2. LITERATURE REVIEW

Several topics should be addressed to preface the study undertaken in this dissertation. In this chapter, I review published literature from six sources. First, because code switching by definition involves bilingual speech, a review of the literature which attempts to define bilingualism will be sketched. Following this, I review some recent and influential treatments of code switching, narrowing in on grammatical aspects of code switching, the topic of this dissertation. I then touch upon important distinctions in language contact phenomena (borrowings, calques, pidgins and creoles) to help properly discriminate code switching from other contact phenomena. A sketch of recent developments in syntax is then provided, focusing on the Minimalist Program which will form the theoretical basis for the discussion in chapter 5. Finally, I briefly discuss the linguistic literature available on Spanish and Nahuatl, the languages which comprise the corpus reported in chapter 4, and then sketch the historical and social context of Spanish-Nahuatl bilinguals in contemporary Mexico.

2.1 What is Bilingual Proficiency?

I will be concerned in this dissertation with what has been called “individual bilingualism,” in distinction from “societal bilingualism” (Baetens Beardsmore, 1986). The latter is concerned with language policy and the social rewards of language choice

(Ferguson, 1959; Kjolseth, 1978), while the former is concerned with the representation of language and language use in the individual who speaks more than one language.[33]

2.1.1 Some Definitions

Haugen's (1953: 7) definition of individual bilingualism is extremely inclusive and may therefore be a good starting point:

> Bilingualism is understood . . . to begin at the point where the speaker of one language can produce complete meaningful utterances in the other language.

There are some difficulties, it seems, with this very generous view of bilingualism. For instance, while traveling in Paris, I was able to produce basic questions in French to find my hotel (*Oú est le boulevard St. Germain?*) but would not have been able to count to ten, conjugate a single verb, or even use greetings appropriately. Similarly, if college freshmen may be described as "bilingual" after a first course in a foreign language, then it will be very important to distinguish this sort of bilingualism from others.

Halliday, McKintosh and Strevens (1970) have used "ambilingualism" to describe an individual who is capable of functioning equally well in either of his languages in every domain of activity and without any trace of one language in the other. According to Fishman *et al.* (1971), instances of bilingualism of this type are extremely rare:

[33]Bilingualism is generally taken to denote the use of more than one language, rather than strictly *two* languages (but see qualifications regarding relative competence below). I will use the term with this meaning throughout.

> Bilinguals who are equally fluent in both languages (as measured by their facility and correctness overall) are rarely equally fluent in both languages about *all possible topics*; this phenomenon is invariably a reflection of the fact that societal allocation of functions is normally imbalanced and in complementary distribution rather than redundant.

According to Hakuta (1986) and Romaine (1989), bilingualism is encouraged and maintained when different functions of language are assigned to different languages. This observation is consistent with Fishman's suggestion noted above; bilinguals may be better at talking about homelife in the language they use at home and better at talking about school life in the language used at school, precisely because language is the product of varied experience. Thus, a complementary distribution often emerges with respect to a bilingual's facility in different topics.

Cummins (1981) has also introduced a typology of bilingualism in connection with his Threshold Hypothesis, discussed at some length in section 1.3.2. A "proficient bilingual," in his terms, is simply one with native-like ability in both languages. A "partial bilingual" has native-like ability in one language and partial ability in another. Cummins' view that a third sort of bilingualism exists, called "limited bilingualism," in which an individual is said to have "low ability in both languages," is disputed in section 1.3.1 and in Valadez, MacSwan and Martínez (1997).

2.1.2 Critical Period Effects and Language Proficiency

Another factor influencing bilingual proficiency, often associated with partial bilingualism, is the critical period for language acquisition. A number of studies have reported extreme negative effects due to late exposure to primary linguistic data, often to the extent that universal principles of human language are inoperative.

The best known and most extensively studied case of this sort is Genie, a young girl who was tragically isolated in a small room from the age of 20 months to past 13 years. Curtiss (1977, 1989), who studied Genie's linguistic development following her rescue, found that Genie's speech never passed beyond "primitive syntactic and morphological ability," essentially "stringing together content words." Unlike normally developing children, Genie has never been able to use such functional elements as pronouns and prepositions. Examples of her utterances, excerpted from Curtiss (1989), include

(1a) Applesauce buy store
'Buy some applesauce at the store'

(1b) Genie full stomach
'I have a full stomach'

(1c) Want Curtiss play piano
'I want you to play the piano'

(1d) Father hit Genie cry long time ago
'When my father hit me, I cried, a long time ago'

Genie could understand such notions as predictions and conditionals, but she apparently expressed these only by a strategy of linear order. Despite her inability to acquire a grammar, Genie nonetheless demonstrated a "well-developed semantic ability" which included colors and numbers, shape and size terms, supraordinate, basic and subordinate class terms. In addition, she could distinguish among objects in visual and functional terms, could speak about nonpresent persons and objects, as well as past, future and imagined events.

In another study of critical period effects, Mayberry (1993) attempted to test experimentally her hypothesis that "the timing of language acquisition may have greater effects on the outcome of first-language acquisition than second-language acquisition . . ." (36). Mayberry gathered a subject population consisting of thirty-six signers who had used ASL (American Sign Language) continuously for at least 20 years, assigning each subject to one of four experimental groups according to acquisitional history. The subjects in three of these groups were all prelingually (congenitally) deaf but had acquired ASL as a first language in three different age ranges: 0-3 (native learners), 5-8 (childhood learners), and 9-13 (late first language learners). The fourth group of subjects was postlingually deafened and had learned ASL as a second language (late second language learners).

Mayberry's (1993) subjects were presented with thirty ASL sentences on videotape. They were asked to repeat back each sentence after hearing it ("elicited imitation"); eight of the thirty sentences (randomly selected) were then transcribed