneously by a shared component and a pair of separated components.

The evidence is far from clear, but all signs point to the principalled interaction of the two grammatical bases representing the languages of the bilingual speaker. The fact that forced switching requires additional production time is irrelevant to the degree of integration of the grammars, since it represents an artificial situation in which additional demands are placed upon the linguistic mechanisms of the speaker. More weight should be placed upon the observation, repeated for several different language pairs, that reception time of switched utterances is essentially the same as for monolingual material, which indicates that at some level of processing the two grammars are not being kept apart; but that both are subject to some central mechanism for extracting the information content. A detailed examination of the syntactic structures of code-switched utterances may therefore be valuable in tracing the course of this bilingual processing mechanism. It appears that many if not most bilingual speakers form their sentences first at some abstract semantic level. Haugen (1973: 530) suggests that such abstract formation may even occur before the speaker knows what language the sentence will come out in, while

Clyne (1972: 41) offers certain conditions in which a sentence may be originally formulated in one language but produced in another.

As a result of the important theoretical implications of intra-sentential code switching, the following chapters will present a detailed analysis of cases of bilingual language behavior among Hispanic bilinguals in the United States. Fundamental emphasis will be placed on the linguistic restrictions which govern such switching, including individual grammatical constraints, large-scale syntactic patterning and the variable or transitional nature of certain elements in the bilingually switched utterance. Several useful models of code switching will be discussed and refined as far as the data permit.

## NOTES

1. For example Poplack (1980, 1981); Zentella (1981), Valdés-Fallis (1978), Scotton (1979), Lambert (1969).
2. Cf. Sánchez (1983) for a discussion of similar possibilities.
3. Peñalosa (1980:46) observes that under normal circumstances coordinate bilingualism among United Hispanics tends to evolve into compound bilingualism, although the effects of the school system may arrest this tendency. See also Cárdenas (1972).

**Situational Factors in Language Switching****THE ROLE OF EXTRALINGUISTIC VARIABLES**

The study of structure in bilingual code switching begins with the search for factors which exercise definitive influences on the switching mechanism. These factors are divided into grammatical or linguistic and situational. Even though it is extremely difficult to incorporate situational variables into formal models, for reasons that will be detailed in succeeding chapters, it is vital to give at least an overview of the pragmatic ramifications of bilingual code switching, since the phenomenon itself is strictly subordinated to a unique set of sociolinguistic parameters. The largest body of research on Spanish-English code shifting has been directed toward the sociolinguistic dimension, both in conversational contexts and in literature, since this bilingual behavior is strongly identified with the ethnicity of United States Hispanics. In view of the impossibility of considering bilingual code switching as a purely linguistic phenomenon devoid of situational connotations, we shall briefly pass review of some major observations and theories, with the caveat that no definitive conclusions have yet been reached regarding the total significance of this mode of linguistic production. The reader interested in a detailed account of the socio-linguistic context of Spanish-English bilingualism in the United States is urged to consult the excellent studies by Peñalosa (1980) and Sánchez (1983).

**CONTEXTUAL CONSTRAINTS**

The first question to be considered is whether the social roles assigned to code switching are the same in the various Hispanic communities of the United States, or whether significant qualitative differences exist among these bilingual communities. In comparing the observations of Mexican Americans and Puerto Rican Americans in several areas of the United States, there do not appear to exist such qualitative differences, although this statement must be hedged somewhat when dealing with the catchall category of ethnic identification and solidarity, since Mexican Americans and Puerto Rican

Americans live in different situations. Intragroup differences also exist according to area, length of residence in the United States, socio-economic status, and perhaps other factors. Poplack (1981:182), in comparing data dealing with both Chicano and Puerto Rican speakers in the United States, speculates that the same constraints will hold for both groups, even in grammatical models, perhaps differentiated by quantitative ramifications on the application of variable rules, much as occurs with many phonological variables. While the same claim is not made explicitly for the case of sociolinguistic variables, the accumulated evidence points to a similar conclusion.<sup>1</sup> Let us therefore enumerate some commonly observed contexts and situations which circumscribe Spanish-English code alternation in the United States.

It is normally assumed that in order for code switching to occur in a conversation, all participants must be functionally bilingual in the languages being used to an extent that would permit comprehension of bilingual utterances as well as monolingual sentences in each language. Language switching is frequently taken as a sign of ingroup solidarity and monolingual discourse may be triggered by the appearance of a speaker who is unknown to the participants, whose bilingual abilities are either unknown or in doubt, or who is identified with only one of the two language domains. Nonetheless, there do exist situations in which code switching may be a deliberate strategy to exclude certain individuals from a conversation, or to 'test' a participant in some fashion. Scotton (1979) gives some examples from Africa in which language switching may momentarily exclude certain participants in a conversation while at the same time not causing the impression of deliberate rudeness and rejection that would follow from monolingual discourse in a non-mutually shared language. Di Pietro (1978) notes that some parents use their native language as a 'secret code' in front of their minimally bilingual children, who in turn develop a passive bilingual competence which is not revealed to their parents; the present writer can personally vouch for the existence of such situations. McMenamin (1973:483) reports that some young Mexican Americans told him that they would deliberately code switch with a Mexican American policeman, "to see if he could understand me--they're mostly vendidos." However, such cases are in a minority, and the presence of individuals who are not bilingual or whose bilingual abilities are doubtful is often sufficient to inhibit all code switching. Even small children have been shown to be sensitive to such factors,<sup>2</sup> while Sánchez (1983:142) postulates code switching as an intermediate developmental or transitional stage for an individual or a speech community, as English replaces Spanish in an increasingly large number of domains.

Mere bilingualism of the interlocutor may not be sufficient, since code shifting often conveys connotations of ingroup identity, ethnicity and solidarity. In particular, membership in the ethnic group characterized by code switching is frequently a necessary ingredient for code switching to be used. When the listener is not part of the ethnic group, no matter how fluent in Spanish or how functionally bilingual, spontaneous code switching is inhibited or entirely suppressed.<sup>3</sup> If the listener is of Anglo American background and/or identified with a professional domain (teachers, social workers, psychological investigators, etc.), the element of formality may be foremost, while in other less formal situations in dealing with Anglo Americans or non-United States Hispanics, lack of code switching may be due to perceived norms of courtesy.

When Anglo American or non-American Hispanic listeners are addressed in a code-switched format, important qualitative and quantitative differences may ensue to distinguish such behavior from conversations with ingroup members. For example, Poplack (1981:180-1) compared code switching among Puerto Rican Americans between an ingroup interviewer and a nongroup interviewer and found, among other things, that switches involving nouns represented 65% of the switches when talking to a nongroup member, but only 24% of the total switches when addressing ingroup members. Similarly, of the inserted nouns, 89% were 'ethnically identified' when speaking to nongroup members, while only 49% had such ethnic connotations when speaking to ingroup members. Intrasentential code switching, exhibiting the highest degree of bilingual integration and characteristic of the linguistic habits of truly bilingual communities, represented only 12% of the switches when speaking to nongroup members, and 31% of the switches among ingroup members. On the other hand, switching at full sentence boundaries occurred in 23% of the cases with nongroup members and in 45% of the cases with ingroup members. These figures are significant in that they indicate the degree to which the ethnicity of the researcher can skew quantitative and even qualitative results on language shifting.

The interview situation itself may also profoundly affect the results of language shifting: Zentella (1981:115) found that the percentage of switches involving nouns and NPs was considerably higher in interviews than in spontaneous speech, noting that this may be the result of the interviewer's selection of the topic and leading the speaker into discussions of his/her choice. In order to maintain the language of the interview, the speaker must temporarily change to the other language for the unknown term. The speaker might avoid this switch for a single item in spontaneous conversation by switching to the other language in anticipation of it (i.e., at the initial boundary of the sentence that includes the term. Kimple et al. (1969) found that among Puerto Rican speakers in New York City, use of English was equated with period of residence in the city, social status and higher occupational status, all factors which can significantly influence the outcome of an interview.

At this point it is necessary to discuss to what extent lexical unavailability is responsible for noun switching. Several investigators<sup>4</sup> have indicated that code switching may occur to fill in lexical gaps in a speaker's repertoire, to provide a signifier for referents which have no ready translation in one of the languages, or to shift the overall language of discourse into one temporarily or permanently more comfortable for one or more participants in a conversation. Insecurity as to the correct usage or pronunciation of a word in one language may also trigger a lexical switch, as in the example from Sánchez (1983: 151):

Lloró porque lo pelis . . . pelis . . . he pinched him.

There may also be significant differences in connotation between 'equivalent' terms in the two languages (Sánchez 1983:163); for example between *bartender* (in an American-style hotel) and *cantinero* (in a local Hispanic bar). Certain words, even though known in both languages, may be identified preferentially with one of the languages,<sup>5</sup> and the frequent successive translation of nouns in code-switched discourse indicates that for many speakers, the switch serves to emphasize, underline or comment on the indicated item, rather than indicating

lack of lexical availability.<sup>6</sup> At the same time, it has been observed that for many speakers, switches in one direction predominate. In the case of United States Hispanics, a shift from Spanish to English has more often been noted; this is consistent with the general observations of Macnamara and Kushir (1971) that bilinguals do not usually translate from the weaker language to the stronger one. Such a preferential directionality may indicate not lexical unavailability but relative facility of retrieval or the extent to which the item in question is integrated into each of the two monolingual grammars.<sup>7</sup> As such, directionality of bilingual transitions may ultimately be used as a form of linguistic-dominance measure, but at present it is only possible to note that the issue of lexical retrievability is complicated by factors of connotation and context. Many speakers indicate that they switch languages (most frequently from Spanish to English) when a word doesn't 'come' in the first language, but actual observations of code-switched discourse in natural settings belie the simplicity of this claim, and indicate rather a complex array of mode shifts, metaphorical connotations and attempts to influence opinion, mitigate requests or otherwise assume an active role in the communicative interchange.

Scotton (1979), in reviewing many African situations, speaks of code switching as a 'safe choice' in choosing a *lingua franca*, where several choices are available and each carries with it certain liabilities and advantages. She notes (71), "We reason code-switching often takes place because the switcher recognizes that the use of either of the two languages has its value in terms of the rewards and costs which accrue to the user. The switcher chooses a 'middle road' in terms of possible rewards and decides to use both languages in a single conversation." Scotton's analysis does not apply in all of its details to the Spanish-English interfaces in the United States, since in Africa a speaker is often faced with the need to express himself in a number of indigenous languages, including standard and nonstandard varieties of each, and in one or more 'colonial' languages, each with corresponding connotations of prestige or ethnicity. Precisely one of the motives for code switching in Africa is the need to impress others with virtuosity in several languages or in one prestige language. This situation may have a parallel in parts of Latin America, where merchants and others in contact with the public speak smatterings of English and other European languages, in the hopes of favorably impressing potential customers. In the United States, however, at least a functional bilingualism is presupposed in the Hispanic communities in which code switching predominates, although the actual degree of proficiency in each language may vary widely at the individual level. Other types of liabilities and rewards may nonetheless be suggested; for example, speaking English in a given situation automatically carries a connotation of 'business,' 'professionalism,' and so forth, since English for the Hispanic-American is the 'they' code as opposed to the 'we' code of Spanish.<sup>8</sup> Conversations with teachers and public officials preferably take place in English, even when all participants are bilingual and even though some may have greater proficiency in Spanish. At the same time, unrestricted use of Spanish even in intimate conversations, may also connote excessive stiffness and "may carry a tone of formality and may in fact sound pedantic" Sánchez (1978:192). Thus code switching is often seen as a way of taking the sharp edge off a situation which, if maintained in a single language, might sound excessively formal or intimidating.

A closely related function of language switching is the mitigation of requests or the subtle insinuation of favors or concessions. In the United States Hispanic communities it is usually a shift from English to Spanish which conveys the insinuation, since Spanish is felt to be "closer to the heart."

A language shift, by its very nature, calls attention to itself, since it represents a radical departure from the norms in either of the monolingual systems. As such, its very function may be an attention-getting device. McClure and Wentz (1975) give several examples, including use of *mira*, *look*, *a ver*, *hey*, etc. Frequently a bilingual repetition, giving the same item in each of the two languages, will add special emphasis, as in:<sup>9</sup>

yo soy segundo . . . I'm second.

va a haber un baile . . . un toys for tots dance  
stay here, Roli. Te quedas aquí.

Clarification of a previous statement not heard or misunderstood may also trigger a bilingual translation:

Qué tiene . . . will you watch your cards? Fíjate en las cartas . . .

Bilingual shifts are common in narrating, when shifting from the narrative description to an actual quote; this usage is common both in speech and in literature:<sup>10</sup>

In the next breath, le preguntó a Ollie: Do you belong Ollie? I mean, are you affiliated? ¿Y Ira? Sonriendo like a cat . . .

Jacobson (1978) has described one motive for language switching, which he calls 'imaginary content.' Jacobson's data indicated that "hearsay, remoteness of occurrence, mystery, the unexplainable or incredulity all contain a number of elements that transport the individual to a domain that is removed from the experiences of members of the mainstream culture," and that Spanish is used in these cases:

. . . we'll talk about la Llorona when, ah, I can tell you de cómo de que dicen que aquí en Espada Park se aparece

yeah, a certain day pero, este, que llueva, con que ese día llueva mucho, y que esté crecido el río pero a medianoche dicen que la han oido

Oh, yes, but I don't know and if that is true or not. De ese que pasó allí en la catedral, también. Andaba pintando la catedral . . . You saw her face in the egg? Exact replica, I mean, just like having a picture you know. y vide a la señora esta (that witch) y esta mujer no la conogía yo . . .

A final factor sometimes associated with the domain of bilingual language switching is the topic of the discussion. Although certain contexts favor code switching, namely informal discussions between members of the same solidarity groups, within the confines of these constraints, the actual topic of the discussion normally has little to do with language-switching rates.<sup>11</sup> If such observations are valid over a wide spectrum of bilingual language situations, then this

behavior is significantly different from such topic-dependent phenomena as use of diminutive forms.

#### SUMMARY OF CONTEXT-DEPENDENCY

In terms of the psycho-social motivation for language shifting, two general areas have been discovered: a response to lexical unavailability either temporary or permanent, and a metaphorical or stylistic code shifting to add a special metalevel of connotation through choice of language and the presence of a shift. Noun insertion is most commonly associated with lexical unavailability, as well as ethnic solidarity when certain highly charged items are involved. Intersentential code switching involves a changing focus, which may be used to mitigate requests, insinuate solidarity, comment on a preceding remark or change the tone of the conversation from formal to informal or vice versa. Intrasentential code shifts, indicating the highest degree of integration of the two monolingual grammars, may follow from either or both of these two motivating factors. The need to supply a particular  $L_2$  lexical item when an  $L_1$  equivalent is unavailable may trigger either a shift at the point of the insertion or an 'anticipatory shift,' as will be described in the following chapters. A subtle shift in the tone of a conversation, often barely perceived consciously by the speakers, may trigger an intrasentential shift, perhaps only in the form of a tag such as *verdad?*, *este*, *I mean, right?* or an attention-getting signal such as *oye* or *mira*. Qualitatively, intrasentential shifting is different from intersentential shifting in that major constituent boundaries are permeated and the shifting is more subtle, to the point where retrospectively the actual switch point may not always be accurately identifiable. At the same time, both forms of switching respond to a variety of personal needs on the part of the speaker, and the speaker's perception of the listeners' needs, and as such are constrained primarily by psycho-social affective factors. Overall sociolinguistic context and the affective parameters have to provide a favorable configuration for a language shift to take place; if these conditions are met, a code shift may then take place, subject to grammatical constraints which will be dealt with in the following chapter. Without the prerequisite psycho-social configurations, code shifting is unlikely and can only be forced in artificial situations. Given a compelling personal urgency to switch languages, however, such a switch can literally occur at any point in the discourse, although the inseparable relationship between grammatical constraints and psychological pressures normally postpones such a shift until the first grammatically acceptable point, which is never long in coming. As a result, while the following chapters will deal only with grammatical parameters, the discussion will always be set against the backdrop of situational and sociolinguistic considerations, without which bilingual discourse is not possible. For reasons to be discussed shortly, no attempt will be made to quantify or otherwise formally incorporate these affective parameters into the linguistic models, but their importance is nonetheless primordial, and they must always be included as a necessary first step in evaluating any proposed language-switching model.

#### NOTES

1. See Sánchez (1983) for Mexican Americans, Poplack (1981) and Fishman et al. (1971) for Puerto Rican Americans.
2. Fantini (1978), Lindholm and Padilla (1978).
3. Poplack (1981), Fishman et al. (1971), McMenamin (1973).
4. Lance (1976), Aguirre (1981), Barkin (1976, 1978), McMenamin (1973), Redlinger (1976), Scotton (1979), among others.
5. Huerta Macías (1981).
6. Cf. Sánchez (1983) for an elaboration of this theme.
7. Cf. McClure (1981:86).
8. Zentella (1981:124), Kimple et al. (1971).
9. These examples come from McClure and Wentz (1976), Sánchez (1983) and from the corpus collected and reported in the following chapter.
10. From Rolando Hinojosa, *Mi querido Rafa* (Houston: Arte Público Press, (1981).
11. McClure and Wentz (1975), McMenamin (1974), Redlinger (1976), Huerta-Macías (1981); however, switches in the topic which inherently lead to greater formality may change the code-switching configurations. Cf. McClure (1981:79), Sánchez (1983:141). Huerta (1980) gives some evidence which may be interpreted as indicating thematic sensitivity of code switching, although in these examples it is impossible to separate out the variable of degree of confidence/formality.

## Grammatical Constraints on Language Switching

### THE ROLE OF GRAMMATICAL CONSTRAINTS

The first step in studying the linguistic properties of bilingual code switching is to discover the grammatical constraints which affect the discourse. Early studies of language switching assumed that nearly any possible grammatical configuration could result from an intrasentential language shift.<sup>1</sup> For this reason, the opinion prevailed that code switching was a form of uncontrolled interference which resulted from improperly learned languages and/or from the inability to keep the languages separated when speaking. Once attention was directed at language switching with the view toward discovering patterns among the existent data rather than exercising a corrective influence on language behavior, it was discovered that code shifting falls within clearly delimited psycho-social boundaries, some of which have been outlined in the preceding chapter. Observations of the sociolinguistic functions of code switching led in turn to the study of grammatical structures, to determine if similar internal coherence could be discerned. From the first, definite grammatical restrictions became evident, which indicated that code switching is a rule-governed form of linguistic behavior, and not an unprincipled confusion.

### TYPES OF LANGUAGE SWITCHING

Investigators have frequently attempted to subdivide the categories of language switching to account for quantitative and qualitative differences which characterize bilingual shifting in differing circumstances. The major bifurcation is between the insertion of individual lexical items, to fill small gaps due to lexical unavailability or temporary unretrievability, and fluent intrasentential switching in which the discourse passes smoothly from one language to the other and remains in the second language for long enough to change the linguistic focus of the discourse. The first type of switching, defined by Barkin (1978a) as *mechanical switching*, often occurs unconsciously, and fills in unknown or unavailable lexical items. Jacobson (1978a) speaks of

semi- or lexical code switching, in which only individual lexical items are inserted, either to fill gaps or for some more subtle stylistic reason. McClure (1981:86) refers to this type of language switching as *codemixing*, which is "the individual's use of opposite language elements which cannot be considered to be borrowed by that individual. It occurs when a person is momentarily unable to access a term for a concept in the language which he is using but can access it in another code, or when he lacks a term in the code he is using which exactly expresses the concept he wishes to convey . . . sentences containing codemixes are generally perceived [by the children in McClure's experiment] . . . to be sentences of one language containing elements of the other."

The other type of language switching, fluent intrasentential shifts which transfer the linguistic focus from one language to the other, is called *codechanging* by McClure, who defines this mode as "motivated by situational and stylistic factors [being] . . . the alternation of language at the level of the major constituent . . . the code change is a complete shift to another language system." All function words, morphology and syntax are abruptly changed:

[1] I put the forks en las mesas.

[2] Let's see qué hay en el dos.

Barkin (1978a) refers to this type of switching as *connotative* or *deliberate*, with the emphasis on the conscious nature of the switching (which in Barkin's categorization may involve one or more lexical items). Jacobson (1978a) calls the same process *true code switching*, by which is meant switching of syntactic elements and not merely individual lexical items. Experimental evidence also bears out this methodological separation, for the relative linguistic abilities of the speaker have a significant effect on the proportion of the different types of switches.<sup>2</sup> There exists a definite hierarchy of code-switching types as the speaker's linguistic abilities range from strongly Spanish dominant to strongly English dominant. Spanish dominant speakers switch isolated items only occasionally, while English dominant speakers who know little Spanish only introduce occasional identity markers, such as *chicano*, *raza*, etc. Bilingual speakers who speak English slightly better than Spanish utilize rapidly alternating Spanish/English switches with great frequency, and the frequency diminishes somewhat among bilingual speakers whose Spanish abilities are slightly stronger. Gingras (1974) suggested that judgments of code switching are not so much due to the fact of being bilingual, but rather the age at which a person became bilingual.

The real or perceived linguistic abilities of the listener may also affect the form of a speaker's code-shifting discourse, as has been seen in the preceding chapter. Poplack (1980) found that the bilingual speech of Puerto Rican Americans contained a much higher proportion of inserted nouns when the listener, although fluent in Spanish, had an Anglo-American outgroup status when the listener was another ingroup member. Similarly, intrasentential shifts jumped significantly among ingroup members, as did shifts at full sentence boundaries. Zentella (1981) noted that in interview situations the proportion of inserted nouns and NPs may be higher due to the forced nature of the discourse, where the interviewer chooses the topic and the speaker must switch languages to access lexical items not available in the main language of the interview. In spontaneous conversation, such lexical

strategies might entail an anticipatory switch before the actual inserted lexical item.

Also important is the directionality of the switch, from Spanish to English or from English to Spanish. The preceding chapter has given examples of the situational variables which entail language shifts, such as mitigation of requests, gaining attention or highlighting certain elements or personal feelings. This fact notwithstanding, quantitative studies of bilingual Spanish-English code switching in the United States have disclosed a significantly higher proportion of switches into English than into Spanish. Poplack's (1980) corpus contains a total of 1176 switches from Spanish to English, as opposed to 659 from English to Spanish. In the category of tag expressions the proportion is even more unequal: 403 shifts from Spanish to English and 9 from English to Spanish. Spanish to English noun insertions outnumber English to Spanish insertions 141 to 34. Aguirre (1981:315) found that among predominantly Spanish speakers, switches from Spanish to English are more acceptable than from English to Spanish when dealing with ungrammatical items. Pfaff (1979) discovered a predominance of Spanish to English shifts in the case of inserted nouns and NPs, while similar results were reported in Valdés-Fallis (1978), and McClure and Wentz (1975). McMenamin (1973) describes an overall preference for switches from Spanish to English in subjects studied in California, and only the oldest speakers representing an earlier generation, switched freely from English to Spanish.

#### GRAMMATICAL CATEGORIES OF SWITCHES

The reasons for language preference are not to be sought entirely in the linguistic dimension, since the nature of the connotative and metaphorical factors described in the preceding chapter sets the stage for use of English to reinforce certain elements in the discourse. In addition, given the traditional educational programs in the United States, many Hispanic Americans have a more extensive vocabulary in English than in Spanish, particularly in areas relating to business, technology and other professional spheres of activity. At the same time, it is not unimportant that, in comparison to Spanish, English is relatively uninflected, at the level of nouns, adjectives and verbs, and therefore a switch from Spanish to English can be effected without transferring or introducing a complicated grammatical concordance. McClure (1981:86) notes that "codemixing takes place within constituents, and there is usually at some level an indication that the code-mixed item is marked for use in a sentence of another code." Reyes (1976:185-6) and Sánchez (1972) speak of the various patterns which have been created in California Spanish utilizing *hacer* + English bare infinite or gerund:

- [3] io dejan hacer improve/lo dejan improve
- [4] va a hacer reenlist/va a reenlist
- [5] los están busing pa otra escuela/los están haciendo bus.

Such constructions would be more difficult to adapt in going from English to Spanish, since additional concordance would have to be introduced in an artificial manner. Pfaff (1979:314) further elaborates a restriction involving Spanish-to-English switches: switches to

morphologically unadapted English verbs are permitted only when preceded by an inflected Spanish verb, unless tense/aspect is overtly marked elsewhere in the sentence. Reyes' examples fit this criterion, as do most others which have appeared in studies on bilingual language switching.

Following the case of inserted lexical items, the most frequently commented structural constraint on intrasentential language switching is that it occur at a major constituent boundary, usually defined to occur before a full sentence (S-node), a NP, VP, PREP, etc. McClure (1981:86) goes so far as to use the presence or absence of a shift at a constituent boundary to define a true code shift (codechanging) as opposed to a lexical insertion (codemixing). For example, a sentence such as

- [6] No van a aceptar a una mujer que can't talk business

would be classified as a codemix, since the relative pronoun, which introduces the subordinate clause, is in L<sub>1</sub>: the corresponding codechanged sentence would be:<sup>3</sup>

- [7] No van a aceptar a una mujer who can't talk business.

In fact syntactic boundaries rank second only to full-sentence boundaries as preferred points for bilingual language shifts. In the experimental mode, Wakefield et al. (1975) created a situation in which subjects had to listen to a sentence containing an intrasentential code shift and then answer a true-false question based on that sentence. Typical sentences were:

- [8] The number of persons who have taken up parachuting as a sport está definitivamente aumentando.  
[9] The number of people who have taken up el deporte del paracaídas está definitivamente aumentando.

Wakefield and his collaborators found that response time in providing the true-false judgments was significantly longer when the switch did not occur at a definable syntactic boundary, as in [9], than when the shift occurred at a syntactic boundary.

Nonetheless, syntactic boundaries are permeable to intrasentential shifts, particularly involving the head elements such as conjunctions, relative pronouns, and prepositions, which casts doubt on McClure's distinction when applied to such pairs as [8] and [9]. Aguirre (1980:302) found that English and Spanish are equally acceptable in a sentence such as:

- [10] Se me hace que/that I have to respect her.

Jacobson (1978b) stated that conjunctions like *que* seem to be below the speaker's level of awareness. Pfaff (1979:311-2) noted that in her collected data conjunctions did not always occur in L<sub>2</sub>, although the proportion of L<sub>2</sub> conjunctions was much higher in the literary work of Ricardo Sánchez and in the lecture style of Alurista. Poplack (1980) reported similar variability of subordinate conjunctions. While it is possible for the conjunction to occur in either language, certain constraints must be respected: Spanish allows for the optional elimination

of a subject pronoun in the dependent clause, while in English such deletion is not possible. Therefore, a shift from an English subordinating conjunction to a Spanish subordinate clause in which the subject has been deleted will be judged unacceptable in most instances, and this type-form is rarely found in spontaneous speech:

- [11] It seems to me que la tenemos que respetar/  
tenemos que respetarla.  
[12] \*It seems to me that tenemos que respetarla/  
la tenemos que respetar.

Coordinating conjunctions such as *and*, *or*, *but* and the Spanish equivalents are similarly variable; the claim to the effect that the conjunction will always appear in the second clause is not borne out by the evidence.<sup>4</sup> It may even happen that one of the coordinating conjunctions will be inserted in L<sub>1</sub>, flanked by L<sub>2</sub> elements, to serve as a momentary metacommentary on the ongoing discourse.<sup>5</sup>

One general constraint operates between a pronominal subject and a conjugated verb, a boundary which is usually impermeable to bilingual switching;<sup>6</sup> thus one cannot say \**El* went, \**He* vino, etc. Gumperz (1976:32) has suggested that it is not the pronoun itself which is responsible for the prohibition, but rather its brevity, and the fact that it cannot occur in stressed or contrastual form. Gumperz has found phrases such as

- [13] That one es el más viejo

to be somewhat more acceptable, due to the increased length and more complex internal structure of the pronominal construction.

Shifts between L<sub>1</sub> articles and L<sub>2</sub> nouns are common; this is the basic pattern of an L<sub>2</sub> noun insertion. The presence of modifiers in the NP causes complications, since English prefers prenominal position for adjectives, while postnominal modification predominates in Spanish. In the area of nominal modification the variability is greatest as regards bilingual switches, and only the most general patterns may be elicited, none without counterexamples in actually existing data. Pfaff (1979:306) claims that in combinations of nouns and adjectives, the mix must match the surface word order of both the language of the adjective and the language of the head noun, a constraint which rules out most potential English/Spanish combinations. Thus, the following would be acceptable, according to this criterion:

- [14] the great caudillo

- [15] the poor huérfano

while the following would be less acceptable, given the reduced probability of prenominal modification in Spanish:

- [16] my red carrota

- [17] su importante composición

the following would be impossible:

- [18] el caudillo important

- [19] the president destacado

In general, such constraints do hold, although in many cases it is difficult to separate lexical borrowing of such words as *pickup*, *hang-glider*, and so forth from lexical insertion of L<sub>2</sub> nouns: [16] and [17]

are generally accepted type-forms of bilingual switching, although [18] and [19] do not frequently occur. Poplack (1980) has given many examples from Puerto Rican-American speech, including *pechos flat*, which undermine the absolute nature of Pfaff's suggested constraint.

Timm (1975) has offered a series of observations on three-word combinations of DET, N and ADJ; the following combinations were judged unacceptable in the order D + A + N:

- SSE: su favorito spot
- SES: su favorite lugar
- ESS: his favorito lugar
- ESE: his favorite spot

in the order D + N + A:

- ESE: his lugar favorite
- EES: his spot favorito
- SES: su spot favorito

The combination EES (*his favorite lugar*) was regarded as acceptable; whereas SEE (*su favorite spot*) and ESS (*his lugar favorito*) were only marginally acceptable. Some of these examples may be rejected out of hand as embodying too much categorical diversity in a short expanse of discourse, since continuous  $L_1 + L_2 + L_1 + L_2 + \dots$  alternation is not normally permitted in any bilingual combination. The remainder of the 'unacceptable' phrases may be subsumed under Pfaff's criterion. However, the judgments reported by Timm are by no means universal, and any investigator attempting to replicate these results will find differing patterns of acceptability response, both with regards to general patterns and with respect to specific noun + adjective pairs.

Prepositions also are variable with respect to choice of language, despite their clear status as the head element of a prepositional phrase.<sup>7</sup> Only when in the equivalent English sentence the preposition would be shifted to the end of a relative clause is a switch to Spanish prohibited as a categorical rule.

Timm (1975) notes other constraints, dealing with the internal structure of verb phrases. For example, negative elements must be in the same language as the main verb, thus ruling out combinations such as \**I do not quiero*, *"El no want*, etc. A switch between auxiliaries and main verbs is normally ruled out: \**I must esperar*, \**he has visto*, \**debo wait*, \**I was caminando*, etc. Sánchez (1972) observes that English syntax may not be used in code-switched questions such as

- [20] \**con quién Peter come?*
- [21] \**cuándo is Mary coming?*

Puerto Rican Spanish does permit this type of sentence order (although not the separation of AUX and main verb), but only when the subject is a pronoun, normally in the second or third person:<sup>8</sup>

- [22] \**Qué túquieres?*
- [23] \**Cómo usted se llama?*

However, it has already been noted that switches between pronominal subjects and verbs are rare; it appears that switches between interrogative words and subject pronouns are equally rare.

Sánchez (1972) gives actually occurring examples where the gerund occurs in Spanish in a fashion reflecting English, rather than Spanish syntax:

- [24] Estoy por lowering the standard
- [25] I'm talking about conociéndonos.

The present writer has observed:

- [26] . . . porque ella está going to have a baby
- [27] we were tratando de ver . . .
- [28] You ought to ammarrarte los pantalones
- [29] Debemos wait in line.

Only the boundary between *haber/have* and the past participle seems to be a truly insurmountable obstacle to intrasentential shifting. Reyes (1976) has noted that some of the above combinations are being lexicalized in stable bilingual communities, so that these restrictions, if valid, presumably apply only to non-lexicalized combinations of English and Spanish elements.

Timm also notes the impossibility of a switch between a finite verb and an infinitive complement, ruling out \**they want a venir*, \**quieren come*, and so forth. Some of these forms have been adapted in bilingual communities, but even those which retain their monolingual status do sometimes participate in shifts, subject to the constraints mentioned by Pfaff (1979: 314). It is not impossible to find examples such as:

- [30] I'm going to romper la cara
- [31] He would mentarme la madre
- [32] Vamos a go hang-gliding.

However, a switch before Spanish *a* is rarely observed.

The nature and scope of the previously mentioned constraints indicates that, while a definitive characterization of the grammatical properties of language switching may not yet exist, there are principled linguistic mechanisms at work. The literature on Spanish-English code shifting indicates that several extensive sets of data have been collected.<sup>9</sup> However, only a few partial texts have been published, and only Poplack (1980) has furnished a complete numerical breakdown of the results. The need was thus felt to collect an additional corpus of data to corroborate the results reported above, and to serve as an empirical basis for more elaborate syntactic analyses contained in the following chapters.

#### THE COLLECTED MATERIALS

A study was made of bilingual interactions in Houston, Texas, combining recordings of spontaneous conversations and materials which resulted from the fortuitous existence of a radio station which offers a variety program directed specifically at Houston's Mexican American community. The key feature of this program is that the announcers are extremely informal, which they take as a source of pride in not being a 'commercialized' radio station but rather a publically sponsored station. Their stated aim is to represent *la onda chicana*, both in music and in commentary, as opposed to the several commercial Spanish-language stations in Houston, which play exclusively Mexican music and whose program format and announcers are nearly always Mexican. The Mexican American listener is left out by such programs, claim the staff of the program in question, who seek to fill the gap by playing music by local Mexican American artists, and who invite

all and sundry to come and visit the station and talk over the air. The resulting programs are informal in the extreme, with false starts, mis-cued tape machines, knocked-over chairs, and at times inconsequential ramblings. The linguistic results are a goldmine for the investigator, for the announcers engage constantly in language switching, as do nearly all the occasional participants in the program. No attempt is made to eliminate the switching; to the contrary, the announcers have indicated that this is an essential part of "la onda chicana" and such behavior is actually encouraged. Some of the occasional participants exhibit a reluctance at first to engage in this linguistic mode over the air, but the easy patter of the announcers soon puts them at ease and nearly all participants, after a few minutes in the homelike atmosphere of the studio, drop their self-consciousness and speak in a normal and relaxed style. The unrehearsed nature of the program occasionally results in some un'-bleeped' material going out over the air, but this happens infrequently enough so that no attempts have been made to curb the program's spontaneity.

In order to prepare a corpus for analysis, approximately 30 hours of program material were analyzed, and the bilingual switched portions were culled out. Only bilingual speakers from Texas served as sources of data (these facts were verified personally by the present writer) and musical segments obviously were discarded. The resulting editing yielded approximately 12 hours of concentrated bilingual material, which was then transcribed and further analyzed. Of the radio material, 22 speakers provided data; 8 were regular program personnel, and the remaining 14 provided a minimum of 15 minutes apiece of transcribed material. The age range was 16 to over 60, and there were 14 men and 8 women. Topics were extremely varied, and included popular music and bands, social activities, the prison system, politics and social issues, public schooling, questions of Mexican American identity and other topics related to daily activities. The remainder of the corpus consisted of approximately 3 hours of spontaneous bilingual interaction which had been recorded on various occasions, involving native Texas residents. This material was analyzed with the primary aim of determining whether the materials collected from the radio were qualitatively different from those resulting in normal conversation. Despite the frequently commented 'paradox of the observer' which occurs commonly in the presence of tape recorders, radio studios and the like, no significant differences were observed between the two sets of data, either in terms of relative frequency of shifts or of the frequency of given shift types. This demonstrates the natural atmosphere which the radio program described above has managed to achieve, and suggests an additional source of linguistic materials for bilingual interaction.

In classifying switch types, the frequent pause-word *este* has not been included, since it was indiscriminately in both Spanish and English discourse by numerous speakers and could not be correlated with the remaining discourse in any significant manner. Since use of this pause-filler has been included in some other analyses of code switching, the comparative figures to be reported below must be viewed in the light of this exclusion. Similarly, despite the high degree of English syntactic influence on Texas Spanish, none of these cases was considered as true language switching (e.g., *hacer su parte* for 'do one's part'; *estar envuelto en* for 'to be involved in' etc.); nor were what were judged to be assimilated borrowings, such as *chicano*, *salsa*, *pachuco*, *pitcher*, *sherriff*, and so forth. The remain-

ing switches were analyzed according to the syntactic function characterizing the point of the switch. Table 3-1 gives the results for the tabulations computed on the 2319 true language switches that were analyzed. In Table 3-2, the figures from Poplack (1980) relating to New York Puerto Rican speakers are given for comparison. As can be seen, the degree of correspondence is quite high, despite the different classificatory schemes which have been utilized: these correspondences, representing the combined total of more than 4000 examples of Spanish-English language switching, suggest a high level of confidence surrounding the data reported above, which will therefore be referred to in the succeeding chapters, where more detailed syntactic properties will be explored.

TABLE 3-1  
Type-forms of code switches, Mexican American  
(N = 2319)

Category	Percent
<b>Adverb Phrases</b>	
And/or/but (before & after)	5.17
Article + L <sub>2</sub> noun	11.19
Between subject and verb	5.34
Between verb and object	1.55
Bilingual repetition	2.99
Copula (before and after)	5.63
L <sub>1</sub> + coord. conj. (L <sub>2</sub> ) + L <sub>1</sub>	1.04
L <sub>1</sub> + sub. conj. (L <sub>2</sub> ) + L <sub>1</sub>	1.44
L <sub>1</sub> adjective + L <sub>2</sub> noun	1.21
L <sub>1</sub> noun + L <sub>2</sub> adjective	0.17
L <sub>2</sub> noun	0.46
Miscellaneous	8.27
Noun in apposition	0.25
Other subordinate conj. (before & after)	4.71
Porque/because (before & after)	3.62
Preposition (before & after)	2.30
Relative Pronoun (before & after)	16.13
Sentence Boundary	3.10
Tag Phrases	15.67
	9.76

TABLE 3-2  
Categories of code switching among New York Puerto Ricans  
(N = 1835)

Category	Percent
Adjective	0.8
Adverb	2.6
AUX	0.0
Conjunctions (sub., coor., rel. pronoun)	2.7
Determiner	0.2
Filler	1.1
Idiomatic Expression	1.7
Independent Clause	4.3
Interjection	6.3
Object NP	7.6
Phrases (prep., adj., inf., adv.)	5.1
Predicate Adjective	2.3
Preposition	0.1
Quotation	1.9
Sentence	20.3
(Single) Noun	9.5
Subject NP	3.8
Subordinate and Relative Clause	4.1
Tag	22.5
Verb	1.0
VP	2.2

#### TYPE-FORMS OF BILINGUAL SWITCHES

As based on the corpus reported above, shifts were most common in the following circumstances:

A. The largest single category (15.67% of the total) represents switches occurring at sentence or clause boundaries. In some cases L<sub>2</sub> material was inserted into an L<sub>1</sub> discourse in a fashion that may be analyzed as intrasentential switching:

[33] Aquí con Hope I'm tryin' to figure out what's wrong with the microphone.

In other cases, a true sentence boundary characterizes the beginning of the switch, even though the discourse passes smoothly from L<sub>1</sub> to L<sub>2</sub> with no discernible pause and only an intonational contour indicating the presence of two sentences:

[34] He was responsible for my knowledge of music.  
Carlitos da para la música.

Poplack (1980) analyzed such switches in terms of the preceding constituent; thus [33] would be analyzed as a switch involving a prepositional or adverbial phrase. In the present analysis, only the immediate surface structure has been taken into account, and the syntactic function of the constituent immediately preceding a switch has not been calculated.

B. Tag phrases were frequently introduced (in L<sub>2</sub>) in the midst of an L<sub>1</sub> discourse: *I'm sorry, verdad, you know, ándele, I mean, etc.*

C. Inserted L<sub>2</sub> nouns, most frequently an English noun in the midst of Spanish discourse.

D. Switches occurred immediately preceding (1.61%) or following (2.01%) the conjunction *que/that*; normally it has been said that the remainder of the sentence will continue in the language of the conjunction, but exceptions do occur:

[35] I'm not sayin' that son chuecos, yo no digo eso.

E. Switches occurred before (1.21%) or after (1.09%) the conjunction *because/porque*:

[36] No podemos hacer nada porque we don't have the power

[37] Me tiene envidia because I'm better lookin' than he is.

The same pattern also occurs with other subordinating conjunctions.

F. Switches occur before (2.64%) or after (0.46%) the relative pronoun *que/that*:

[38] Escucharon a un señor que has been around for a long time

[39] She told me to make a dedication to her son que le dicen el pachuco de Rosenberg.

G. The subject-verb boundary (1.55%) and the verb-object boundary (2.99%) were frequent sites of shifting. In ordinary transitive or intransitive sentences, no element may exist precisely at the syntactic boundary between subject and predicate, and therefore if a switch is to occur, it will divide these two constituents. However, in the case of a copula, this element in a sense straddles the subject-predicate boundary, and is highly variable with respect to category inclusion. In the data considered, the copula occurred in L<sub>1</sub> (0.75%) or in L<sub>2</sub> (0.29%). One interesting exception was provided by

[40] . . . porque ella está going to have a baby.

Clearly the *está* is not an auxiliary verb, since one never says \**está yendo a tener un bebé*; rather, because of the way the sentence was pronounced, with a slight pause after *está*, it is clear that *está* is used as a copula, as in *está encinta*, then followed by an equivalent English expression which only coincidentally contains *be*. In general, as noted above, the boundary between auxiliary *be/estar* and a following gerund is not conducive to a language switch, particularly from English to Spanish. However, in the case of [40], the switch is facilitated by a categorical ambiguity of another sort, between copula and auxiliary verb. This latter confusion is also evidenced by:

[41] . . . the vacant property que es adjacent to a church.

Spanish would normally use *estar* or *quedar* to express the location of a piece of property; however in the sentence, it is possible that the speaker was originally thinking of a defining phrase, such as *es propiedad de la iglesia*, which would highlight the lexical split occasioned by the equivalence of English *be* to Spanish *ser* and *estar*. It

is also possible, of course, that the slight pause which followed the verb represented an attempt to rectify an inappropriately used form.

H. Switches occurred immediately preceding adverbial phrases (5.17%), usually involving a gerund:

- [42] We used to go until three o'clock platicando con los borrachos
- [43] The Houston fire fighters will have their dance el trece de diciembre.

I. Switches occur before (12.11%) or following (4.02%) a preposition:

- [44] He's from Corpus Christi, con una canción titulada . . .
- [45] Ben viene con the sports news.

J. Switches occur preceding (11.77%) or following (4.42%) the coordinating conjunctions *and/y*, *or/o* and *but/pero*.

K. Noun-adjective combinations represented few of the code switches. In view of the categorical restrictions mentioned earlier, and in many cases the nouns involved may be on the way to lexicalization. However, nouns in apposition represented nearly 5% of the total.

L. Another case of language switching, representing more than 2.5% of the total, is the insertion of a coordinating or subordinating conjunction in L<sub>2</sub>, in the midst of an otherwise L<sub>1</sub> discourse.

M. The final area to be mentioned is the bilingual translation. Although it may be the case that certain cases of language switching, particularly lexical insertion, reflect unavailability of given lexical items in one of the languages, the frequent occurrence of immediate bilingual translations indicates that such is not always the case. It remains to be demonstrated, however, that the L<sub>1</sub> form and the equivalent L<sub>2</sub> form are always equally retrievable, and it is necessary to consider such factors as retrieval time and the direction of the bilingual translations. In the present corpus, bilingual translations represented 5.6% of the total, of which 4.8% were from English to Spanish and only 0.8% from Spanish to English. Examples include:

- [46] . . . otra vez a las one o'clock, a las una [sic.]
- [47] Nos dieron la oportunidad de . . . raise some money, recoger fondos
- [48] una canción recently released que apenas salió no hace mucho
- [49] it all depends on what part of the state de qué parte del estado it is, you know
- [50] The people identify with the music more than, you know, they identify with the group o los grupos que vienen del valle.
- [51] Por las palabras, cómo está arranged, verdad, cómo, está compuesto, it's progressive
- [52] You don't just say hey, here, aquí está, you know
- [53] Vamos a eschuchar a very pretty young lady, una señorita que viene de San Antonio . . .
- [54] Va a haber un baile, un Toys for Tots dance . . .

[55] Yeah, sí, ya está liste

[56] Mezcal va a tocar this coming Friday, este viernes.

In many of these cases, the bilingual translation leading from English to Spanish has the function of allowing a switch to L<sub>2</sub>, indicating perhaps some overriding desire to change languages. In most cases, the words involved are common and easily available in both languages, and the repetition of an English term in Spanish has the effect of reinforcing the bilingual nature of the conversation, and leaves open the possibility of further discourse in L<sub>2</sub> (usually Spanish). Since such bilingual translations are often performed unconsciously and the speakers are not subsequently aware of having uttered the same material successively in two languages, more than simply linguistic factors must be brought into consideration. It could be noted that even in normally monolingual discourse, when an L<sub>2</sub> word is introduced, it is not uncommon for the L<sub>1</sub> word to immediately follow, if it is known:

[57] cómo está arranged, verdad, cómo está compuesto.

Even speakers who do not normally engage in code switching and even those who are not bilingual in a balanced fashion may exhibit this same behavior, perhaps as a semi-conscious apology for having introduced a foreign expression. In the data under consideration, 53% of the bilingual translations represent the introduction of an L<sub>2</sub> term following discourse in L<sub>1</sub> (i.e., the discourse immediately preceding the first instance of the bilingual pair is in L<sub>1</sub>, while the bilingual pair takes the form L<sub>1</sub>/L<sub>2</sub>); the remaining translations are of the form L<sub>1</sub>/L<sub>2</sub>.

## NOTES

1. For example Lance (1975), Espinosa (1917).
2. See Valdés (2978) for an elaboration of this topic.
3. A preference for shifts at major syntactic boundaries has also been reported by Reyes (1976), Poplack (1980), Jacobson (1978b), Dearholt and Valdés-Fallis (1978), McMenamin (1973) and Timm (1975), among others.
4. The original claim was made by Gumperz and Hernández-Chávez (1975); counterexamples are found in Poplack (1980), Pfaff (1979) and others.
5. Cf. Valdés-Fallis (1978), Barkin (1978b).
6. Pfaff (1979), Timm (1975).
7. The categorical variability of prepositions has been commented by Pfaff (1979), Poplack (1980), Timm (1975) and others.
8. Cf. the discussion in Lipski (1977b).
9. For example, by Sánchez (1972), Timm (1975), Jacobson (1978a), Valdés-Fallis (1974), Poplack (1980), Aguirre (1981), Wentz (1977), McMenamin (1973).

**The Role of Syntactic Congruence****BILINGUAL PHRASE STRUCTURES**

Establishing the word-level constraints on intrasentential code switching demonstrates the nonrandom nature of language alternation among Spanish-English bilinguals, at the same time highlighting some possible aspects of the grammars that might be written to characterize this type of bilingual behavior. On the other hand, the delimitation of such constraints does little in the way of resolving the question of the degree of integration of grammars representing Spanish and English, since it is possible to account for the categorical restrictions on language switching without positing any form of unified bilingual grammar. It is merely necessary to specify the phrase-structure restrictions for each language individually, together with a list of possible switch types, in a fashion which may be represented either in the phrase structure rules themselves, or conceivably even as a form of lexical diacritic. In order to illustrate the inadequacy of such an approach, let us construct a first approximation to a generative bilingual grammar, and then pass to possible refinements as well as actually proposed models.

A typical phrase structure rule is the one generating prepositional phrases in English and Spanish, roughly

{1} PP → PREP NP

together with

{2} NP → DET N ADJ

Since language switches occur from English to Spanish and from Spanish to English between preposition and object, it may be possible to simply mark the phrase structure rule itself with a diacritic indicating the feasibility of a language shift, using, for example, 1 to represent  $L_1$  and 2 to represent  $L_2$ :

{3} PP → PREP<sub>1</sub> NP<sub>1/2</sub>

This representation permits the generation of bilingual prepositional phrases such as *en the laundromat*, *to la iglesia*, etc. The phrase

structure of the NP also admits a language switch between determiner and noun, so that phrase structure rule [2] may be rewritten to include the possibility of a bilingual code shift:

$$[4] \quad NP \rightarrow DET_1 N_{1/2} ADJ_{1/2}$$

A phrase of the form PREP<sub>1</sub> DET<sub>2</sub> N<sub>1</sub> would be disallowed by general constraints regulating the density of interlanguage variation within a given expanse of discourse.

This simple bilingual phrase structure rule approach will permit the generation of many bilingual language shifts, particularly between article and noun, copula and subject or predicate, and preposition and object. This approach does not strongly militate in favor of either two separate grammars or a single bilingual grammar, since many of the basic phrase structure configurations are essentially the same in both Spanish and English. However, if one attempts, for example, to incorporate adjectives into the NP phrase structure component, some difficulties arise immediately, since English almost invariably preposes adjectives, while Spanish permits both preposed and postposed adjectives, depending upon several complex criteria.<sup>1</sup> For example, a typical English phrase structure rule including adjectives would take the form:

$$[5] \quad NP \rightarrow DET \text{ ADJ}_a \text{ ADJ}_b \dots \text{ ADJ}_n \text{ N}$$

where the presence and relative order of the adjectives filling the various slots are fixed by general constraints. The equivalent phrase structure rule in Spanish takes the form:

$$[6] \quad NP \rightarrow DET \text{ ADJ}_a \text{ ADJ}_b \text{ N ADJ}_c \text{ ADJ}_d$$

Certain Spanish adjectives always precede the noun, while others nearly always follow. In a few cases, preposition or postposition of adjectives appears to vary more freely, such as with ordinal numerals applied to such words as *capítulo*, *lección*, etc. In the remaining cases, adjective placement is constrained by questions of reference (part vs. all of a group), contrast, or general considerations of euphony and style, particularly when two or more adjectives modify a single noun. The result is a considerably smaller degree of fit between Spanish and English as regards adjective placement vis-a-vis the head noun, a fact which has immediate repercussions for the grammatical structure of bilingual language switching.

To take a simple example, consider the phrase *the red truck*, which in Spanish would be *el camión rojo*. A phrase paralleling the English structure, *el rojo camión* would not normally be possible in Spanish in the absence of an additional context removing the contrastive adjectival force of *rojo*. Regardless of the theoretical or methodological approach taken to account for the placement of *rojo*, the fact remains that in most cases, *rojo* will follow and not precede *camión*. Similarly, *red* may never follow the head noun in English, so that \**the truck red* is not only uncommon, it is impossible.<sup>2</sup> As a result, it is difficult to code switch within this type of phrase, since nearly all possible permutations result in grammatical unacceptable configurations:

- [7]    \*el camión red,
- \*el red camión,
- \*the truck rojo,
- the red camión.

In fact, only a switch after the article is generally acceptable:

- [8]    el red truck,  
            the camión rojo.

It might be argued that *the red camión* and *el truck rojo* might be possible if *camión* has become quasi-lexicalized in English, or if *truck* has achieved similar status in Spanish (as it has in the assimilated Anglicisms *troca* and *troque*). This would be a defensible position with more specialized terms, such as *pickup truck*, *van*, *four-wheeler*, *off-road vehicle*, *station wagon*, etc. but *truck* and *camión* are part of the fundamental vocabulary of all speakers of Spanish and English, and it is unlikely that lexical borrowing of these terms would occur among speakers bilingual enough to freely engage in language switching.

How, then, is such a restriction on possible language switches to be incorporated into the grammatical component characterizing the speaker's linguistic competence? It would be possible to insert an *ad hoc* diacritic into the phrase structure rules in English and Spanish to prohibit such a switch, indicating that the presence of a postnominal adjective in an NP would restrict the phrase to Spanish, while the presence of most (but not all) prenominal adjectives would restrict the phrase to English. There remains, however, the question of smoothly incorporating this information into a single phrase structure rule, as could be done in the case of a simple prepositional phrase with no adjectives. It appears that the phrase structure rules themselves must be bilingual, or at least some of them, rather than simply representing the intersection between phrase structure rules in English and in Spanish.

Consider also the relative placement of prepositions and relative pronouns. A preposition, by definition, precedes its object, at least originally, and in normative English grammar, such is indeed the case. At the same time, in colloquial spoken English, particularly in North America, such structures are becoming less and less common and are even considered overly affected. Thus, while the following sentences would once have been considered 'normal,' and are still insisted upon by teachers requiring 'proper English,' they would scarcely ever be heard in daily conversation:

- [9]    This is the car behind which we were parked.
- [10]   I asked him to show me the book about which  
            we were talking.
- [11]   I bought the picture at which we had looked  
            at for so long.

Rather, it would be much more common to hear:

- [9a]   This is the car (that) we were parked behind.
- [10a]   I asked him to show me the book (that) we  
            were talking about.
- [11a]   I bought the picture (that) we had looked at  
            for so long.

In the [a] examples, a transformation has applied to the previous sentences, moving the preposition to the end of the clause. It may even be possible to replace *which* by *that* or even (in a style which many consider substandard) by *what*. Yet another transformation optionally eliminates the relative pronoun, as shown above, when it is the object of a restrictive clause.

In Spanish, the equivalent sentences maintain the preposition firmly entrenched before the relative pronoun:

- [12] *Este es el carro detrás del cual estacionamos.*
- [13] *Le pedí que me enseñara el libro del cual estábamos hablando.*
- [14] *Compré el cuadro del que estuvimos admirados por tanto tiempo.*

A switch is therefore impossible after the antecedent noun in English, if the preposition-moving transformation has applied. On the other hand, a shift is possible from Spanish to English at this point:

- [15] *Este es el carro that we were parked behind.*

There is no way of incorporating this information into a deep structure rule in a non *ad hoc* manner, since reference must be made to the possible application of transformations which apply later in the derivation. The surface structures must be considered, in a fashion which makes essential mention of the equivalent structures in both languages. In effect, the constraint must indicate that a switch between the relative pronoun in English and the following phrase in Spanish is only possible if the preposition intervenes between the antecedent noun and the relative pronoun. This permits switches such as:

- [16] *I asked him to show me the book about which estábamos platicando*

while disallowing:

- [17] \**This is the car that estacionamos (detrás).*
- [18] \**I asked him to show me the book which (estábamos platicando).*

On the other hand, when going from Spanish into English, this restriction does not appear to hold, since one can certainly hear not only sentences of the type:

- [19] *Le pedí que me enseñara el libro del que we were talking.*
- [20] *Este es el carro detrás del cual we were parked.*

but also:

- [21] *Le pedí que me enseñara el libro that we were talking about.*
- [22] *Este es el carro that we were parked behind.*
- [23] *Compré el cuadro that we looked at for so long.*

In other words, the preposition moving-transformation may only apply if the segment which meets its structural description is entirely in English. Thus, if a transformation T may apply only in one of the two languages, say  $L_1$ , a code switch may not take place if it would shift to  $L_2$  the portion of the sentence to which T has applied. Consequently, it appears that if code switching is to be described by listing possible bilingual phrase structures, these structures must be cross-referenced in some fashion with the transformations which apply to both languages.

#### REFINING BILINGUAL PHRASE STRUCTURE MODELS

Considerations such as those just discussed led Barkin and Rivas (1980) to elaborate a more formalized set of constraints on bilingual phrase structure rules. They start from the determination of the 'base language' of a bilingual phrase structure rule. Formally, this question is: given a bilingual PSR of the form:

$$[24] \quad A \rightarrow B_1 B_2 \dots B_n$$

and given the languages of the daughter nodes  $B_1, B_2, \dots, B_n$ , what is the language of the mother node A? Two answers are tentatively proposed, namely that the mother node has to be in the language of function words (articles, determiners, pronouns and prepositions), and that the mother node is in the language of the verb. Transformations then apply to bilingual phrase structure rules in accordance with the language of the mother node; that is, Spanish transformations apply to bilingual PSRs deemed to be Spanish and English transformations apply to PSRs deemed to have English as their base. Rivas (1981) takes this notion one step further by claiming that the bilingual PSRs generate all possible base structures of the bilingual code, whether they have high or low frequency. "The variability is to be accounted for in the Transformational Component by a set of conditions that determine the manner in which transformations are applied." These conditions are:

- a. Each language has a set of monolingual PSRs and each element in a rule is indexed with a symbol that represents the language to which the rule belongs.
- b. In the process of derivation, an element indexed in one language can only be rewritten using PSRs of the same language.
- c. For each monolingual PSR in one language, bilingual PSRs can be generated by indexing any element in the right-hand side of the rule with the symbol of the other language, but at least one of these elements has to remain indexed in the original language.
- d. When lexical insertion occurs, terminal elements indexed in one language must be rewritten with morphemes belonging to the same language. These conditions are summarized by defining a well-formed bilingual deep structure as a structure in which every element indexed in  $L_1$  has been rewritten according to a rule from  $L_1$  and has in its rewriting rule at least one right-hand element indexed in  $L_1$ .

Taken together, these conditions assure that an English phrase structure will not serve as input to Spanish transformations, and vice versa. The actual application of transformations is constrained by yet another set of conditions, which specify, among other things, that each language has a set of monolingual transformations, and that a transformation of one language can apply only if the domain of application is an element indexed in that language. In other words, it is assumed that bilingual deep structures are primordial, with restrictions affecting the transformational component. In establishing a hierarchy which ranks monolingual utterances as most acceptable and code-switched utterances along a sliding scale of acceptability, it is proposed that a transformation applies with less frequency (and produces less acceptable utterances) when one or more of the elements in the structural description are indexed in the other language, with the ratio

of L<sub>2</sub> elements in the SD to acceptability being inversely proportional.<sup>3</sup> Naturally, it is also maintained that a transformation cannot change the language of an element which has already been generated.

This formalization addresses most of the preceding issues, but certain problematic areas still remain. Defining the base language of a phrase structure as the language to which the articles pertain makes it impossible to account for phrases of the sort *the camión rojo*, which do occur with some regularity, nor of bilingual phrases with no articles, such as *pechos flat*, reported by Poplack (1980). Claiming that the noun + adjective combination is becoming lexicalized in English is begging the question, as well as being rather unlikely in the case of common nouns with ready equivalents in both languages. The notion that transformations cannot change meaning and that bilingual phrase structure rules cannot be entirely rewritten in L<sub>2</sub> are not really grammatical constraints, but rather well-formedness conditions on the overall description; if one's rules can entirely change the language of discourse, then the PSRs and transformations have in effect unlimited power, and cannot be constrained by any normally accepted linguistic mechanisms. The same is true for lexical insertion, for in reality, it means nothing to postulate a structure such as:

[25] DET<sub>eng.</sub> N<sub>sp</sub> (the maestro)

The consideration that the base language of a sentence be determined by the language of the main verb is problematical; if we consider:

[26] Este es el carro that we were parked behind,

then according to the above proposal, the sentence should be regarded as Spanish, in which case the transformation moving the preposition to the end of the relative clause should not apply. Nonetheless, it is evident that the structural domain of this transformation is entirely in English in the bilingual pretransformation stage, from which it follows that the preposition-moving transformation may apply. This suggests that for each clause, separate values of language assignment must be made in order to insure the consistency of transformational application. The very notion of assigning a base language to a bilingual sentence or clause may be at the root of the problem, for other than the generally uncontroversial assignment of individual words to one language or the other, the assignment of a bilingual sentence to a single language is an inherently paradoxical enterprise. This matter will be considered in the following chapter.

#### A BILINGUAL AUTOMATON

Another attempt at constructing a formal model of code switching is offered by Dearholt and Valdés (1978), in the form of an automaton with generative capabilities which include code-switched discourse. The proposed model reproduces (1) code switches resulting from lexical deficiencies or untranslatability of certain items; (2) apparently 'random' code switching (this apparent randomness is never defined); (3) the dependence of code switching on sociolinguistic context; and (4) the dependence of code switching on observed syntactic constraints. In order to construct this model, the authors adopt the idea of a *chunk* of information, which is "a unit of thought which can be expressed in a few words, and may be recognizable by being expressed in . . .

phrases, clauses, simple sentences." Since normal speech, although considered 'fluent,' is in reality composed primarily of relatively short groupings, the authors make the assumption that "context-invariant chunks of information delineate important boundaries in the synthesis of speech." Other key assumptions include that a bilingual's choice of language can be modelled by a regular grammar, and the following metatheoretical postulates:

A. That the sociolinguistic context of a discourse can be encoded into a finite string of symbols, to be regarded as an input string.

B. That the sociolinguistic context string as encoded largely determines the state of a probabilistic automaton, and thereby also determines the set of probabilities used to choose a language.

C. That the probabilities of switching words, clauses and phrases and sentences can be measured with reasonable accuracy for different types of bilingual speakers.

In none of the above assumptions is concrete evidence given which would permit the quantification of the relevant sociolinguistic and contextual variables, and therefore it remains problematical whether the metatheoretical basis of the model can ever be reasonably achieved, given the failure to achieve similar quantifications in other areas of sociolinguistics. If one does accept such quantifiability in principle, then the assumption that a regular model can reproduce a bilingual's linguistic production is beyond objection. The randomness inherent in the model remains debatable, given the fact that investigators have indicated that individual code switches may be described by a combination of grammatical, sociolinguistic and affective variables. Detailed observation of the output of bilingual speakers does not reveal any element of language switching that could be described as truly random. It is true that certain formal models in which a random factor is built in do achieve results which asymptotically approach those of human speech, but randomness as the motive force behind code switching is the weakest of all linguistic hypotheses, and should be rejected in favor of more structured models, unless available evidence fails to justify alternative hypotheses.<sup>4</sup>

The Dearholt-Valdés model also makes the assumption of parallel construction in Spanish and English, so that the order of the chunks is the same in both languages. This assumption, essentially the same as the congruence constraints to be discussed shortly, immediately stacks the deck in favor of the formal generation of language-switched utterances. From this point on, it is relatively easy to establish a few syntactic constraints, but the formal incorporation of syntactic congruence or co-grammatical equivalence is not nearly so easy. As a final step, the Dearholt-Valdés model proposes to weight each lexical item with a quantified value of 'difficulty,' derived largely from word frequency tables. This neo-Zipfian notion is introduced without question, despite the as yet unsettled dispute regarding the 'difficulty' (in terms other than simple phonetic length or morphological irregularity) which may be correlated with the position of a given element in a text-frequency table.<sup>5</sup>

In the end, only a flowchart, not a formal algorithm, is given, which schematizes the generation of language switched utterances if all the quantitative requirements could be fulfilled. Since such quantification has not been performed, and is based on questionable assumptions about the basic nature of code switching, this model remains a

theoretical curiosity more suited to cybernetic experiments than to the modelling of real-world bilingual linguistic production.

### TOWARDS A COMPLETE BILINGUAL GRAMMAR

A more ramified phrase-structure model of bilingual language switching is offered by Sankoff and Poplack (1980, 1981), who reject the possibility of providing empirical evidence for a single 'base language' in discourse which contains one or more essential syntactic shifts. At the same time, the authors note that the majority of code-switching constraints deal with surface structures and not with postulated deep structures to which transformations have not yet applied. For example, if one accepts that underlying subject pronouns must be in the same language as the verb, then the following type of sentences, which have been observed in bilingual Hispanic communities, should be excluded:<sup>6</sup>

[27] You estás diciéndole la pregunta in the wrong person.

Similarly, assuming that the passive sentence is generated via a transformation, the following sentence-type, also attested in actual discourse, should be disallowed:

[28] La pregunta fue dicha by you.

To these examples one may add the increased acceptability of a phrase such as:

[29] We left immediately after tú llegaste . . .

as opposed to

[30] \*We left immediately after llegaste,

since in English the conjunction *after* is obligatorily followed by a subject, either nominal or pronominal, and a switch to Spanish must necessarily preserve this postswitch configuration. In transformational analyses of Spanish, it has been assumed that subject deletion is a transformational operation, in which case there is no way of excluding [30] if only deep structure constraints are considered.

Assuming, then, that constraints on intrasentential code switching must be at least partially if not entirely based on superficial configurations, it must be decided whether one grammar or two will be used in the base component. Poplack and Sankoff note that this is an academic question, since "any finite set of rules and procedures for generating an infinite set is a grammar . . . so that any set of rules for constructing a set of sentences containing code switches is a grammar." (Sankoff and Poplack 1981:10). At the same time, such a liberal definition is useless for the actual task of constructing a grammar which will generate all and only the permissible language shifts which have been independently observed in bilingual Hispanic communities. As a first step toward combining two monolingual grammars to allow for intrasentential shifting, Sankoff and Poplack describe a *free-union grammar*, constructed as follows. Two translatability conditions must be satisfied in order to effect the free union of grammars  $G_1$  and  $G_2$ , both of which are satisfied by the pair Spanish/English (and presumably by most if not all other pairs of natural languages as well). The first is that there exist two context-free phrase structure grammars  $G_1$  and  $G_2$  for languages  $L_1$  and  $L_2$  such

that the nonterminal grammatical categories of one generally have corresponding categories in the other (e.g., verbs, nouns, prepositions). The second condition is that each rule is  $G_1$  can be functionally translated by at least one rule in  $G_2$ . In the case of Spanish/English, this condition seems intuitive, since although there do exist transformational differences between the two languages, they are minimal. Thus, for example, the Spanish phrase structure rule which results in postposed subjects

[31]  $S \rightarrow VP\ NP$

can always be 'translated' by the English rule

[32]  $S \rightarrow NP\ VP$ .

Despite the high degree of rigor which characterizes this model, no formal definition is given for the second translatability condition, and it is possible that in certain constructions, for example highly colloquial modalities involving 'expletives,' such translatability might not be achieved so readily:

[33] This is un-bloody-believable.

[34] He's the vice-fucking-president.

[35] ¡Qué paciencia ni qué carajo!

Another possible area of difficulty would be the use of *vos* and *usted* as vocatives in Central America:

[36] Ese maneja como loco, *vos*. (a commentary to a friend).

[37] No tengo los papeles, *usted*. (as said, for example, pleadingly to a customs inspector).

In such cases, code switching would be rare, since these expressions are highly idiomatic and language-specific; in the majority of instances of basic phrase-structure patterns, Spanish and English do have reasonably similar configurations.

Given the two translatability conditions, the free union of  $G_1$  and  $G_2$  is the set of grammatical categories common to the two grammars, plus the combined rewrite rules and lexicons for  $L_1$  and  $L_2$ . The resulting grammar will generate  $L_1$  and  $L_2$  sentences naturally, as well as all occurring intrasentential code shifts. Unfortunately, the free-union grammar will also generate all unacceptable shifts, for there are no constraints on contextual violations: "The basic problem is that the code-switching constraints are . . . conditions on adjacent constituents, but the essence of context-free generation of sentences is that the internal structure of one constituent does not condition that of another. To solve this problem, we must ensure that for any two neighboring constituents whose boundary could potentially involve a code switch, violating one of the constraints, suitable restrictions must already be coded into the symbol for the grammatical category heading each constituent . . . as these symbols are rewritten, the restriction information must be passed on . . . the restriction will be realized by a compatible choice of languages for neighboring lexical terms." (Sankoff and Poplack 1981:15-16). The method chosen is the use of superscripts on the symbols for the various grammatical categories in phrase-structure and rewrite rules, to restrict application of certain rules to symbols with appropriate superscripts. These superscripts appear only in certain derivations and only at certain key nodes, and they carry information which prevents violation of the equivalence principle, namely that juxtaposition of  $L_1$  and  $L_2$

materials may not violate a grammatical constraint in either language. The new grammar involving these constraints contains a lexicon which combines the lexicons of  $L_1$  and  $L_2$  and the grammatical categories of  $G_1$  and  $G_2$ . Given any rule  $r_1$  in  $G_1$ , one looks for any pair of symbols in the output such that there exists at least one translation in  $G_2$  which does not reverse the order of these symbols. If this condition is met,  $r_1$  will be included among the rules of the code-switching grammar, possibly in a number of different versions. Rules stemming from  $G_2$  are included in the new code-switching grammar using a similar criterion. If in the output of a rule  $r_1$  there exist two *obligatory* symbols ordered in such a way as to exclude all corresponding  $G_2$  rules, then the equivalence constraint will prohibit a switch from  $L_1$  to  $L_2$  after the constituent headed by the first symbol and before a constituent headed by the second symbol. The *free morpheme constraint*, stating that bound morphemes may not be switched away from the language of the base morpheme, indicates that if in the output of a  $G_1$  rule there exist two consecutive symbols the first of which represents a morpheme bound to the second or vice versa, a switch cannot occur from  $L_1$  to  $L_2$  between these two symbols.

In practice, the determination of which rules will be included in the code-switching grammar entails the comparison of an enormous number of permutations, since certain basic configurations may have ten or more 'translations' in the other language. Assuming that these permutations can be worked out, and the circumscribed nature of normal colloquial discourse of the type characterizing intrasentential code switching suggests that this computation will be possible, the concrete nature of the grammatical representation is as follows. Each superscript contains a language category and a grammatical category; for example, the Spanish rewrite rule for a noun phrase:

$$[38] \quad NP \rightarrow DET \ N \ ADJ$$

will contain for  $N$  the superscript  $sp:n$  and the adjective will have the superscript  $sp:adj$ , so that when lexical insertion occurs, only Spanish items can be inserted in these slots. The article remains unspecified, however, so that either English or Spanish determiners may occur:

$$[39] \quad NP \rightarrow DET \ N^{sp:n} \ ADJ^{sp:adj}$$

$$[40] \quad my \ camioneta \ grande$$

$$[41] \quad this \ campeonato \ de \ natación$$

This phrase structure grammar constrained by superscripts suitably chosen to reflect observed grammatical restrictions will presumably generate all and only the acceptable bilingual sentences, as well as all monolingual sentences in  $L_1$  and  $L_2$ . There remains the question of the quantitative dimension, since not all switches occur with equal frequency, nor is the frequency of switched discourse as opposed to monolingual discourse in one of the two languages specified by the form of the bilingual grammar. In order to indicate the numerical dimension, Sankoff and Poplack propose the inclusion of statistical values based on the actual observation of code switching among bilingual speakers. The bilingual phrase-structure rules would then take the form of *variable rules*, as used in variational studies.<sup>7</sup> The authors state as their ultimate goal the establishment of "the mechanism for combining probabilities . . . will be the key to truly integrated inductive/deductive research on the relationship between probabilistic monolingual and code-switching grammars. That is, statistical properties of the code-switching grammar will not only be empirically

observable in code-switched discourse, but will also be predictable from the statistical properties of the monolingual grammars." At present, the authors concede that only observation of actual corpora of code-switched data can be utilized to provide the requisite statistical parameters. In view of the inherent dependence of bilingual code switching on social and psychological parameters, it may never be possible to predict the values of frequency of shifting based on monolingual grammars. At the same time, the mere assignment of probability values based on observation of previously produced utterances assumes a randomness of production, circumscribed by overall statistical limits as well as grammatical constraints, but subject to no internal structure which would account for the motivation of individual shifts. This model thus shares the aspect of randomness found in the Dearholt-Valdés model, although the latter model does include several categories which combine to form the total numerical value of probability. The Poplack-Sankoff model is at present the most elaborate phrase-structure model of Spanish-English intrasentential shifting and addresses the problems noted at the beginning of this chapter. In terms of formal generative power, the model appears to meet the requirements of descriptive adequacy normally imposed on linguistic descriptions, providing that some way can be found to unequivocally code all cases of functional translatability of equivalent structures. This does not preclude the continued search for an internal structure motivating the individual assignment of language values, to account in a more adequate fashion for given code switches in specific contexts. This matter will be returned to in the following chapter.

#### SYNTACTIC CONGRUENCE

Clearly it will not be possible to find a non-*ad hoc* representation of restrictions on code switching if separate mention must be made of the grammars of Spanish and English. Rather, some means must be found of stating the restrictions on the bilingual interaction of the two grammars which does not in and of itself presuppose a single undifferentiated grammar, but which does presuppose some metalevel of linguistic structure at which superficial configurations generated by the grammars of Spanish and English are compared for possible combination in a bilingual utterance. It is the formulation of such bilingual syntactic constraints that represents the next step in the evolution of linguistic models of bilingual intrasentential code switching.

One major attempt at characterizing such bilingual interaction was Poplack's (1980) *uniform structure constraint*, stating that juxtaposition of  $L_1$  and  $L_2$  elements may not violate a surface syntactic rule in either language. In considering examples of spontaneous speech of New York City speakers of Puerto Rican origin, Poplack found that fewer than 1% of the code switches violated the uniform structure constraint, although it is not indicated what criteria were employed to judge the putative existence of a grammatical violation. As stated, this constraint needs some clarification, for it is not apparent in which of the various possible bilingual configurations the violation is disallowed. Presumably by definition, the  $L_1$  portion and the  $L_2$  portion will not contain any internal grammatical violations, since taken separately both portions must represent sections of grammatically acceptable monolingual discourse. If one begins with  $L_1$  and then shifts to

$L_2$  at some point in the sentence, then another possible interpretation of the uniform structure constraint is that the  $L_2$  elements must behave as though they were  $L_1$  elements, at least at the point of the switch. Thus in

- [42] *Este es el carro detrás del cual we were parked*

the English phrase *we were parked* replaces the Spanish *estacionamos* or *estábamos estacionados*. However, since the subject pronoun is optional in Spanish, and in the first-person plural *nosotros* rarely occurs in sentences, the inverse switch from English to Spanish would violate the preceding interpretation of the uniform structure constraint:

- [43] *This is the car behind which estábamos estacionados/estacionamos.*

This switch-type appears to be somewhat less common than in [42], although it is by no means nonexistent, at least in the data representing Mexican-American discourse. Similarly, one cannot normally say<sup>8</sup>

[44] \**This is the car that estábamos estacionados detrás with or without nosotros*, since in Spanish it is impossible to end a clause with a preposition. Superficially, at the switch point there is a violation if *nosotros* is employed, since the resulting combination is quite possible, as in

- [45] *This is the car that nosotros trajimos desde México.*

It is rather at the end of the following clause that the violation occurs, for in order to match the stylistic variant begun with the juxtaposition of the antecedent and the relative pronoun in English, the Spanish continuation would have to end with a preposition, an impossible configuration. On the other hand, it is quite impossible to insert the preposition in Spanish at the point of the switch:

- [46] \**This is the car that detrás estábamos estacionados/estacionamos*

- [47] \**This is the car that detrás del que estábamos estacionados/estacionamos.*

The two possible violations take radically different forms. In the case of [44], the switch would produce an unacceptable ending if the Spanish continuation duplicated the English word order immediately following the shift, whereas in [46] and [47] the violation is produced immediately, since although the internal structure of the Spanish sentence may be consistent, a preposition may not follow a relative pronoun in English. It is not clear if Poplack's intent was to include both types of possible violations under the same rubric.

Another possible interpretation of the uniform structure constraint is that the  $L_1$  elements must behave as though they were  $L_2$  elements; that is, the post-switch portion of the sentence is taken as primal. This interpretation would account for impossible combinations such as [46] and [47], since a Spanish initial portion would demand that the preposition precede the relative pronoun. This interpretation would also account for the acceptability of

- [48] *Este es el carro that we were parked behind*

- [49] *This is the car detrás del cual estábamos estacionados/estacionamos*

and even predicts the complete acceptability of [43], despite the problematic nature of this sentence.

The strongest interpretation of the uniform structure constraint would be a combination of the preceding two variations, namely that the  $L_1$  portion must behave as though it were a  $L_2$  structure and vice versa; this interpretation would require complete syntactic identity between the equivalent structures in Spanish and English. This constraint is obviously too strong if left unmodified; for example, it will eliminate phrases of the type [48], although such sentence-types are relatively common.

In the Mexican-American data reported previously, utilizing the least controversial interpretation of the uniform structure constraint, namely that the  $L_2$  portion must behave as though it were in  $L_1$ , 3.33% of the collected examples violated this constraint. However, if we eliminate those cases in which a pause or other hesitation phenomenon preceded the non-equivalent switched segments, indicating less than total linguistic self-assuredness with regard to the possibility of a switch, only .92% of the total indicates fluent violation of the uniform structure constraint.

Sridhar (1980) proposed a slightly different constraint, namely that the internal structure of an inserted  $L_2$  sequence need not conform to  $L_1$  rules, only the end points; this is known as the *dual structure principle*. In practice, this constraint is scarcely differentiable from the uniform structure constraint, since in applying the latter, it is impossible to demand complete syntactic equivalence in the internal portion of the  $L_1$  and  $L_2$  portions. To cite only a single example, in

- [50] *Le pedí que me diera el libro that we had talked about*

the internal structure of the Spanish portion is not matched by the equivalent English construction (*I asked him to show me the book*); whereas the English portion contains a major deviation from the Spanish equivalent, in that the preposition occurs at the end of the phrase. At the same time, the dual structure principle does not explain the impossibility of examples like [44], since in this sentence it is precisely the requirement of a grammatically impossible internal structure in Spanish which accounts for the impossibility of this particular switched configuration.<sup>9</sup> In considering the Mexican-American data, 1.03% of the examples represented violations of the dual structure principle, discounting cases of hesitations or pauses, including all cases of violation of the uniform structure constraint.

Proposals such as the two just described deal essentially with the switching mechanism itself, that is, with the point of the switch and to a limited extent with the discourse preceding the point of the switch. As such, these constraints represent a hypothesis regarding those elements of bilingual discourse which affect bilingual language switching: in addition to the syntactic potential of a given switch point, only discourse preceding the point of the shift is instrumental in establishing the possibility of a switch at any given point. This hypothesis regards the linguistic competence of the bilingual speaker as a finite-state process, which provides a decision procedure for the feasibility of a switch at any given point based on the discourse which has occurred heretofore, but which has no look-ahead power. Despite the fact that some of the possible interpretations of the uniform structure constraint and the dual structure principle apparently entail consideration of post-switch portions of a bilingual sentence, it has been

seen that in certain cases this extension represents too strong a constraint, in that it rules out actually occurring switch-types. At the same time, the available examples and counterexamples suggest a combination of factors both preceding and following the point of a switch which regulate intrasentential language behavior. As an attempt toward considering these combinations and formulating the strongest possible hypothesis consistent with the data, the following proposal was offered by Lipski (1977a, 1978):

Given an underlying semantic representation  $S$ , let  $X_1$  and  $X_2$  be the actual realizations of  $S$  in {languages}  $L_1$  and  $L_2$ , respectively. Furthermore, for any point  $p_n$  and  $X_n$  (where  $n = 1, 2$ ), let  $a_n$  indicate that portion of  $X_n$  lying to the left of  $p_n$  and  $b_n$  indicate that portion of  $X_n$  lying to the right of  $p_n$ . In order to effect a code-switched utterance by combining  $X_1$  and  $X_2$  with a break at  $p_1/p_2$ , it is necessary that  $b_1$  and  $b_2$  be syntactically equivalent.

This hypothesis requires that not only the insertion of the switched material produce no ungrammaticality at the point of the switch, but also that in order for a switch to occur, the major syntactic configurations of both Spanish and the equivalent English sentence be essentially equivalent following the point of the switch. If the alternative sentences would take approximately the same form in both languages, a switch is facilitated, while if a radical syntactic difference would quickly separate the equivalent sentences, switches are less likely to occur. This is not to suggest that speakers are really thinking so far ahead as to have two fully planned sentences in their mind for each sentence they utter, switching at will when parallelism permits.<sup>10</sup> Indeed, the fact that switching only occurs when parallel structures would result suggests that no such pre-analysis has been performed, for if it had there would be no overwhelming reason for prohibiting switches even where parallel constructions were absent. The notion of syntactic parallelism is perceived only vaguely, and only in short, relatively simple sentences, which typify the speech styles in which code shifting is most frequently found. In order to add substance to this theoretical proposal, it is necessary to specify more carefully what is meant by 'essential syntactic equivalence.' Consider the sentence

[51] My uncle Joe es (un) carpintero.

The English version of the post-switch portion would be *is a carpenter*, following the same relative syntactic pattern as in Spanish. A similar case is:

[52] There are many families on the block que tienen chamaquitos

where the English post-switch portion would be *who have kids*, again with the same essential configuration. It is not difficult to imagine cases where such congruence would not hold, since such examples have already been considered. In the English sentence

[53] This is the car (that) we were parked behind

a switch is nearly impossible after *that*, for the reasons just discussed, whereas a switch before *that* is possible, since the Spanish continuation, beginning with *detrás del cual* or a similar phrase, corresponds to the more formal but nevertheless possible English continuation *behind which we were parked*. The impossibility of a switch after *that* is due to the fact that the major constituents following the shift differ radically in their order. English contains NP (subject) + V + PREP, and Spanish contains, in order to make an acceptable relative clause: PREP + REL (+ NP subj.) + V. On the other hand, we have seen that [48] is a possible switch-type, despite the fact that the English post-switch portion differs noticeably from the Spanish equivalent in the position of the preposition. The significant difference is that in this sentence, the internal grammatical structure of the post-switch portion is completely acceptable, as is the point of transition between Spanish and English. Moreover, the only deviation between the English and Spanish post-switched portions comes at the very end of the continuation, in the case of the preposition. Prior to this point, the Spanish sentence proceeds just as the English sentence does, and in fact there are many possible Spanish continuations which would exactly duplicate the English continuation, minus the final preposition, for example:

[54] Este es el carro que vimos detrás de la gasolinera.

The constraint of essential congruence of post-switch constituents must therefore be tempered by the notions of distance from the point of the shift, and with the amount of congruent material which occurs following the switch point before a radical syntactic departure intervenes to break the essential syntactic congruence. In the case of a switch from the English following *that*, there is no way to construct a Spanish continuation which is both grammatically consistent internally and which represents a grammatically acceptable transition from English to Spanish.

There are many other cases where syntactic congruence following the point of a switch may be violated, when dealing with stereotyped syntactic patterns that are equivalent to different patterns in the other language. One common case involves the placement of object pronouns. Consider:

[55] No sé, porque I never used it

where the Spanish continuation would be *nunca lo usé/usaba*, but where the clitic pronoun is not a separable part of the verbal phrase and thus is susceptible to comparison with equivalent English syntactic patterns. Discounting the placement of the pronoun (which is relatively unstressed in both languages), the syntactic structures following the switch point are identical in Spanish and English, and both the English continuation and the Spanish equivalent are internally grammatical. Another case is

[56] They're still meeting at Ripley house every Thursday night y la gente se está juntando ahí

where the English continuation would be *people are still getting together there*, or something similar. The major constituents are identical following the switch point: NP (subj) + AUX + V + ADV, but the internal structure of the constituents varies somewhat. Moving further away from complete congruence, which nonetheless facilitates language switching, consider:

- [57] She told me to make a special dedication to her son, que le dicen el Pachuco de Rosenberg

where the phrase *que le dicen* roughly corresponds to English *that (who) they call*; the major difference is the absence of the subject pronoun in Spanish, since the verb is used impersonally in the plural, and the employment of the clitic *le*. Grammatically, the Spanish continuation does not follow traditional normative models, but it is quite in line with popular tendencies which utilize *que* instead of the combination PREP + *quien* or *cuyo*. A similar example is:

- [58] One more time Ruth, pa que la gente se cuente y then can call you at . . .

Finally, complete syntactic parallelism may fail to hold in the case of short idiomatic expressions or *modismos*, as in the often-cited example:

- [59] It's the first shag que se me hace que looks good on a girl

where the expression *se me hace* corresponds to English *I think, it seems to me, etc.* An even stranger example is:

- [60] Escucharon a Lisa López's latest album, una canción titulada . . .

where despite a Hispanic pronunciation of *Lisa López*, the English possessive pattern was introduced into the midst of the Spanish sentence. This sentence represents a major syntactic violation, presumably triggered by the nontranslatable Hispanic name *Lisa López* in conjunction with the stereotyped expression often used by disc-jockeys in the format "X's latest album." In this case, the same announcer was heard, on other occasions, to use the expression

- [61] El latest album de . . .

There are even cases where a language switch in one speaker so totally orients the interlocutors to that language that the continuity of syntactic parallelism becomes lost. For example, in our corpus, one speaker said:

- [62] I love la música chicana

to which his interlocutor replied:

- [63] Y a Carlos también

thus corresponding to the hypothetical Spanish sentence

- [64] A mí me encanta la música chicana.

Preserving the pattern of English *I love*, the reply would have been *yo también*, but the noteworthy lack of a transitive Spanish verb which is semantically equivalent to *encantar* impeded this hypothetical possibility and produced a response which, while semantically congruent at the discourse level, is syntactically inappropriate.

Internal structures may vary categorically, as in

- [65] It's gonna be Sonny and the Starliners, on the trece de febrero . . .

where Spanish uses the cardinal numeral and English would use the ordinal *the thirteenth*. Articles may also be left out, as in

- [66] Si compran tickets en Saint Josephs, you don't have to go to purgatory

where the Spanish continuation would be *no tienen que ir al purgatorio*, containing several internal deviations from the English pattern. A similar example is:

- [67] We care about other people como los pintas

where English might say *like prisoners*. Short expressions are often introduced virtually as automatisms, regardless of their internal structures, such as in

- [68] So dale, Ana María (where *dale* corresponds to *go ahead*)

- [69] Ahora es puro disco, you know

where in Mexican Spanish the tag *you know* would not be translated, or perhaps would be rendered as *verdad?* although Caribbean dialects might utilize *tú sabes*. Similarly, the use of the filler *este* corresponds to English *uh*, and is not a syntactically classifiable word.

In view of the above examples, it is not unlikely that the mental imaging which precedes the production of any well-formed and semantically coherent sentence, particularly a short one, includes the choice of the most typical syntactic pattern which would realize the idea. Moreover, there is every reason to suppose that there are certain syntactic/semantic/pragmatic clues in the discourse preceding a bilingual switch that hint at the following structures and thus indirectly, at the appropriateness of a switch to the other language. Consider:

- [70] They only use English when they have to, like cuando van de compras.

The *when* of the English portion suggests that this same word or some equivalent adverbial expression will appear later in the sentence; it is also suggested, given the usual word order of the sentences and also the implicit subject of the first portion, that following the equivalent of *when* will come a verbal expression with or without a stated subject, and perhaps an adverbial complement as well. This is the case with Spanish *cuando van de compras*: the English equivalent would be *when they go shopping*, where the two-word verb phrase *go shopping* is equivalent to *ir de compras*. We can hypothesize that even before the sentence has proceeded from the cognitive image to a concrete realization in either language, the speaker/thinker had an image of people going shopping; that is, doing something and not, for example, passively experiencing a situation. In the majority of such cases in both languages such a mental image will correspond to a simple subject + verb (+ adverbial complement). Since this is one of the most frequent sentence-types in both Spanish and English, the transference from mental image to syntactic realization is essentially identical regardless of whether Spanish, English or a mixed version is the output. Thus this particular switch, while not specifically predictable, is at least possible right from the moment of conceiving the thought, and does not rely on a specific well-formed sentence in each language for its realization. Consider also sentence [55]. If the sentence had been continued in Spanish, it would be:

- [71] No sé, porque nunca lo usé/usaba.

The relative order of verb and object pronoun varies, but the order of the adverb and main verb is identical in both languages. Following the conjunction *porque/because*, both languages utilize the normal declarative order. In reality, it is not even necessary to fully specify the remaining portions of a sentence in order to substantiate the acceptability of a shift following the conjunction. In other words, we may postulate that following a conjunction nearly any final clause could theoretically occur, since it is already known implicitly that a declarative order will result in both cases. The subject pronoun is optional in Spanish but if it occurs it will be found in the same place as in English. Adverbs will also occupy the same relative position. Clitic pronouns present the only major difference (although fully specified direct and indirect objects are identically placed in both languages), but in view of their status as unstressed clitics--inherent parts of the verb phrase--it is not as significant that the order of the individual constituents is somewhat different.

In a similar fashion, any sentence containing two clauses conjoined by *and/y*, *or/o* or *but/pero* represents a mental image in which discrete situations are postulated from the outset. As a result, the exact syntactic structure of the second clause is relatively unimportant in determining the feasibility of a switch at the point of the conjunction, since the conjunction represents two equivalent clauses and also two equivalent types of context. Most coordinating conjunctions could just as well be represented by pauses or other inflectional elements and therefore in most cases sentences containing coordinating conjunctions must be analyzed as containing independent sub-sentences as far as the possibility of language switching is concerned. Therefore, it is most prudent to amend the congruence constraint to indicate congruence of major constituents up to and including the element(s) which introduce the following clause occurring in the same sentence.

Additional evidence for a vague but nonetheless real perception of material about to occur in a sentence comes in 'anticipatory embedding' where a switch is caused by a lexical element occurring later in the sentence, which for some reason is untranslatable in the language of the first portion.<sup>11</sup> In most cases, these lexical items are proper nouns or other words which are so language- or culture-specific that no adequate equivalent exists in the other language. Jacobson (1978b:22-3) gives examples like *Sonny's* (the name of a bar), which cannot be translated into Spanish, and which causes a sentence originally begun in Spanish to switch into English:

- [72] Allá en el parque there's a little place  
called Sonny's.

In this sentence the anticipatory switch occurs between the adverbial clause (preposed) and the beginning of the main sentence, although it would theoretically be possible to switch later in the sentence, for example after *called*. Jacobson also notes that certain idiomatic expressions or metaphors may trigger a switch for the same reasons of untranslatability. "to achieve unilinguality within the boundaries of a given syntactic unit." However, while an item like *putter* (in golf) or *take for granted* may not be so easily rendered in Spanish so as to achieve syntactic fluency, it is harder to accept that the word *pimples* in

- [73] Porque tenía como . . . porque he had a  
lot of pimples

represents a concept sufficiently foreign to Hispanic life as to force the sentence into English. Thus it is necessary to avoid circular reasoning, to assume that switches necessarily occur because of the untranslatability or lexical unavailability within a given group of speakers, of certain items. On the other hand in the unequivocal cases where a proper noun or other manifestly unique expression draws the preceding context into a language switch, it is difficult to avoid the conclusion that the speaker's cognitive apparatus has the sentence worked out before it has been produced, and that semantic/lexical information, as well as the congruence of syntactic configurations, may be part of the material that must be manipulated and balanced in order to stabilize the eventual output, which may be unilingual or may be switched in mid-sentence.

In the data studied for the present analysis, 3.16% of the examples were of anticipatory switches, with switches from Spanish to English outnumbering the opposite switches by a margin of 3 to 2. In nearly all cases, the anticipatory switching occurred at the boundary of the largest possible constituent which included the L<sub>2</sub> element triggering the switch, thus vindicating Jacobson's observations and the claims of Barkin and Rivas (1980) with regard to the maximal acceptability of bilingual phrase structures. Examples include:

- [74] Va a haber un benefit at the Starlight Ballroom  
(switch occurs to create a monolingual prepositional phrase).
- [75] Mucha gente no sabe where Manchester is  
(switch creates a monolingual embedded sentence),
- [76] Y también I know where the feedback is comin'  
from (switch creates embedded monolingual clause),
- [77] . . . many girls, cuando tienen sus quinceañeros  
(switch creates monolingual embedded clause),
- [78] There was this guy, que era un vato de México  
(switch creates monolingual relative clause),
- [79] . . . for the altar boys de la iglesia Guadalupe  
(switch creates monolingual prepositional phrase),
- [80] . . . and acaban de escuchar a César Ramírez  
(switch creates monolingual clause),
- [81] el número es . . . so you can call Ruby (switch  
creates monolingual adverbial phrase),
- [82] Todas las palomillas that work at American Hospital  
Supply (switch creates monolingual relative clause),
- [83] Está localizado on seventy-one. . . Avenue H  
(switch creates monolingual prepositional phrase),
- [84] Están aquí from Reagan Senior High (switch creates  
monolingual prepositional phrase),
- [85] Pero, I'm originally from Kingsville, Texas (switch  
creates monolingual prepositional phrase),
- [86] Corré twelve miles for a marathon (switch creates  
monolingual predicate), and
- [87] Not only en el estado de Texas pero en todo Aztlán  
(switch creates monolingual prepositional phrase).

In only a few cases were the anticipatory switches carried out so as to divide the sentence into monolingual segments smaller than the longest possible major constituent:

- [88] para go and see Sonny and the Starliners (a pause following *para*, highlighting the violation of syntactic congruence).
- [89] Estaba advertised in *Low Rider* magazine, and
- [90] On the lookout for el grupo Casablanca.

The fact of anticipatory language switching does not eliminate the constraint of essential syntactic congruence following the switch point, since in nearly all of the anticipatory switches in the present corpus, such post-switch congruence also obtains. It is probably not coincidental that the largest number of such anticipatory switches involves prepositional phrases, where the basic phrase structure is similar or identical in Spanish and English. In the case of a switch after a coordinating conjunction, as in [80], the presence of two syntactically equivalent clauses obviates the requirement of syntactic congruence, although in this particular sentence such congruence does hold except for the idiomatic expressions *acabar de/have just*.

If the constraint of syntactic congruence following the shift point is a legitimate restriction on intrasentential code shifting, we may ask if it is possible to incorporate this anticipatory information in a bilingual phrase structure model such as that of Poplack and Sankoff, which is essentially limited to the constituents immediately preceding and following a shift point. Strictly speaking, a finite-state model such as the context-free grammar cannot take into account material following the point of the switch, since in order to do so, it would be necessary to consider the (nearly or totally) infinite set of possibilities for a continuation of the discourse following any given point. There are several means of formally incorporating look-ahead information in the phrase-structure grammar, none of which is entirely satisfactory. One is the assumption of some sort of cycle of production, by which an entire sentence is first generated and then is scanned for acceptable shift points. In addition to the theoretical problems which have occurred with the notion of cycle in other areas of linguistics, this suggestion presupposes a production time which would have to be significantly increased over the sentence in which no intrasentential switch occurs. Observations do not bear out the existence of such a differential, since intrasentential switching normally entails no increased production time. Another possibility is the establishment of some sort of restriction on the application of transformations, such as subject movement in Spanish or preposition movement in English, which would produce surface structures divergent in the two languages. Such an approach might rule out certain types of shifts, but would be too restrictive, since there is no reason to limit application of these transformations if no language shift in fact occurs. In practice, it is possible to account for most if not all intrasentential code shifts without any formal recourse of looking ahead behind the boundary of the clause in which the switch occurs, since as noted above, the immediately succeeding configurations are normally predictable from the discourse produced up to the point of the switch. Just as the fact that the need to express a certain concept or the upcoming presence of a non-translatable lexical item will often trigger an anticipatory switch, the anticipated presence of certain types of syntactic structures will also inhibit a language shift. Neither process need to be incorporated formally into the grammar, since in the case of an anticipatory inhibition of a shift, a limitation at the point of the shift is sufficient. For example, to

prevent a switch to Spanish in a sentence such as [53], it will be noted that the word *that*, as relative pronoun, must necessarily have the superscript specification *eng:pro* corresponding to the rewrite rule generating the relative clause. At the same time, one may specify that the element following *that* must also be specified as *eng*; since there is no Spanish phrase-structure rule which will generate a subject pronoun as part of the embedded prepositional phrase, a continuation in Spanish will be disallowed. In general, it is sufficient to require that if a switch occurs in the middle of a major syntactic constituent, then the remainder of the constituent be syntactically equivalent in  $L_1$  and  $L_2$ , that is that there be no essential superscript for the remaining elements. Also, it must be required that if a switch occurs at a constituent boundary, the elements of the immediately following constituent have no essential language tagging superscripts. This requires a small amount of look-ahead, but this type of short-range chunking occurs in all types of discourse, and is necessary in such monolingual operations as concordance, as in

- [91] Antes de conocerla, Roberto sabía que María sería la mujer de sus sueños
- [92] Esta no es mi casa.

The basic requirement embodied in the equivalence constraint will itself account for most cases of post-switch congruence by default, since most widely divergent post-switch configurations will automatically be ruled out.

The preceding discussion has indicated the need for going beyond simple grammatical constraints on language switching to include more global notions of syntactic congruence. The construction of an integrated bilingual grammar will need to include the transitional grammatical structures, since it has been demonstrated that the juxtaposition of  $L_1$  and  $L_2$  elements normally respects the syntactic constraints of both languages. It is also evident that some degree of anticipatory analysis does take place, and that a certain degree of syntactic congruence following a shift must also be postulated as a constraining condition on language switching. These global syntactic conditions provide one set of necessary conditions for bilingual language switching, and must be conceded a position of primary importance. The matter of establishing sufficient conditions for a language switch is not addressed through global syntactic congruence, given that in many cases more than one position will produce syntactically congruent patterns. Moreover, considerations of congruence do not address the quantitative nature of certain switching patterns, such as the variable behavior of conjunctions, relative pronouns, articles, prepositions and the like. For possible approaches to these questions, it will be necessary to consider additional dimensions of bilingual language behavior.

## NOTES

1. See for example Wallis and Bull (1950), Bull (1950; 1954; 1965; chap. 17).
2. In a few cases an English adjective may become lexicalized in Spanish, in which case it may follow a noun. For example, *nice*, *cool*, *full*, etc. are often used in this fashion. Midstream switches may occasionally cause an English adjective, which is only partially lexicalized in Spanish, to occur after a noun. Poplack (1980) gives the example *pechos flat*. See also the discussion in Pfaff (1979) as well as the presentation in the preceding chapter.
3. This does not conform to the acceptability judgments reported by Gingras (1974).
4. One fully worked-out example in which random elements have been rather whimsically introduced in order to partially reproduce certain types of professional jargons is given by Beckmann (1973). The results are startlingly realistic, but of course no true semantic generative power is claimed for this model.
5. The claimed correlation between text frequency of a given item and its 'difficulty' in some empirically verifiable fashion was first proposed by Zipf (1965). This notion, despite the controversy that it has aroused, has never totally disappeared from linguistic reasoning, but must be tempered by other constraining factors, or else all languages would reduce themselves to a set of maximally differentiated forms corresponding to Zipf's mathematical formulas and 'laws.'
6. For example Hadlich (1971:63-64).
7. See the complete characterization of variable rules offered by Cedergren and Sankoff (1974).
8. Similar structures in Spanish, ending with a preposition, do occasionally occur in substandard speech of certain dialects. For example, in Sevilla this type of structure has been reported by Carbonero (1982:50-51).
9. Sridhar derived his conclusions from bilingual interactions in India, involving neither English nor Spanish. This may possibly account for the different theoretical conclusions.
10. As apparently understood by Sridhar (1980).
11. This type of switching was first explicitly described by Jacobson (1978b), and has subsequently figured in other studies on code switching. Similar concepts may be found in Clyne (1967), who, however, had grouped this type of code shifting under the general category of 'triggering.'

## The Variable Category Nature of Language Switching

### THE ROLE OF VARIABLE ELEMENTS

Congruent Spanish-English shifts are not distributed randomly, but rather reveal a number of groupings, in which certain words, principally conjunctions, prepositions, copulas, articles and pronouns play pivotal roles, providing entry or exit ports between the two languages. This suggests that the speaker who performs such intrasentential switching is aware, consciously or unconsciously, of which language he is using, or at least of the necessity to maintain the discourse within the same language, except at indicated points. At the same time, it is an observable fact that following a switch, many speakers are unaware of having switched, or of the precise point in the sentence where a switch occurred, particularly if there is more than one syntactically acceptable switch point within the sentence. The fact that speakers will, when presented with sentences containing unacceptable switches reject or even correct these sentences does not contradict the observation that the speakers themselves may not be aware of where a switch has taken place in their own spontaneous production.

Also of significance in the description of bilingual language behavior is the quantitative dimension, the fact that certain elements are switched more or less frequently, while still other elements appear in either Spanish or English with nearly equal probability. In the corpus studied in this investigation, we have the following overall frequencies of shifts, for major variable elements:

- A. Preceding (1.61%) or following (2.61%) the subordinating conjunction *que/that*:
  - [1] I'm not sayin' that son chuecos.
- B. Before (1.21%) or after (1.09%) the conjunction *because/porque*:
  - [2] No podemos hacer nada porque we don't have the power.
- C. Before (2.64%) or after (0.46%) the relative pronoun *que/that*:
  - [3] Escucharon a un señor que has been around for a long time . . .

- [4] She told me to make a dedication to her son, que le dicen el pachuco de Rosenberg . . .

D. Preceding (12.11%) or following (4.02%) a preposition:

- [5] He's from Corpus Christi, con una canción titulada . . .  
 [6] Ben viene con the sports news

E. Preceding (11.77%) or following (4.42%) the coordinating conjunctions and/y, or/o and but/pero:

- [7] Sometimes te pones serio and you know that, you make good points.

Such quantitatively variable behavior has motivated the formulation of generative models which reflect bilingual language switching, but in many cases the models have little explanatory power. For example, in the model of Dearholt and Valdés (1978), a mechanical algorithm assigns probabilities to certain types of shifts and generate them in the absence of context-sensitive factors such as language tagging, syntactic congruence and lexical availability. Sankoff and Poplack's models (1980, 1982), including the modifications suggested in the preceding chapter, incorporate variable rules in the production of intrasentential language shifts, and are based more directly on the calculation of transition probabilities resulting from an actual corpus of data. This model is likely to generate much more realistic utterances. Nonetheless, such a model does not address the question of explanatory value, nor is there an explicit algorithm for language tagging in the bilingual phrase. The bilingual phrase-structure model of Rivas and Barkin (1980) provides for the incorporation of quantitative data, but utilizes an arbitrary method of assigning language values to the entire phrase structure, based on the lexical identification of certain key items.

In a typical English or Spanish sentence, there exist pivotal points which represent a convergence in the mental representations of the two languages, for bilingual speakers. Intrasentential shifts occur with greatest regularity precisely at these pivotal points. Occurring and potential language switches may be grouped into two categories: lexical substitutions and syntactically congruent shifts. Lexical replacement usually involves nouns or idiomatic expressions, less frequently other elements. The L<sub>2</sub> element may be inserted into an L<sub>1</sub> discourse due to lack of familiarity with the L<sub>1</sub> term, if an appropriate L<sub>1</sub> term does not exist, or from psychological pressures induced by a particular configuration:

- [8] Era el winning pitcher  
 [9] Me gusta la onda de Low Rider . . .  
 [10] We have to look out for la raza chicana

The important point is that the inserted L<sub>2</sub> items are effectively bracketed by L<sub>1</sub> elements, enough so as to be regarded as insertions. In congruent shifts, the discourse passes smoothly from one language to the other and remains in L<sub>2</sub> at least long enough to shift the linguistic focus from L<sub>1</sub> to L<sub>2</sub>. The duration of the shifted discourse in L<sub>2</sub> varies widely; the discourse may return to L<sub>1</sub> in the same sentence:

- [11] It's gonna be Sonny and the Starliners on the trece de febrero, que va a ser the day before Valentine's,

Alternatively, considerably more discourse may occur in L<sub>2</sub> before another shift, this time considered independently of the previous one, returns the focus to L<sub>1</sub>. The key defining feature is that following the switch there is every indication that the speaker has transferred attention entirely to L<sub>2</sub> and is no longer merely inserting items in the midst of an otherwise L<sub>1</sub> context. In most instances of a congruent shift, not only is the duration in L<sub>2</sub> considerably longer than in the case of an insertion, but the structural complexity of the L<sub>2</sub> inclusion is also greater. Frequently, the action of transformations may be seen, as in [11], and the requirement of essential syntactic congruence following the switch becomes successively weakened as the distance from the shift point is increased.

Returning to the precise point of an intrasentential language switch, it is obvious that the switch point itself cannot pertain exclusively to one language or the other, regardless of the lexical source of the word or words involved; this is because the elements in question serve as pivotal points or ports between the two linguistic codes, and must therefore pertain to a meta-system of relational elements which, in a sufficiently well-defined context, permit a language shift to occur. The fact that only certain categories of elements permit shifts, and that the linguistic tagging of many elements is variable, suggests the existence of such a relational system, whose categorical definition transcends normal language-tagging of simple sentence elements. On the other hand, to postulate an entirely separate meta-system pertaining neither to L<sub>1</sub> nor to L<sub>2</sub>, solely to accommodate the pivotal points (which in the case of actual words do, after all, belong to real languages available to the speaker), is entirely gratuitous, especially since these same elements in other contexts may not participate in switches. There is nothing intrinsically special, for example, about prepositions, conjunctions or relative pronouns, yet in certain contexts these elements cannot be analyzed as simple sentence constituents, but rather serve as relational units in an integrated bilingual system.

#### MODELING VARIABLE CATEGORY INCLUSION

In order to model the means by which the grammars of Spanish and English (or any two other languages) combine to form intrasentential switches with certain elements constituted as pivotal points, a system is needed in which key units facilitate a language switch through a form of category inclusion, more complex than mere assignment of lexical classifications. This stage of research into the psycholinguistics of bilingual language switching may be represented as a cybernetic black box, in which the input and output of the system are observable, but where the internal mechanism remains hermetically sealed out of the range of direct observation. The system is the linguistic organization of two codes and the process permitting intrasentential language shifts, while the black box is the assignment of appropriate category values to the specific elements of phrases and sentences.

In the formulation of a model which will portray this situation in a fashion subject to empirical testing, model formation and comparison with established research findings, greatest ability may be derived from the mathematical theory of fuzzy subsets. This theory is a generalization of the notion of set inclusion, and has been

developed as a result of dissatisfaction with the practical applications and results of theoretical logic and set theory.<sup>1</sup> In classical (Boolean) set theory, a particular element  $x$  is either a member of a given set ( $x \in A$ ), or it is not ( $x \notin A$ ). This is but a restatement of the logical law of the excluded middle, which is an intuitively essential part of daily routines. For example, a light is either on or it is off, an engine is either running or it is not, a given response in an examination is either true or false, a plate either broke when it fell or it did not, and so forth. Certain apparently mutually exclusive categories, however, may show themselves to be more permeable under closer scrutiny, or when technological advances permit refinement of descriptive methods. It is no longer a simple matter to decide if a person is alive or dead, given current legal and medical controversies, nor have scientists accurately and unequivocally assigned life to viruses. One of the earliest precursors of the formal theory of fuzzy sets is quantum physics, in which the fundamental impossibility of precisely specifying certain physical quantities concurrently with others is postulated, while the advent of quantitative methods in psychology and sociology has prompted further dissatisfaction with categorical models which do not permit variable inclusion in mutually disjunctive sets.

Fuzzy logic allows the coefficient of set inclusion to range continuously between 0 (representing nonmembership) and 1 (representing full membership). The intervening values represent the 'fuzzy' portion of the spectrum, where an element in a sense both does and does not belong to a given set. An elaborate computational apparatus has been evolved in order to turn this intuitively satisfying addition to set theory into a formal calculus, but in order to apply these conventions to particular circumstances, the theoretical notions must be given a substantive content. Fuzzy models have already been applied to linguistic problems, in a qualitative manner, in order to characterize semantic 'hedging' or ambiguous qualifiers<sup>2</sup> and to describe certain grammatical phenomena resulting from the intersection of conflicting structural analyses.<sup>3</sup> In these cases no formal computations were applied, since insufficient data were available, and no causative or explanatory value was claimed for the fuzzy analyses. In the case of Spanish-English language shifting, the variable category nature of language tagging, and the inherent difficulty of unequivocally assigning language values in the face of the shifting at pivotal points, permit the formulation of a fuzzy logical model to portray the categorical instability represented by intrasentential shifts.

#### CONSTRUCTION OF A SPECIFIC MODEL

The construction of one such model is as follows. The entire universe of discourse is assumed to consist of the union of the two languages  $L_1$  and  $L_2$ , represented formally as  $L_1 \cup L_2$ . The complement of  $L_1$  may therefore be defined as  $L_2$ , and vice versa. In the case of Spanish and English, this simplification is unobjectionable, since the use of nonassimilated words not belonging to English or Spanish is virtually absent from bilingual behavior. It may furthermore be assumed that the intersection of  $L_1$  and  $L_2$  (represented as  $L_1 \cap L_2$ ) is empty, i.e., that there are no elements common to both languages, although strictly speaking, in the case of Spanish and English, this may not be entirely true, since there are a few words which even in the speech of monolinguals, may be analyzed as per-

taining to both languages. Note that only the lexicons are being combined, not the grammars, as in the free-union grammar of Sankoff and Poplack (1980, 1982). It is already assumed that the grammatical structures may be directly generated by some type of bilingual grammar, probably along the lines suggested in the preceding chapter. For the purposes of modeling code shifts, these idealizations do not detract from the qualitative value of the proposed model, which is its major point of significance. As a result of this bipartite division, it is then only necessary to define coefficients of inclusion  $c_n$  in one language, since the complement  $(1 - c_n)$  automatically represents the degree of inclusion in the other language. For every element  $e$ , one may define the coefficient of inclusion  $c_e$ , which represents the extent to which  $e$  belongs to  $L_1$ . The complement  $(1 - c_e)$  then represents the extent to which  $e$  belongs to  $L_2$ . In other words, if  $c_e = 1$ , then  $e$  is entirely and exclusively a member of  $L_1$ ; if  $c_e = 0$ , then  $e$  belongs exclusively to  $L_2$ ; if  $c_e = .5$ , then  $e$  is equally shared between the two languages, and so forth. Whereas the lexical items are drawn from the two monolingual registers, the coefficients are derived from the dynamic interaction of the two languages, and are only incidentally related to the language of origin of the items in question.

These nonbinary coefficients of category inclusion may be used to characterize the putative mental representations of bilingual language-switched utterances, and by extension, of the mechanisms which have produced these utterances. In the majority of bilingual discourse in which intrasentential switches do not occur, it is assumed that the coefficients will either be 0 or 1, or close to these points. The case of inserted nouns or (particularly nonconjugated) verbs is more complex, given the range of variation which has been observed. If only one  $L_2$  item is inserted into an  $L_1$  context, this suggests that the coefficient of the  $L_2$  item is relatively low, i.e., that it is regarded as a foreign introduction. It is, however, probably not zero, given that the following context redirects itself to an  $L_2$  orientation. It is also necessary to consider the degree to which the  $L_2$  item has been integrated into the lexicon of  $L_1$ , since some borrowed words have become quasi-lexicalized. For example, the word *pitcher* (in baseball) is effectively lexicalized for Spanish speakers throughout Latin America, even those knowing little or no English. On the other hand, the use of *winning pitcher*, as in [8] above, must be regarded as at least partially a foreign inclusion, since this configuration would never be heard from a monolingual Spanish speaker not living in the United States. Whereas *pitcher* might have a relatively high coefficient in an otherwise Spanish context, indicating that it is lexically incorporated into Spanish;

[12] El pitcher era . . .

*winning* would have a very low coefficient:

[13] El winning pitcher era . . .

Moreover, as a result of the prenominal modifying pattern and the use of the gerund as modifier, the entire pattern *winning pitcher* is so non-Spanish that for most speakers the low coefficient indicative of non- $L_1$  status would apply to the entire combination. Another example is the word *benefit* as in

[14] Va a haber un benefit para . . .

where the word *benefit* is a partially naturalized lexical insertion, in contrast to

- [15] Va a haber un baile, un toys for tots dance.

Here the whole combination *toys for tots* is untranslatable into Spanish, and moreover represents a well-known cultural artefact specific to the United States. Thus, the word *dance* in [15] would be more closely identified with English than the word *benefit* in [14].

In certain cases, pronunciation of key elements may provide a clue as to category inclusion values. If, for example, when speaking Spanish a bilingual speaker pronounces an inserted English word with a Spanish accent, this may suggest complete integration of the word. This pronunciation, however, may simply be a reflection of the speaker's normal habits in English, since many Spanish-English bilinguals in the United States, while speaking relatively fluent English, give it a distinctively Hispanic pronunciation. It is therefore necessary to have accurate information on the command of both  $L_1$  and  $L_2$  phonology for each speaker in order to assess the relative importance of pronunciation in determining category inclusion values.<sup>4</sup>

So-called 'verbal quotes' may provide another indication: when a speaker verbally brackets an inserted  $L_2$  element by a pause, or an exaggerated pronunciation, or highlights the element by a particular facial or body movement, it may be that the word is still identified as foreign to  $L_1$ . Such determination often requires access to information not contained in the linguistic form of the utterance. Another possible means of identifying elements which are felt to be foreign intrusions is via the observation of those units which do not participate in the normal morphological and syntactic patterns of  $L_1$ . For example, in

- [16] Voy a hacer fix la televisión

the word *fix*, at least at first, most probably was a foreign insertion, although in some instances the combination of *hacer* + English verb has been lexicalized to the point where this combination may no longer be analyzed as violating grammatical restrictions.<sup>5</sup> Other examples include:

- [17] Otra vez a las one o'clock, a las una [sic.],
- [18] Quería, uh, make a little commentary, if it's possible,
- [19] No lo dejan que . . . for us to come and visit him anymore, and
- [20] We couldn't come by a visitar contigo, no man.

In these cases, selectional restrictions are violated and/or the inserted words are not modified to fit the morphological patterns of  $L_1$ ; this is clearer when going from English to Spanish, because in the opposite direction there are many cases where the equivalent English expression is normally undifferentiated morphologically, which makes it difficult to accurately characterize the categorical status of such elements, as in [18], and

- [21] Nos dieron la oportunidad de . . . raise some money, recoger fondos.

While from the standpoint of an analysis of a particular utterance the coefficients of category inclusion are assigned to individual words, it is really the syntactic function of the word which is primarily responsible for this coefficient, together with affective variables and

the nature of the following context. In general it is not possible to compute absolute values for category inclusion of particular boundaries, since to do so it would first be necessary to compute the total frequency of occurrence of such phrases, then the frequency of switches involving such constituents, and most difficult, it would be necessary to assign numerical values to the multitude of extralinguistic factors which also constrain or promote switching.<sup>6</sup> For example, highly emotionally charged elements such as *la onda chicana* and *la raza latina* may inexorably drag a discourse from English to Spanish, at least for a while. The occurrence of a proper noun or a highly language-specific common noun in  $L_2$  will frequently cause the following discourse to remain in  $L_2$ , much as the presence of an  $L_2$  proper noun in succeeding discourse:

- [22] Es que Pat was dry.
- [23] Si compran tickets en Saint Joseph's, you don't have to go to Purgatory.
- [24] I'm a Jiménez, todos los demás son Torres.

In the data studied, 4.59% of the switches involved shifting following proper nouns or language-specific common nouns; Spanish-to-English switches outnumbered English-to-Spanish switches 9 to 1.

Also influential is the already mentioned presence of a proper noun or language-specific common noun closely following the point of the shift; an anticipation of this item frequently triggers a switch preceding the actual occurrence of the form:

- [25] Va a haber un benefit at the Starlight Ballroom
- [26] Mucha gente no sabe where Manchester is.

In the data under consideration, 3.16% were anticipatory switches, with Spanish-to-English switches outnumbering English-to-Spanish switches by a margin of 3 to 2.

Shifting forces such as those just described reduce category inclusion values, in a way which at present may not be precisely computed, and which must be combined with other variables, both linguistic and affective, in order to accurately characterize the coefficient of category inclusion of a given element in specific utterance.

Clyne (1967: 86), in the course of an elaboration of the 'trigger categories' which stimulate intrasentential shifts, speaks of 'overlapping words,' that is, words which have the same essential form in both languages. Such a word "causes the speaker to lose his linguistic bearings. He momentarily forgets which language he is speaking, and later utterances are automatically in the other language." Overlapping words in Clyne's sense are characterized by fuzzy category membership, in this case because of morphological configurations which, however, must be bound up with syntactic configurations, since the same word may be ambiguous in one context and unambiguous in another. It is this type of lexical ambiguity that has been exploited in bilingual dominance tests,<sup>7</sup> where an item which is either graphically ambiguous (*dime, capital, pan*) or phonetically ambiguous (Spanish *ay*/English *I*) is presented to subjects to elicit a category membership value. Barkin (1978b:2) speaks of a mechanical or unconscious switching "caused by a trigger word, usually resulting from paradigmatic association." Jacobson (1978a: 21) notes that conjunctions like *que* seem to be below the speaker's level of awareness and surface therefore unwillingly. Herras and Nelson (1972) found clear evidence of inability to recall the language in which entire phrases or sentences were produced.

Remaining at the lexical level, Albert and Obler (1978:215-17) review pertinent findings and indicate that even under the best of circumstances, language tagging is imperfect, although there is a tendency for accuracy to increase with the speaker's age, to a certain degree.

In the case of the Spanish-English findings, there are a few lexical items which may be considered ambiguous enough to constitute trigger words. Among such items, we may mention *chicano*, *barrio* and even *raza*, words now quasi-lexicalized into English, at least among Hispanics in the United States, but which still retain primary cultural identification with Spanish. Other cultural artefacts such as *low rider*, *benefit*, *high school*, etc. may also be considered as lexically ambiguous, since they may not be readily translated into Spanish, and in some instances the coefficient of category inclusion for such items would be altered accordingly. The notion of fuzzy set membership becomes most useful when considering the pivotal points at which syntactically congruent shifts occur, for it is at these points that one finds the strongest evidence of ambiguous category membership. Syntactic couplings like *porque*, *since*, *and*, *para*, etc. lexically belong to a single language, but in view of their relational function and the potential for redirecting the discourse into  $L_2$ , they are prime candidates for fuzzy set inclusion. In a sentence like

- [27] There are many families in the neighborhood  
que tienen chamaquitos

*que* is a fuzzy joiner, belonging neither entirely to English nor (despite its lexical origin) entirely to Spanish. The precise value of its coefficient of category inclusion may not be precisely determined as yet, but despite the fact of its Spanish origin, it is impossible to argue for a coefficient of less than 0.5, since if *que* were replaced by *that* or *who*, a less acceptable sentence would result:

- [28] \*There are many families in the neighborhood  
that tienen chamaquitos.

In this case, the discrepancy is presumably a question of transition probabilities: in English, *that* is regularly followed by a minimally conjugated form which is often identical to the infinitive. In Spanish, *que* usually demands a conjugated form, which by its very nature conveys more information than the English counterpart. The opposite switch sounds equally strange:

- [29] Hay varias familias en el barrio que have kids.

When dealing with adjective phrases, another frequent source of shifts, it is difficult to separate true language shifts from lexical insertion. In most cases, it is easier to shift from English to Spanish than vice versa:

- [30] his big red camioneta  
[31] \*su camioneta red  
[32] \*su red camioneta

Example [32] may be rejected on general grounds of providing too many switches in a short space. The relative position of adjectives in the two languages does the rest. Spanish is normally characterized by a preference for postnominal modification, although some exceptional adjectives also regularly precede the noun. This occurs especially with combinations of several adjectives, to avoid difficult or impossible constructions:

- [33] las principales obras modernas consultadas por los investigadores

Pfaff (1979: 306) postulates that in noun + adjective combinations, it is necessary to match the word order of both the language of the adjective and the language of the head noun; this is essentially the same requirement as in the phrase-structure restrictions of Barkin and Rivas (1980). In English nearly all adjectives occur prenominally; this is often sufficient to define an adjective formed from a noun or a verb: *a talking bird*, *the photocopy machine*, *his dog-in-the-manager attitude*. In switching from English to Spanish between adjective and noun, a possible although less likely syntactic congruence is formed; one could say

- [34] su roja camioneta/pickup  
although [35] would be more common

- [35] su camioneta/pickup roja

On the other hand

- [36] \*his truck red

would be impossible in English; thus by beginning the phrase with *su camioneta*, the speaker has precluded the possibility of inserting a simple adjective. Only an appropriate prepositional phrase or a clause, following the same relative order in both languages, is possible as a continuation:

- [37] su camioneta that he just got fixed . . .

In such a case, it is not immediately clear how coefficients of category membership are to be defined. In a sentence including the phrase *his big red . . .*, there is nothing in the context to suggest fuzziness of category membership, other than the fact that a similar sentence could be formed in Spanish. It is not until the moment of lexical insertion of a Spanish word such as *camioneta* that the context shifts to Spanish. If *camioneta* is considered by the speaker, at the time of the utterance, to be a lexicalized borrowing originating in Spanish, then a switch has not really occurred and there is no reason for the context to switch to Spanish. If, on the other hand, the word is used with full awareness of its membership in Spanish and its non-inclusion in most styles of English, then a shift has occurred and the coefficient of *camioneta* must be 0.5 or less. What would condition the difference in coefficients for the same item in the same syntactic and semantic context? The only available differentiating factors are pragmatic and affective, which underscores the impossibility of providing a purely mechanical algorithm for assigning category values in a sufficient fashion. In this particular instance, one may propose a multi-stage process. Assuming that the discourse continues in Spanish, the first stage is lexical insertion. If a word is felt to be entirely Spanish then one is dealing with simple lexical replacement. The same word, however, may shift to a higher coefficient nearer to the value of 0.5 needed to facilitate a language switch, if the combination of a permissible environment and the mental frame of mind necessary to shift languages is present. The value of the coefficient of *camioneta* would be a composite arrived at by combining the following partial values: the category value of the item in absolute terms; the intent to switch languages or merely to insert a foreign word, and the permissiveness

of the grammatical environment toward a congruent shift. In the above case, we assume that in absolute terms *camioneta* has a relatively low value, signaling that it belongs to Spanish. If one assumes that the coefficients of strength of the scale of syntactic permissiveness and willingness to switch increase according to the parameters involved, these factors would bring up the coefficient of the item. In a properly weighted mixture, if the resulting coefficient is around the hypothetical value of 0.5, a switch would be facilitated.

This latter approach may be tied directly to the diacritic markings of the Poplack-Sankoff model discussed in Chapter Four, since from the outset one may assign category values of 1 to elements in a phrase structure all of whose elements are marked as *eng.*, and a value of 0 when all elements are characterized as *sp.*, or vice versa. When a bilingual phrase structure exists, this simple procedure is inadequate, because the presence of a code shift in the phrase entails the application of variable category coefficients. However, the very existence of elements in a rewrite rule unmarked as to language indicates the need to consider these elements as possessing intermediate category values. Such values will ultimately be described by an interval of values rather than by a unique coefficient, with the interval being gradually reduced to a single asymptotic value by the other constraining factors mentioned above.

We may represent the internal structure of the coefficient of category inclusion as a formula consisting of positive and negative quantities. The positive quantities act to maintain or raise the coefficient of category inclusion, that is, to regard the possibility of a shift. The negative quantities in turn reduce the category inclusion value and thereby facilitate a switch. The positive quantities identified so far are:

- the presence of an essential  $L_1$  superscript for the element in question.
- a resulting syntactic incongruence in the constituent immediately following the element in question.
- the inverse of the distance from the last preceding  $L_2$  to  $L_1$  switch.

The negative factors thus far identified are:

- membership in the group of variable-category or 'overlapping' words.
- presence as first element in a major constituent.
- a lexical retrieval or pronunciation of the given item as in  $L_1$ .
- the presence of an  $L_2$  proper noun in the context immediately following the given element.
- immediately following an inserted  $L_2$  noun.
- an overriding affective configuration which urges a shift.

In any given case, the relative importance of each quantity will vary in accordance with the parameters which have been described. At present, research has not progressed to the point where the variables may be adequately quantified, but this should eventually be possible in most instances. For the factors which raise the category inclusion coefficient, (a) and (b) almost axiomatically impede language shifts, in a fashion which may be at least quantitatively reflected, whereas

relative distance from preceding shifts is a more flexible parameter, and is interrelated most closely with the factors which reduce the coefficients. Among the latter factors, (u), (v) and (y) should be susceptible to precise numerical calculation, whereas (x) may also be calculated as an inverse function of distance. Determinable only after the fact and only in an approximate fashion, (w) and (z) remain problematic. Just as in the Dearholt-Valdés (1978) model, we may point to these factors as the 'fuzziest' points in the present scheme.

In summary, the coefficient of Category inclusion  $c_e$  for an  $L_1$  element  $e$  may be graphically represented as:

$$c_e = a_e + b_e + c_e - u_e - v_e - w_e - x_e - y_e - z_e$$

Since it is impossible to assign exact numerical values to affective variables such as attitude and connotation, we must compute the relative values of category inclusion, that is, the ease with which switches are permitted at certain boundaries. Moreover, it is possible to compute relative values of category inclusion for elements which may occur in either language during a switch. Examples include conjunctions, prepositions and relative pronouns. Table 5-1 contains the appropriate data for the language-switching examples reported in earlier chapters.

TABLE 5-1

## Some calculated values of relative category inclusion

Category	% Probability in $L_2$
And/or/but	72.7
Because/porque	52.5
Copula	27.8
Other subordinate conjunctions	44.4
Preposition	75.1
Relative pronoun	85.2

The figures of category inclusion are only relative; in other words, they state that if a switch takes place, the relative probability of the indicated elements appearing in a given language is reflected by the coefficient in question. No computation is given of the relative probability that a switch will occur at all at these junctures. Poplack (1980:604) hypothesizes that the frequencies of switches at particular types of boundaries would simply reflect the frequency of any given combination of constituents. The data collected in the present study also suggest a similar conclusion. For example, the frequency of coordinating conjunctions in the speech chain is quite high, as is that of prepositions, and to a lesser extent, of relative pronouns and subordinating conjunctions, and all of these configurations figure prominently among the most common intrasentential shift points.

Despite the fact that coordinating conjunctions conjoin syntactically equivalent clauses, the data in Table 5-1 indicate a significantly higher probability of the conjunctions occurring in the language of the second clause, which appears to belie claims that these conjunc-

tions approximately straddle the lexical categories of English and Spanish. The apparent anomaly disappears if we consider that in many cases, a pause or an inflectional punctuation occurs immediately before the conjunction which separates the two clauses. In the case of *but/pero* this punctuation serves to contrast the two clauses, whereas in the case of *and/y* and *or/o* the pause is optional, and is more closely related to the degree of unity which the speaker perceives between the conjoined clauses. If one resists the data to remove cases of pause or inflectional juncture before the conjunctions, the approximate values of category inclusion are as given in Table 5-2, indicating a more balanced categorical status for the conjunctions.

TABLE 5-3

Corrected category inclusion values for conjunctions  
(code switches with no hesitation phenomena)

Category	% in $L_2$
And/or/but	61.4
Because/porque	53.1
Other subordinate conjunctions	46.9

A prepositional phrase is normally a unitary construction which serves a modifying role, and it is not surprising that there exists a strong tendency for the preposition to be in the same language as the remainder of the phrase. It is in fact rather surprising that in one fourth of the observed cases, this inherent unity is broken by a language shift between the preposition and the remainder of the phrase; this variable behavior reflects the dual status of a preposition within a sentence. On the one hand a preposition is the head element in a modifying clause, and as such may not be easily separated from that clause. In more global terms, a preposition is a relational element linking its object and another unit of the main sentence. As a relational element, its status is similar to that of a conjunction, and variable category membership is therefore less unexpected.

A similar assertion holds in the case of relative pronouns. A relative pronoun is an inherent part of a relative clause, which it introduces, and consequently there is a strong categorical bond between the relative pronoun and the language of the remainder of the relative clause. At the same time, a relative pronoun differs from other pronouns in having its antecedent expressed in the same main sentence, and there is a categorical bond between the pronoun and its antecedent, although attenuated by an intervening elements. Of more direct relevance in the case of relative pronouns is the existence of differing transitional probabilities, since Spanish demands a fully conjugated verb following a relative pronoun, whereas the morphological differentiation is considerably less in English.

The case of *because/porque* and other subordinating conjunctions is the clearest example of unstable category values, since the data show these values to be nearly equally divided between Spanish and English. There is no elegant explanation for the fact that subordinating conjunctions show a weaker bond with the language of the

rest of the subordinate clause than do coordinating conjunctions, although the answer probably lies in a combination of intonational phenomena and transitional probabilities.<sup>9</sup> Subordinating conjunctions are less frequently set off by pauses or inflectional punctuation, and the word order following the conjunctions is often essentially the same in Spanish and in English.

The copula, despite its status as a grammatical equivalence marker, is nonetheless a verb, and is bound closely to the subject of the clause in which it occurs. For this reason, the copula occurs more frequently in the language of the subject than in the language of the predicate nominative, although the discrepancy is not overwhelming. This same categorical instability of the copula is evident even in monolingual Spanish discourse, where the copula exhibits variable concordance with the subject and the predicate nominative in conjunction with a number of contextual factors.<sup>10</sup>

The figures in Table 5-1 and Table 5-2 are not meant to have explanatory value, since they reflect computations performed on a finite corpus of observed data. The numerical coefficients are different in nature from those characterizing variable rules, however, in that they operate on a different meta-level and refer to different aspects of quantitatively variable phenomena. A variable rule describes overall frequencies of application of a given process; in the case of language switching, the variable-rule models described in the preceding chapter place numerical values of probability on the occurrence of a given type of switch. This does not only require calculating all the instances where a similar context does not contain a code switch, but in order to accurately formulate variable rules, it is also necessary to consider the pragmatic factors which influence language shifting over expanses larger than individual sentences. Characterizing language shifting in terms of variable category inclusion offers no statement as to the overall frequency of a shift at a given point, except indirectly given the fact that only coefficients around the hypothetical middle point of 0.5 will actually be conducive to a bilingual switch. The coefficients of category inclusion indicate the relative probability of language assignment, given the presence of a language shift. Variable rules and variable category inclusion values are complementary and mutually dependent. A variable rule is sensitive to fuzzy category inclusion, since if a given element  $e$  is unequivocally identified as either  $L_1$  or  $L_2$ , by definition it will not participate in a language switch. One might propose a strong hypothesis for the interrelationship between variable rules and category inclusion by stating that a variable rule of bilingual intrasentential language shifting will apply just in case the category inclusion value of the appropriate elements is exactly 0.5. Given the difficulties expressed above regarding the numerical calculation of category inclusion values, this assertion is somewhat tautological, but qualitatively the relationship seems beyond objection. Category inclusion values are at least initially determined by calculating the frequencies of certain types of language switches, that is, switches which have resulted from the application of variable rules.

## POSSIBLE EXPERIMENTAL VERIFICATION

In order to refine the notion of variable category membership to the point where accurate numerical characterizations may be made, and to give the notion of variable rule more explanatory power, it will ultimately be necessary to go beyond passive samplings and engage in experimentation in controlled situation. Introspective evaluation of switching is generally inadequate, and may actually be misleading, since speakers are not always aware of their own linguistic production, or have false ideas of the acceptability or frequency of certain configurations. A few early studies of language switching behavior have attempted to supplement observed data by presenting native speakers with putative language-switched utterances and asking them for judgments of acceptability.<sup>11</sup> Some of the test sentences were based on actually observed utterances, while others contained structural manipulations introduced by the investigators in order to probe the boundaries of acceptability. However, a typical response of a linguistically naive informant asked by a linguist if a given form is correct or acceptable is to accept anything, under the assumption that the linguist must know what is going on and would not waste time with incorrect examples. This is especially true in the case of informants with little formal education, who often regard university-trained investigators with a mixture of suspicion and awe, especially if the investigator is not native to the area or even to the country where the investigation is being carried out. It is probable that many linguistic investigations have been distorted by this peculiar tendency of informant response. An alternative informant response is to deny the acceptability of items which form part of the speaker's own repertoire, whether or not this is consciously recognized. Speakers who respond in this fashion are often aware of societal norms of stigmatized forms, and may even engage in hypercorrection to avoid producing what they perceive as unacceptable or 'uneducated' variants.

Another even more unreliable method of eliciting additional data on bilingual language switching is to request that informants spontaneously engage in intrasentential switching. While this behavior may be typical of nearly all speakers of a linguistic community, it is by nature a spontaneous and unrehearsed form of linguistic behavior, and as soon as an element of self-consciousness is introduced, the naturalness immediately disappears. Few code-switched utterances are produced, and those that do occur are usually overburdened with pauses, hesitations and obviously unnatural constructions.

One indirect method of testing the category inclusion models involves repetition of sentences in which switches occur, or identification of switch points, in order to determine how accurately the points of switching can be identified and reproduced. In order to illustrate the potential feasibility of this line of research and to add validity to the numerical coefficients of category inclusion derived from the collected corpus of data, a pilot experiment was carried out, with the purpose of isolating some major points of categorical instability in bilingual Spanish-English sentences.

A repetition test will determine average values of variable category membership of certain sentence elements, particularly those which straddle nodes or represent other transitional points. The working hypothesis is that pivotal elements in a sentence in which a language shift occurs may be produced in either language, subject to the general grammatical constraints which have been isolated, and

therefore that the perception of such a sentence should evidence some instability as regards the language of the pivotal elements. In order to probe this hypothesis, a list of test sentences was prepared (see appendix to this chapter) in which language switches occurred in conformity with already observed bilingual sentences. Included were monolingual sentences in both English and Spanish, sentences in which an L<sub>2</sub> lexical item has been inserted into an L<sub>1</sub> sentence, sentences in which the shift occurs at a node where no pivotal element is present at the node, and shifted sentences where a pivotal element is present at the syntactic node. The key elements to be tested were conjunctions, prepositions, articles and relative pronouns. In addition, a few deliberately contrived sentences were inserted in which the switched format was not in conformity with observed restrictions. The key sentences, were introduced into a series of 'dummy' sentences.

The test sentences were read onto a tape by a Texas native of Mexican background, who is bilingual and who typically engages in language switching in informal conversation. Several readings were made and the tape was edited to provide the most natural sounding utterances.

The subjects of the pilot experiment were 10 university students of Mexican-American background, either natives of Texas or having spent most of their lives in that state, and who reported engaging in language switching and/or who had actually been observed to do so. The experiment was explained as a test to determine the ease of reaction to stimuli in Spanish and English. The subjects were told that they would hear some sentences in Spanish and some in English, although no mention was made that some sentences would contain a language shift. After listening to a sentence, subjects had to wait 15 seconds and then had to repeat the sentence from memory. None of the subjects had known exposure to classes in psychology or linguistics that might have led to attempts to out-guess the experiment. The waiting time was determined by trial and error, since repeating immediately after the stimulus sentence usually provided for nearly perfect recall regardless of the configurations of the sentence, while too long a waiting period led to errors that were more and more random, tending to mix up lexical items, sentence constituents and even the overall meaning of the sentence in ways having nothing to do with bilingual structures.

Table 5-3 gives the rates of correct identification for the test elements, and indicates that, as predicted, instability of category identification does occur with pivotal elements in bilingually switched phrases. Comparing the results of Table 5-3 with those of Tables 5-1 and 5-2 shows a significant correspondence between the data representing the corpus of received speech and the results of the repetition experiment, despite the fact that the latter contained a substantially smaller set of examples. A larger scale experiment should bring the figures into even closer correspondence, and will permit the refinement of relative values of category inclusion for pivotal elements.

The results of this pilot experiment are preliminary and rudimentary, since only a few categories were involved, but they do suggest some possible avenues of approach to the cognitive structures underlying bilingual language switching. First, there is evidence that the underlying prelexical grammatical structure of the sentence is sensitive to category inclusion values, in that certain nodes are facilitative of intrasentential shifts. At the same time, the linear surface

TABLE 5-3

Rate of correct category identification of switched elements (%)

Category	Language	% Correct Identification	N
and/or/but	L <sub>1</sub>	64	50
and/or/but	L <sub>2</sub>	92	50
preposition	L <sub>1</sub>	72	50
preposition	L <sub>2</sub>	84	50
relative pronoun	L <sub>1</sub>	48	50
relative pronoun	L <sub>2</sub>	92	50
because/porque	L <sub>1</sub>	64	50
because/porque	L <sub>2</sub>	58	30
other subordinate conjunction	L <sub>1</sub>	70	30
other subordinate conjunction	L <sub>2</sub>	80	30
copula	L <sub>1</sub>	86	50
copula	L <sub>2</sub>	52	50

production of the sentence also enters into consideration. It has been observed that a degree of superficial syntactic congruence is needed between sentences in each of the two languages in order for an intrasentential shift to occur readily. Moreover, the presence of shifts is tempered by superficial configurations which are difficult if not impossible to reconcile with the exclusive determination at a deep structural level. For example, the optional elimination of subject pronouns in Spanish, presumably carried out by a transformational operation, critically determines the possibility of switches to or from English. Similarly, the optional elimination of *that* in English constrains many potential shifts. Many superficial combinations, such as DET + noun are at least partially constrained by the form in which elements are deployed across a sentence, in ways which more nearly approximate the surface bracketing of a sentence as a linear string than the two-dimensional nodal representation.

These results are promising in the direction of ultimately establishing a quantitative dimension in the area of variable category inclusion. It is equally clear that for elements which are not so highly variable as conjunctions and prepositions, additional methods will have to be devised to incorporate quantitative values for categorical inclusion and instability.

## NOTES

1. For a formal introduction to fuzzy logic, see Kaufman (1975), Zadeh (1975).
2. Lakoff (1973). See also Arbib and Manes (1975).
3. Lipski (1980a).
4. Cf. the discussion in Lance (1975:149-141). It may, however, mean that the speaker simply has poor pronunciation of English. Anisman (1975) has provided examples where prosodic features of the speech of English-Spanish bilinguals serve to define a switch as opposed to a lexical borrowing.
5. Cf. Reyes (1976) for a discussion of *hacer* + English verb constructions and their integration into bilingual grammars.
6. This has been done, at least partially, by Poplack (1980) and Sankoff and Poplack (1981).
7. For example, Keller (1974).
8. This example comes from Gumperz and Hernandez Chávez (1975).
9. For an example of the intonational parameters associated with subordinate clauses in Spanish, see Contreras (1980).
10. Cf. for example Gili Gaya (1961: chap. 2).
11. For example, Timm (1975).

## APPENDIX

### Key sentences in repetition test

Va a llegar Ana and she's going to stay for a week.

I know there's going to be a party next week pero no sé dónde.

My uncle is the vice-principal of the school o algo así.

Margarita salió con Jaime and she said they were going to get married.

Freddy Fender writes all his own songs y toca la guitarra también.

El curso se llama español para secretarias y they're going to offer it this term.

Ricardo es muy feo pero all the girls are crazy about him.

Me gusta la música ranchera y I buy all the records that come out.

Are you going to turn on the television or ¿qué vas a hacer?

Lupe me llamó y she said that she'll be arriving tonight.

Ninfas es uno de los mejores restaurantes in the city of Houston.

Allá viene Carlos with his new Pac-Man t-shirt.

I'm going to read this letter para todos mis amigos en San Antonio.

Dicen que El Salvador esta gobernado by thirteen families.

Cuando estábamos jugando Roberto lanzó la pelota over the fence into the street.

Va a haber una junta con all the important people in the city government.

Esta canción está dedicada para all my friends in Huntsville.

I'm writing a letter to un amigo que está estudiando en California.

I saw that program on el canal 23, que viene de México.

Ayer vino me hermanito con a brand new pair of cowboy boots.

This package is from my best friend, que se llama Ronny Barrios.

Este es el nuevo disco that I bought last weekend in Corpus Christi.

Do you know that guy que le dicen el gorilla de Lamar?

I know a lot of people que ahora viven en el otro lado.

¿Cómo se llama ese carrito that has the little hatch on top?

El es uno de los mejores músicos que are part of the current Houston scene.

Ramon es una persona que has really been around a lot.

Este es el libro que my mother sent me last Christmas.

This is one of those video games that tienen mucha onda ahorita.

Who is that tall guy that todos dicen que es mariguano?

No voy a salir ahorita because it's raining too hard.

Me tiene envidia because I'm smarter than he is.

We can't do anything now porque estamos sin carro.

You can't go in that club because no tienes la tarjetita de miembro.

Los pobres no pueden hacer nada porque they don't have any power.

Your car is better than mine because no consume tanta gasolina.

I'm telling you this para que aprendas de una vez.

Me dijo ayer that he was going to come this afternoon.

I didn't know que te ibas a casar este verano.

Te estoy mandando las fotos para que you can see what we all look like.

Jesse me dijo que we could visit him any time we wanted to.

They don't understand that hay que tener orgullo en las cosas que hacemos.

Este libro is my new bilingual dictionary.

My big brother Rafael es un sargento en los marines.

Our phone number es siete cuatro nueve, cuatro ocho ocho tres.

Estas nuevas canciones are the ones that you see on all the jukeboxes.

I like salsa music because it's poderosa, tiene mucho ritmo.

El dia trece febrero es the day before Valentines.

Juanita y Carmen son the best cheer leaders on the squad.

No me gustan los chicharrones porque son real slimy and greasy.

Ese hombre es the guy who closes up the store at night.

These records are pura música latina. pues así me gustan.

## Code Switching in Literature

### SPEECH AND LITERATURE

An important dichotomy in the study of bilingual language switching separates the spontaneous speech of bilingual individuals and language switching as reflected in written documents, in which case it becomes a rhetorical or literary device. This latter, literary phenomenon, is found most frequently in the works of some contemporary United States Hispanic writers. An analysis of written code switching may be of great value in tracing the psychological variables that come into play, and promises to provide a broader perspective of the affective values of language mixing. Language switching in literature is not the result of confusion or inability to separate the languages, nor even of general lexical unavailability; but rather stems from a conscious desire to juxtapose the two codes to achieve some particular literary effect, in turn presumably reflecting an inner drive which cannot find ready expression by remaining within a single language. On the other hand, because of the fact that written documents--particularly those classed as literary--involve not only conscious reflection but also the inherent correction, editing and rewriting process that accompanies acts of writing, such texts may not be used as specimens of naïve, spontaneous linguistic production. Writing involves a clear self-consciousness, comparable to the linguistic self-consciousness found in stressful situations such as interviews, radio and television announcements, and therefore does not represent the uncontaminated output of the speaker's internal linguistic mechanisms. While perhaps such edited and pondered phrases may be closer to the neo-traditional notion of 'linguistic competence,' they provide only a limited amount of usable data about the active and passive filtering mechanisms in the cognitive apparatus, which account for the vast jumble of mixing, analogy, slips and other forms of linguistic left-overs. These 'performance phenomena,' written off by theoretical linguists, often provide the only real insight into the internal workings of the bilingual mind.

Despite the obvious limitations inherent in the use of literary examples of bilingual code switching, these have received a great amount of attention, due to the overlapping domains of study, repre-

sented by linguistics and literary analysis. In an attempt to incorporate the corpus of literary bilingual code switching into the scope of theoretical discussions, two questions must be addressed. The first concerns the extent to which the code switching exhibited in literature (or in general, in any written format) is truly representative of the norms of the linguistic community it purports to address or depict. The second question is the potential value that may accrue to literary code switches which are manifestly deviant from observed societal norms. While no one has claimed that written texts are equivalent, as sources of data, to spontaneous utterances, there is nonetheless a considerable amount of linguistic, psychological and aesthetic information to be obtained from a careful consideration of code switching in its literary form.

English-Spanish alternation is manifested primarily in literature of United States Hispanics of Mexican and Puerto Rican origin. Within these groups, literary code switching is most common in poetry, less so in narrative and least frequent in essays, although individual exceptions may be found. A key feature of bilingual writing is that it presupposes a readership not only capable of understanding the texts, but also whose own linguistic behavior in some way is reflected by the language of these texts. Mere comprehension is not sufficient, although it is necessary for the literary aims to be achieved. For example, Emilio Díaz-Valcárcel's *Figuraciones en el mes de marzo* contains much pseudo-code-switched material and neologisms formed on English patterns that could be understood not only by Puerto Ricans living in the United States but also by those in Puerto Rico, and by other readers who know both Spanish and English. At the same time, virtually none of the examples represents types of code switches occurring in bilingual communities in the United States. Moreover, such spontaneous code switching is all but nonexistent in Puerto Rico, where the majority of the population is only functionally monolingual and where knowledge of English is confined to written expression. Díaz Valcárcel's intent was not to reproduce spontaneous speech patterns, but rather to comment on the unconstrained penetration of American linguistic and cultural elements in Puerto Rico, and the chaotic patterns of social pretension which it entails. Similarly, *Se está haciendo tarde (final en laguna)* by the Mexican novelist José Agustín contains humorous mixing of English and Spanish which does not reflect the speech habits of any known bilingual community, and yet which is readily understood by those whose linguistic repertoire includes familiarity with both colloquial Mexican Spanish and colloquial American English. In this case, the setting of the novel in Acapulco precludes the existence of a stable bilingual community, at the same time highlighting the tremendous incursions of English in Mexico due to the presence of large numbers of American tourists. Still, the characters of Agustín's novel are mostly Mexican, whose bilingualism is of a more intellectual variety, although tempered by contact with Americans. The readership of this experimental novel is similarly assumed to belong to the intellectual elite, familiar with both English and Spanish and accustomed to reading novels set in the expanded format of the *nueva narrativa*.

Leaving aside such intellectualized attempts at literary code switching, which have come from writers whose principal works have been set outside of the United States, most writers of bilingual literary texts come from communities where bilingual code switching is the norm, at least in colloquial contexts. It is therefore legitimate in

principal to search for examples of linguistically valid alternations in the literary output of such writers, albeit tempered by the element of self-consciousness and self-editing feedback already mentioned. It is a mistake, however, to assume automatically that all examples of code switching in literature, especially in poetry, represent specimens likely to be spontaneously produced in non-literary circumstances.

#### ANALYZING CODE-SWITCHED LITERATURE

Keller (1979:268) defines two assertions with respect to literary code switching. The strong assertion, which would be difficult to sustain, is that *all* literally acceptable code switching would need to sound natural to the members of the bilingual community which it purports to represent. The weaker assertion, preferred by Keller, allows for the literary acceptance of code switches which create 'powerful bilingual images' and which sound acceptable to members of the bilingual community, even though in active production modes such forms might never occur. Keller thus proposes to extend the analysis of Valdés-Fallis (1976a, 1977), who places more emphasis on the application of the strong hypothesis, that is, deviations from societal norms in code switches as detracting from literary value. While Valdés never demands application of the stronger hypothesis, she does include it among preferred criteria for literary evaluation of bilingual poetry.

Even a judgment of the accuracy of literary code switches is difficult in practice, since the only method of verifying the existence of such a configuration in the bilingual community, other than the fortuitous circumstance of finding a pre-existent example in another linguistic corpus, is to present the literary examples to members of the speech community for approval, asking them in effect if these examples sound like something they or their acquaintances might say. Such a methodology is notoriously inaccurate, since a large number of speakers will accept, in passive tests, materials which deviate substantially from observed norms of the linguistic community. It is not uncommon for the speakers to accept as 'correct' anything that they can reasonably understand, regardless of appropriateness, commonness or regional acceptability. An equally common response is to reject as 'unacceptable' virtually all examples containing language shifts, on the grounds that such behavior is socially stigmatized and is therefore not to be encouraged. In practice, each investigator of bilingual code-switched literature has employed personal criteria of acceptability, perhaps obtaining additional verification from colleagues. As a result, except for broad categories which verge on being truisms, there is no accurate measure of the extent to which code-switched literature deviates from societal usage. This in and of itself might be sufficient reason for assigning literary examples of code switching to a position subordinate to or at least secondary to spontaneous utterances, in writing theoretical grammars. It is only by extrapolating from established corpora of spontaneously produced utterances that linguistic criteria applicable to literature may be derived; to proceed in the opposite direction is to put the cart before the horse.

### THE LITERARY VALUE OF CODE SWITCHING

In accepting the caveat that literary examples do not necessarily reflect societal usage, one still faces the thorny philosophical problem of the relations between literary norms, aesthetic values and grammars of natural languages. Although this issue is peripheral to the structure of bilingual grammars, it does form an essential part of the discussion surrounding bilingual language shifting, and as such must be treated, however briefly.

Keller and Valdés-Fallis speak of the redeeming value of 'powerful bilingual images' which may be formed from code switches that are deviant with respect to actual societal patterns. Defining precisely what is meant by a powerful bilingual image, and how to decisively recognize one if it occurs constitutes the key difficulty in spanning the gap between linguistic analyses of literary texts and linguistic analyses of natural language texts. Keller (1979: 270), elaborating upon the work of the Russian formalists and the New Critics, states that literary texts . . . while they may be related to contexts, are separate from them. They are complete in themselves, and their textual significance is accordingly enclosed within the limits of the form they take. Expanding upon this notion, he further remarks (275) that the literary texts create a secondary language system that is the result of organizing deviations from the communal language into patterns that are critically discernible in those very texts . . . the text subsumes, rather than replaces, the functions of what in sociolinguistics would be distinguished as code and context.

More recent approaches to literary analysis which have not yet been systematically applied to Spanish-English bilingual literature, might, however, not be so tolerant of the idea of textual hermetism and closure. The Marxist perspective which underlies much contemporary criticism of Hispanic literature and also of language usage in the United States and other areas,<sup>1</sup> would vehemently deny the possibility of separating the texts from the social structures which permitted their composition or even forced them into existence. Similarly, a deconstructionist approach, which has not yet been attempted for bilingual code switching per se, would undoubtedly view such language behavior as undermining the categorical singularity of normal discourse, which in turn must be related to the psychological profile of the authors and readers of such works. Even an orthodox structuralist approach, in delineating the mechanisms of textual production and coherence, would have to step outside the text in order to give substance to the inherent oppositions and transformations.<sup>2</sup> This is not the proper place to pursue the theory of bilingual code switching in literature; one may only signal the need to augment the New Critical approach with a recognition of the unique circumstances which have led to code-switched texts.

Those investigators<sup>3</sup> who have applied linguistic criteria to literary code switching are certainly correct in maintaining that one is first and foremost confronting an aesthetic artefact, which must be evaluated as such, and not solely as a natural language text. However, we must not seal off the literary artefact entirely from sociolinguistic criteria, for to do so would be to create the paradoxical configuration of two grammars, one for literature and one for normal speech contexts, with the reader (who is assumed to possess, perhaps, only latent abilities as a literary writer) somehow bridging the gap between the two grammars in order to relate the literary message

to the environment in which he lives. Even in the case of highly deviant poetry (e.g., Cummings, Dylan Thomas) or prose (James Joyce), it is difficult to justify the existence of an entirely separate literary grammar to generate the texts in question.<sup>4</sup> In the case of bilingual literature (poetry or prose), which comes much closer to societal norms for bilingual communities, such a contention would be almost impossible to seriously sustain. Only in exceptional cases, usually parodies or highly experimental works, could a truly separate 'deviant' grammar be postulated, but such a grammar would serve only to generate the text that spawned it, and could not be used to generate a wider range of literary examples.

Russian formalism provides one pathway out of this dilemma, and several investigators have taken up the theme of foregrounding. Keller (1979: 283) notes that: the very act of switching from one language to another constitutes a radical moment of foregrounding . . . Codeswitching is so radical a phenomenon that in itself it constitutes foregrounding. Keller postulates that literary code switches are valid to the extent to which the foregrounding implied by the switch is justified in terms of other literary parameters. This simple characterization is potentially useful providing that it is tempered by some restraints. First of all, it must be borne in mind that it is the switch itself, the dynamic act of changing, that constitutes the foregrounding, and not necessarily the particular choice of languages involved, nor even the direction of the switch. Many investigators of bilingual Spanish-English literature have attempted to attach literary and sociological significance to the absolute language categories of lines of poetry, to the fact that a given word or line is in English or in Spanish. In the specific examples cited as evidence,<sup>5</sup> it may be possible to postulate a definite literary value for the choices involved, but this is not axiomatic. In examining specimens of spontaneous code switches, it is often difficult to discern a valid reason for the switch, and many speakers are unaware of having switched, even when the alternation has been pointed out to them. Jacobson (1978a: 21) in defining a set of criteria which permit code switching establishes the catchall category of 'preference,' which must be applied when no other overt reason is apparent. The literary value of code switching is largely one of relative weighting, the proportions of Spanish to English, in the entire text, and in the immediate context of a switch. It is the tension represented by the difference that is more significant, from a linguistic and a literary point of view, than the specific choice of languages, although clearly the latter choices are far from arbitrary. As long as foregrounding is taken to mean foregrounding of the switch, the act of transition, and an underscoring of membership in a bilingual community, then the above characterization is unobjectionable. One must be wary, however, of confusing language switching with choice of languages, since the two factors, while related, are nonetheless distinct.

Another caveat involves specifying the exact moment of the switch. It has been seen that it is difficult and sometimes impossible to precisely specify the points at which a code switch may occur, other than in general syntactic/semantic terms. Much work has been done to determine syntactic criteria for intrasentential switches, and the results indicate a considerable latitude of variation, with only the construction of a bilingual grammar ultimately holding out the possibility of a complete linguistic explanation. On the other hand, the critical studies cited above, despite the implicit or explicit New Critical

orientation, point the way toward a text-grammar or pragmatics of intrasentential literary code-switching, establishing criteria and values for the assignment of entire sentences or fragments to distinct languages.

The concept of foregrounding must also be supplemented by a distinction between intersentential code switching, where a complete logical proposition is being highlighted, and intrasentential switching, where the bilingual grammar itself rises to prominence. There are three categories to be defined, each of which presupposes a different series of linguistic, psychological and social connotations. Type I is the monolingual text, perhaps with a handful of L<sub>2</sub> words thrown in for flavor. Much Chicano and Boricua poetry written entirely in English or in Spanish falls into this category, and does not necessarily presuppose a high degree of bilingualism among readers or writers, although biculturalism is clearly assumed. Some writers (and readers) of such poetry speak little or no Spanish, and the Spanish words that crop up from time to time are frequently common to the lexicon of most Americans and serve as cultural/ethnic markers and not necessarily as part of a well-integrated bilingual grammar.

Type II bilingual literature exhibits intersentential code switches, where entire lines of poetry or entire sentences of prose are produced in a single language, with switches occurring at principal phrase or sentence boundaries. This type of writing, while possible from any bilingual individual, including one who has learned a second language late in life, is most typical of individuals who have learned each of the languages in a different cultural setting, and who associate contexts and consequently propositions with a specific language. It is impossible, in such texts, to establish the full extent of the writer's active bilingual competence, but the fact that entire propositions are expressed monolingually indicates that the writer sees fit to separate the two linguistic and cultural domains, and the works should be evaluated accordingly.

Type III bilingual literature exhibits intrasentential code switches, typical of individuals who have learned and/or used both languages approximately in similar or identical contexts. It is in such texts that the high degree of bilingual integration becomes most apparent, and the texts, while perhaps readable by a wide range of bilingual individuals, are truly representative of only a relatively small segment of the national population. Since intrasentential code shifts are most closely constrained by linguistic limitations, it is in this type of literature that the faithfulness to societal linguistic norms may be most easily measured; it is also here that one may most profitably apply linguistic criteria to determine the degree of integration in the writer's bilingual grammars.

In each type of text, the role of foregrounding is different. In type I, there is no real L<sub>2</sub> discourse, but merely insertion of individual items, for a variety of effects. The impression is that of a superposition of isolated elements. In type II, two complete sets of logical propositions are offered, and a switch of language automatically entails a shift of domain of discourse. In type III, a more or less balanced bilingual grammar is presupposed, with the extent of balance being reflected by the relative proportion of discourse in each language. While it may be possible to apply the same criteria as in type II to this type of language switching, the prime feature being foregrounded is the fact of bilingualism, the existence of a linguistic in-group where mere knowledge of two languages is not sufficient to share

the bilingual grammar, constructed of a finely integrated blend of the two languages. It is not coincidental that the type III texts, which have also attracted the attention of linguists and psychologists when found in spontaneous speech, provide the richest and most rewarding terrain for literary analysis.

#### LINGUISTIC CONSTRAINTS ON LITERARY CODE SWITCHING

In addition to providing the basis for the study of the literary effects of bilingualism, an examination of written texts will usually reflect the major constraints governing code switching, and as such may be used to corroborate observations made on spontaneous utterances. The data of most immediate interest to the linguist wishing to construct a theoretical competence-based model remain those representing the spoken language, which exhibit a greater diversity in terms of switching types and which presumably can be taken as evidence of the unreflecting output of whatever internal linguistic apparatus the speaker possesses to store what for the linguist are two distinct codes.

In several individual cases, the criteria of general acceptability to bilingual speakers have been used to evaluate literary specimens of code switching, but such a determination has never been made for a large corpus of literary examples. It is unlikely that such a large-scale inquiry could ever be usefully carried out, due to the highly personal and variable reactions to literary examples. As a result, the evaluation of the linguistic acceptability of literary examples will remain the province of specialized investigators dealing with isolated examples. At the same time, it is possible to consider a corpus of literary examples in terms of overall syntactic criteria, to determine to what extent the literary examples fit in with code switching patterns characteristics of spontaneous discourse. To this end, a survey was conducted on the following literary works characterized by intrasentential code shifting:

Author	ethnic group	title	genre
Rolando Hinojosa	Mexican-American	<i>Mi querido Rafa</i> (13 chaps.)	novel
Evangelina Vigil	Mexican-American	<i>Thirty an' Seen a Lot</i>	poetry
Tito Laviera	Puerto Rican-Am.	<i>La carreta made a U-turn</i>	poetry
Alurista	Mexican-American	<i>Timespace huracán</i>	poetry
Alurista	Mexican-American	<i>Spik in Glyph?</i>	poetry
Miguel Algarín	Puerto Rican-Am.	<i>On Call</i>	poetry

In addition, 25 poems by Mexican-American and Puerto Rican-American writers were selected from the issues of the *Revista Chicano-Ríquena*, with the only criterion for selection being the inclusion of a significant amount of intrasentential code shifting. In total, 1947 code shifts were tabulated, a corpus sufficiently large as to permit comparison with the examples collected in spontaneous discourse. The breakdown of categories is given in Table 6-1, where for purposes of comparison the percentages of shift types from spontaneous discourse are also reproduced. In these figures, examples of tag phrases have not been included in the literary examples, since they were culled out of the corpus of spontaneous utterances. In the literary examples, tags represented 2.4% of the total, a much lower percentage than in spon-

TABLE 6-1  
Frequency of shift-types (%)

Type	Literature	Spontaneous Discourse
bilingual repetition	1.23	6.23
L <sub>2</sub> noun insertion	12.36	9.16
and/or (before or after)	13.30	17.94
preposition	12.26	17.88
relative pronoun	1.23	3.44
sentence boundary	36.98	17.37
adverb	4.25	5.73
article + L <sub>2</sub> noun	0.94	5.92
L <sub>1</sub> adj. + L <sub>2</sub> noun	1.98	0.19
L <sub>1</sub> noun + L <sub>2</sub> adjective	1.70	0.51
subject/verb boundary	4.34	1.72
verb/object boundary	1.89	3.31
copula	1.60	1.15
porque/because	0.09	2.54
other subord. conj.	3.02	4.01
N in apposition	2.83	5.22

staneous discourse, so that a comparison including tags in both categories would reveal slightly different relative values. It may be seen from Table 6-1 that the literary values are significantly different from those representing spontaneous discourse. Only in Hinojosa's novel, which comes the closest to reproducing actual speech (as was the author's intent) are tag expressions found to a significant degree, making up nearly all of the 2.4% of the total corpus. Moreover, there are significant discrepancies between prose and poetry, with the former being closer to spontaneous speech patterns, as indicated in Table 6-2.

TABLE 6-2  
Frequency of literary shift-types (%)

Type	Poetry	Prose
bilingual repetition	2.29	0.00
L <sub>2</sub> noun insertion	16.37	7.72
and/or	8.27	19.11
preposition	14.26	9.96
relative pronoun	1.94	0.41
sentence boundary	28.17	47.15
adverb	5.81	2.44
article + L <sub>2</sub> noun	0.88	1.02
L <sub>1</sub> adj. + L <sub>2</sub> noun	3.52	0.20
L <sub>1</sub> noun + L <sub>2</sub> adjective	2.46	0.81
subject/verb boundary	5.81	2.64
verb/object boundary	2.82	0.81
copula	0.53	2.85
porque/because	0.18	0.00
other subord. conj.	2.11	4.07
N in apposition	4.58	0.81

More significant than the relative percentages of the switch types, however, is the reduced categorical instability of key pivotal items in the literary examples. For example, in the literary specimens co-ordinating conjunctions nearly exclusively occur in L<sub>2</sub>, as do prepositions, reflecting a much more stable category inclusion than in the case of spontaneous utterances. Only copulas and subordinating conjunctions exhibit a more characteristic categorical variability, while relative pronouns occur almost exclusively in L<sub>2</sub>. Also not found in this corpus of literary code switching is the insertion of L<sub>2</sub> conjunctions in the midst of an otherwise L<sub>1</sub> discourse. The frequency of code switches at sentence boundaries is quite high in literature, which reflects the preference for separating propositions by language categories, even in highly integrated bilingual literature. The percentage of inserted nouns is also high, in accordance with the foregrounding value customarily associated with such insertions. The proportion of shifts in noun + adjective combinations is also significantly higher in literature than in spontaneous speech, and nearly invariably reflects stylized combinations which violate both the uniform structure constraint and the requirements of syntactic congruence following the shift point. The combination of English adjective + Spanish noun is relatively common, as is the combination of Spanish noun + English adjective; these configurations rarely occur in normal code-switched conversational styles. Finally, it may be noted that the overall percentage of major structural violations, as measured by the uniform structure constraint and the requirements of syntactic congruence, is significantly higher in the literary examples (4.93%) than in spontaneous discourse (0.9%). This is not surprising, given the aesthetic values inherent in literature, but the overall panoramic presentation should cast a large measure of doubt on the feasibility of evaluating literary code switches by the same linguistic criteria utilized in studying normal spontaneous discourse.

## NOTES

1. Cf. the presentation in Sánchez (1983), who, however, does not deal explicitly with literature. Bruce-Novoa (1982) deals exclusively with Chicano literature, utilizing an implicit Marxist framework which manifests itself outwardly in a post-structuralist analysis.
2. Most recently, Bruce-Novoa (1982) has provided the most progressive analysis of Chicano poetry, utilizing a poststructuralist model which incorporates aspects of deconstructionist theory, particularly the notion of the critical maintenance of a fundamental dichotomy as the generative force of the text (in Bruce-Novoa's model, the opposition Life-Death). Although many of the texts studied by Bruce-Novoa contain examples of code switching, this type of linguistic behavior is not studied explicitly, but Bruce-Novoa notes at the outset (pp. 12-13) that bilingual language behavior among Mexican-Americans constitutes an 'interlingual and intercultural space.'
3. This includes, although not exhaustively, Keller (1976, 1979), Valdés-Fallis (1975, 1976a, 1977, 1978), Timm (1978), all of whom offer analyses based on or at least mentioning Spanish-English bilingual literature.
4. Cf. the discussion of this problem in Lipski (1976).
5. Most particularly, in the works of Keller and Valdés-Fallis, cited above.
6. Timm (1978) found that code switching examples in Tolstoy's *War and Peace* closely approximate those found in the spontaneous speech of Mexican-Americans, again indicating that code-switched narrative prose may at least partially establish congruence with spontaneous discourse.

## In Conclusion

The preceding chapters have described some of the constraints governing Spanish-English intrasentential switching and hinted at others. At the same time, it is all the more evident that much research must still be done before the theoretical questions addressed in Chapter One can be definitely answered, and before the goal of a realistic grammar of bilingual code switching can be reached. The data and analyses which have been presented in the foregoing chapters suggest the following tentative conclusions:

1. *One grammar or two?* The nature of linguistic constraints on bilingual language switching indicates the presence of two separate grammatical mechanisms, each of which generates acceptable monolingual utterances. A single undifferentiated grammar would of necessity have to include combinations which would be ungrammatical in one or the other of the languages and/or which would produce unacceptable code-switched configurations. Regardless of the fashion in which a bilingual speaker has learned Spanish and English, we may postulate the existence of a Spanish grammatical component and an English grammatical component.
2. *Language tagging.* As in the case of grammatical structures, bilingual language switching provides strong evidence in favor of the existence of separate lexicons, one for each language, with a tagging mechanism keeping the lexical domains separate when context dictates the need for strictly monolingual discourse. There is at the same time a considerable overlap between the two lexicons, largely in the form of borrowed nouns and noun phrases which have become nearly or totally lexicalized in a language. More English nouns have penetrated the Spanish lexicon than vice versa, but the number of lexical elements of Spanish origin which have entered English, at least in the case of bilingual Spanish-English communities in the United States is also considerable.
3. *The integrative mechanism.* Strict application of Occam's Razor requires that gratuitous meta-structures be avoided whenever possible, and that bilingual language behavior be described as much as

possible in terms of already existing monolingual grammars. As a result, preference must initially be given to modifications of existing grammars of Spanish and English, rather than to the formulation of a special bilingual generative mechanism, unless experimental evidence inexorably militates in favor of the latter alternative. Among the proposed integrative models which have been examined, the bilingual tagging mechanism of Sankoff and Poplack (1980, 1981) offers the greatest degree of promise, since it deals directly with bilingual surface structures and adds no special meta-system to control bilingual language shifting. This type of grammar must be supplemented by constraints which account for the observed syntactic congruence of major constituents surrounding the point of an intrasentential switch. It has also been suggested that the variable rules which form an integral part of such bilingual grammars ultimately derive their statistical coefficients not from observation of performance, but from theoretical considerations which reflect the motivation for shifts in determinable circumstances.

4. *Categorical variability.* Under certain conditions, bilingual language shifting is characterized by an inherent categorical instability with respect to the tagging of certain items in the bilingual phrase. These items normally represent grammatical transitions, and include conjunctions, prepositions and relative pronouns. In other cases, the instability results from the presence of a grammatical node or juncture, which embodies a bilingual shift on either side of the juncture. Experimental evidence reinforces the observations of performance, to the effect that under specifiable conditions, some sentence elements are best analyzed as having variable category membership, neither entirely English nor entirely Spanish. The theory of fuzzy subsets or categories has been suggested as a metatheoretical basis for the description of categorical instability, which would ultimately provide many if not all of the variable coefficients for a bilingual code-switching grammar. The embryonic state of research into the psycholinguistics of code shifting presents a black-box situation, in which one may observe the input and output of a system, but in which the internal mechanism remains sealed off from direct observation and manipulation. The variable category inclusion model is designed to reproduce these black-box conditions, by providing a formalism for depicting the quantitatively variable nature of bilingual language switching as well as bilingual recall of previously produced or perceived material. At present category inclusion values, like variable rules, come predominantly from the properties of observed discourse, but the groundwork is being laid for a more comprehensive multiplexing, which could incorporate components relating to situation, grammatical category, syntactic congruence, anticipational switching, bilingual repetitions, and so forth. In order for such a fully ramified model to come into existence, it will be necessary to gain a clearer understanding of the relations between an initial cognitive image and the logico-semantic representation which ultimately becomes realized in one or the other of the bilingual speaker's languages, or perhaps in both. The models discussed in this study in no way purport to explain language switching from a developmental standpoint, but rather describe language usage in a bilingual context. Semantic factors alone are not sufficient to account for allowable language switches; neither are syntactic congruences or psychological desires (although the latter could, if forceably enough applied, cause a switch under nearly any

conditions). What is needed is a synthetic approach which provides for an evaluation of the contribution of these various factors, modeling the manner in which they interact to provide a smooth and often unconscious transition from one language to the other. The proliferation of models and methods employed to study code shifting, even when remaining within the limited purview of Spanish-English bilingualism, demonstrates that the ultimate goal of a unified linguistic description of bilingualism has not yet been attained. However, with respect to the psycholinguistics of bilingual behavior as reflected by language switching, significant progress is being made, and several distinct research programs are proceeding in parallel fashion and may eventually be unified. A key feature in the contemporary analysis of language switching is the abandonment of doctrinaire insistence on extremist models, which postulate either totally separate grammars or a single homogeneous underlying grammar. Researchers are coming to accept that all bilingual speakers, regardless of their ethnic background, the manner in which they learned their languages or the community in which they live, exhibit characteristics of both separate grammars and of a single unified underlying system, and that it is fruitless to force a choice between what are in reality two aspects of a single phenomenon. Once this hurdle has been cleared, it is possible to accept language-switching behavior as a manifestation of the degree of integration in the bilingual modality and to utilize the data derived therefrom to model the linguistic competence of the bilingual speaker. This competence is both subject to formalizable linguistic constraints and governed by the infinitely variable psychological forces which interact and unite to produce the singular behavior that is intrasentential language switching. Finally, it is perhaps not out of order to remark that there is a certain poetic justice in the fact that the language behavior of groups for so long considered linguistically deficient is now proving to be the most fertile territory for the investigation of all levels of bilingualism; even more importantly, the interaction of English and Spanish is seen as a structured process, forging new norms and creating a definite place for itself among the linguistic communities of the world.

## BIBLIOGRAPHY

- Aguirre, Adalberto. 1978. An experimental socioinguistic analysis of Chicano bilingualism. San Francisco: R&E Associates.
- \_\_\_\_\_. 1981. Toward an index of acceptability for code alternation: an experimental analysis. *Aztlan* 11.297-322.
- Albert, M., and L. Obler. 1978. The bilingual brain. New York: Academic Press.
- Anisman, P. 1975. Some aspects of code-switching in New York Puerto Rican English. *Bilingual review* 2.56-85.
- Arbib, M., and E. Manes. 1975. A category-theoretic approach to systems in a fuzzy world. *Synthèse* 30.381-406.
- Barkin, F. 1976. Language switching in Chicano Spanish: linguistic norm awareness. LEKTOS: interdisciplinary working papers in language sciences, special issue, pp. 46-64.
- \_\_\_\_\_. 1978a. Language switching in Chicano Spanish: a multi-faceted phenomenon. In H. Key, G. McCullough, J. Sawyer, eds., SWALLOW VI. Long Beach: California State University, pp. 1-10.
- \_\_\_\_\_. 1978b. Loanshifts: an example of multilevel interference. In A. Lozano, ed., SWALLOW VII. Boulder: University of Colorado, pp. 1-10.
- Barkin, F., and A. Rivas. 1979. On the underlying structure of bilingual sentences. Paper read at the III Symposium on Spanish and Portuguese Bilingualism.
- Beardmore, H. 1974. Development of the compound-coordinate distinction in bilingualism. *Lingua* 33.123-127.
- Beckmann, P. 1972. The structure of language: a new approach. Boulder: Golem Press.
- Bruce-Novoa, John D. 1982. Chicano poetry: a response to chaos. Austin: University of Texas Press.
- Bull, W. 1950. Spanish adjective position: present rules and theories. *Hispania* 33.297-303.
- \_\_\_\_\_. 1954. Spanish adjective position: the theory of valence classes. *Hispania* 37.32-38.
- \_\_\_\_\_. 1965. Spanish for teachers. New York: Ronald Press.
- Carbonero, P. 1982. El habla de Sevilla. Sevilla: Ayuntamiento de Sevilla.
- Cárdenas, D. 1972. Compound and coordinate bilingualism/biculturalism in the southwest. In R. Ewton and J. Ornstein, eds.,

- Studies in language and linguistics 1972-1973. El Paso: Texas Western Press, pp. 165-180.
- Cedergren, H., and D. Sankoff. 1974. Variable rules: performance as a statistical reflection of competence. *Language* 50.333-355.
- Clyne, M. 1967. Transference and triggering. The Hague: Martinus Nijhoff.
- \_\_\_\_\_. 1972. Perspectives on language contact. Melbourne: Hawthorne Press.
- Contreras, H. 1980. Stress, word order and the notion of subject in Spanish. In L. Waugh and C. van Schooneveld, eds., The melody of language. Baltimore: University Park Press, pp. 45-53.
- Dearholt, D., and G. Valdés-Fallis. 1978. Toward a probabilistic automata model of some aspects of code-switching. *Language in society* 7.411-419.
- Diller, K. 1974. "Compound" and "coordinate" bilingualism: a conceptual artifact. *Word* 26.254-261.
- Di Pietro, R. 1978. Code-switching as a verbal strategy among bilinguals. In M. Paradis, ed., Aspects of bilingualism. Columbia, S.C.: Hornbeam Press, pp. 275-282.
- Ervin, S., and C. Osgood. 1954. Second language learning and bilingualism. *Journal of abnormal and social psychology*, supplement 49.139-146.
- Espinosa, A. 1917. Speech mixture in New Mexico: the influence of the English language on New Mexican Spanish. In H. Stephens and H. Bolton, eds., The Pacific Ocean in history. New York: Macmillan, pp. 408-428.
- Fantini, A. 1978. Bilingual behavior and social cues. In M. Paradis, ed., Aspects of bilingualism. Columbia: Hornbeam Press, pp. 283-301.
- Fishman, J., E. Herasimchuk, and R. Ma. 1971. Bilingualism in the barrio. Bloomington: Indiana University Press.
- Genesee, F., J. Hamers, W. Lambert, L. Mononen, M. Seitz, and R. Starck. 1978. Language processing in bilinguals. *Brain and language* 5.1-12.
- Gingras, R. 1974. Problems in the description of Spanish-English intrasentential code-switching. In G. Bills, ed., Southwest areal linguistics. San Diego: Institute for Cultural Pluralism, pp. 167-174.
- Gumperz, J., and E. Hernández-Chávez. 1970. Cognitive aspects of bilingual communication. In W. Whitley, ed., Language use and social change. London: Oxford University Press, pp. 111-125.

- Hadlich, R. 1971. A transformational grammar of Spanish. Englewood Cliffs: Prentice-Hall.
- Haugen, E. 1973. Bilingualism, language contact and immigrant languages. In T. Sebeok, ed., Current trends in linguistics, Vol. 10, Part 1. The Hague: Mouton, pp. 505-591.
- Heras, I., and K. Nelson. 1972. Retention of semantic, syntactic and language information by young bilingual children. *Psychonomic science* 29.391-392.
- Huerta, A. 1980. Code-switching among Spanish-English bilinguals. In R. Nash, and D. Belaval, eds., Readings in Spanish-English contrastive linguistics. San Juan: Inter-American University Press, pp. 206-234.
- Huerta-Macías, A. 1981. Code-switching in Chicano six-year-olds. In R. Durán, ed., Latino language and communicative behavior. Norwood: ABLEX, pp. 153-168.
- Jacobson, R. 1977a. How to trigger code-switching in a bilingual classroom. In B. Hoffer, and B. Dubois, eds., Southwest areal linguistics then and now. San Antonio: Trinity University, pp. 16-39.
- . 1977b. The social implications of intra-sentential code-switching. In R. Romo, and R. Paredes, eds., New directions in chicano scholarship. Special issue of *The new scholar*, pp. 227-256.
- . 1978a. Code-switching in South Texas: sociolinguistic considerations and pedagogical applications. *Journal of the linguistic association of the Southwest* 3.20-32.
- . 1978b. Anticipatory embedding and imaginary content: two newly identified codeswitching variables. In A. Lozano, ed., SWALLOW VII. Boulder: University of Colorado, pp. 16-25.
- Kaufman, A. 1975. Introduction to the theory of fuzzy subsets, translated by Swanson. New York: Academic Press.
- Keller, G. 1974. Ambiguous verbal stimulus test to measure language dominance in Spanish-English bilinguals. Published by Bilingual review, Jamaica, New York.
- . 1976. Towards a stylistic analysis of bilingual texts: from Ernest Hemingway to contemporary Boricua and Chicano literature. In Beck, Davis, Hernández, Keller, and Tarán, eds., The analysis of hispanic texts: current trends in methodology. Jamaica, N.Y.: Bilingual Press, pp. 130-149.
- . 1979. The literary strategems available to the bilingual Chicano writer. In F. Jiménez, ed., The identification and analysis of Chicano literature. Ypsilanti: Bilingual Press, pp. 262-316.

- \* Kimple, J., R. Cooper, and J. Fishman. 1969. Language switching and the interpretation of conversations. *Lingua* 23.127-134.
- Kirstein, B., and A. de Vincenz. 1974. A note on bilingualism and generative grammar. *IRAL* 12.59-61.
- Kolers, P. 1966a. Interlingual facilitation of short-term memory. *Journal of verbal learning and verbal behavior* 5.314-319.
- . 1966b. Reading and talking bilingually. *American journal of psychology* 79.357-376.
- . 1968. Bilingualism and information processing. *Scientific american* 218.78-86.
- Lakoff, G. 1975. Hedges: a study in meaning criteria and the logic of fuzzy concepts. *Journal of philosophical logic* 2.458-508.
- Lambert, W. 1969. Psychological studies of the interdependencies of the bilingual's two languages. In J. Puhvel, ed., Substance and structure of language. Berkeley: University of California, pp. 99-125.
- Lance, D. 1975. Spanish-English code-switching. In E. Hernández-Chávez, A. Cohen, and A. Beltramo, eds., *El lenguaje de los chicanos*. Arlington: Center for Applied Linguistics, pp. 138-153.
- Lenneberg, E. 1967. Biological foundations of language. New York: Wiley.
- Lindholm, K., and A. Padilla. 1978. Language mixing in bilingual children. *Journal of child language* 5.327-335.
- Lipski, J. 1975. A semantic-syntactic shift in Spanish. *Folia linguistica* 7.149-163.
- . 1976. Poetic deviance and generative grammars. *PTL* 2.241-256.
- . 1977a. Proposed subjects in questions: some considerations. *Hispania* 60.61-67.
- . 1977b. Code-switching and the problem of bilingual competence. In M. Paradis, ed., *Fourth LACUS Forum* (Columbia, S.C.: Hornbeam Press, pp. 263-277).
- . 1978. Code-switching and bilingual competence. In M. Paradis, *Aspects of bilingualism*. Columbia, S.C.: Hornbeam Press, pp. 250-264.
- . 1979. Bilingual competence and code-switching. *Langue et l'Homme* 42.30-39.
- . 1980. Surface deviation as grammatical competition. In J. Copeland, and P. Davis, eds., *Seventh LACUS Forum*. Columbia, S.C.: Hornbeam Press, pp. 73-83.

- . 1982. Spanish-English language switching in speech and literature: theories and models. *Bilingual review* 9.191-212.
- ~~Macnamara, J., and S. Kushnir.~~ 1971. Linguistic independence of bilinguals: the input switch. *Journal of verbal learning and verbal behavior* 10.480-487.
- ~~McClure, E.~~ 1981. Formal and functional aspects of the code-switched discourse of bilingual children. In R. Durán, ed., *Latino language and communicative behavior*. Norwood, N.J.: ABLEX, pp. 69-94.
- ~~McClure, E., and J. Wentz.~~ 1975. Functions of code-switching among Mexican-American children. In R. Grossman, L. San, and T. Vance, eds., *Papers from the Parasession on Functionalism*. Chicago Linguistic Society, pp. 421-432.
- ~~McMenamin, J.~~ 1973. Rapid code-switching among Chicano bilinguals. *Orbia* 22.474-487.
- ~~Peñalosa, F.~~ Chicano sociolinguistics. Rowley, Mass.: Newbury House.
- ~~Penfield, W., and E. Roberts.~~ 1959. *Speech and brain mechanisms*. Princeton: Princeton University Press.
- ~~Pfaff, C.~~ 1979. Constraints on language mixing: intrasentential code-switching and borrowing in Spanish/English. *Language* 55.291-318.
- ~~Poplack, S.~~ 1980. Sometimes I'll start a sentence in English y termino en español. *Linguistics* 18.581-618.
- . 1981. Syntactic structure and social function of code switching. In R. Durán, ed., *Latino language and communicative behavior*. Norwood, N.J.: ABLEX, pp. 169-184.
- ~~Redlinger, W.~~ 1976. A description of transference and code-switching in Mexican-American English and Spanish. In G. Keller, R. Teschner, and S. Viera, eds., *Bilingualism in the bicentennial and beyond*. Jamaica, N.Y.: Bilingual Press, pp. 41-52.
- ~~Reyes, R.~~ 1976. Language mixing in Chicano bilingual speech. In D. Bowen, and J. Ornstein, eds., *Studies in southwest spanish*. Rowley, Mass.: Newbury House, pp. 183-188.
- ~~Riegel, K.~~ 1968. Some theoretical considerations of bilingual development. *Psychological bulletin* 70.647-670.
- ~~Rivas, A.~~ 1981. On the application of transformations to bilingual sentences. Paper presented to IV Symposium on Spanish-Portuguese Bilingualism. University of Massachusetts working papers on Hispanic linguistics and bilingualism.
- ~~Sánchez, R.~~ 1972. Nuestra circunstancia lingüística. *El grito* 6.45-74.
- . 1978. Denotations and connotations in Chicano code-switching. In A. Lozano, ed., *SWALLOW VII*. Boulder: University of Colorado, pp. 187-198.
- . 1983. Chicano discourse. Rowley, Mass.: Newbury House.
- ~~Sankoff, D., and S. Poplack.~~ 1980. A formal grammar for code-switching. Centro de Estudios Puertorriqueños, CUNY, working paper no. 8.
- . 1981. A formal grammar for code-switching. *Papers in linguistics* 14.3-46.
- ~~Scotton, C.~~ 1979. Codeswitching as a "safe choice" in choosing a lingua franca. In W. McCormack, and W. Wurm, eds., *Language and society: anthropological issues*. The Hague: Mouton, pp. 71-87.
- ~~Sridhar, S.~~ 1989. The syntax and psychology of bilingual code mixing. *Canadian journal of psychology* 34.407-416.
- ~~Timm, L.~~ 1975. Spanish-English code switching: el porqué y how not to. *Romance philology* 28.473-482.
- . 1978. Code-switching in War and Peace. In M. Paradis, ed., *Aspects of bilingualism*. Columbia, S.C.: Hornbeam Press, pp. 302-315.
- ~~Tulving, E., and V. Coirota.~~ 1970. Free recall of trilingual lists. *Cognitive psychology* 1.86-98.
- ~~Valdés-Fallis, G.~~ 1975. Code-switching in bilingual Chicano poetry. In G. Canton Harvey, and M. Heiser, eds., *Southwest languages and linguistics in educational perspective*. San Diego: Institute for Cultural Pluralism, pp. 143-170.
- . 1978a. Code-switching in bilingual Chicano poetry. *Hispania* 60.877-885.
- . 1978b. Social interaction and code-switching patterns: a case study in Spanish/English alternatives. In G. Keller, R. Teschner, and S. Viera, eds., *Bilingualism in the bicentennial and beyond*. Jamaica, N.Y.: Bilingual Press, pp. 53-85.
- . 1978a. Code-switching and language dominance: some initial findings. *General linguistics* 18.90-104.
- . 1978b. Code-switching and the classroom teacher. Arlington: Center for Applied Linguistics, document language in education: theory and practice, number 4.
- . 1979. Is code-switching interference, integration or neither? In E. Blansett, and R. Teschner, eds., *A festschrift for Jacob Ornstein*. Rowley, Mass.: Newbury House, pp. 314-325.

- \_\_\_\_\_. 1981. Code-switching as deliberate verbal strategy: a micro-analysis of direct and indirect requests among bilingual Chicano speakers. In R. Durán, ed., Latino language and communicative behavior. Norwood, N.J.: ABLEX, pp. 95-107.
- ~~W~~akefield, J., P. Bradley, B-H Lee Yom, and E. Doughtie. 1975. Language switching and constituent structure. *Language and speech* 18.14-19.
- Wallis, E., and W. Bull. 1950. Spanish adjective position: phonetic stress and emphasis. *Hispania* 33.221-229.
- Wentz, J. 1977. Some considerations on the development of a syntactic description of code-switching. Ph.D. dissertation, University of Illinois.
- Zadeh, L. 1975. Fuzzy logic and approximate reasoning. *Synthèse* 30. 407-428.
- ~~W~~entella, A. 1981. Tá bien: you could answer me en cualquier idioma: Puerto Rican code-switching in bilingual classrooms. In R. Durán, ed., Latino language and communicative behavior. Norwood, N.J.: ABLEX, pp. 109-131.
- Zipf, G. 1965. The psycho-biology of language. Cambridge: M.I.T. Press. First edition 1935.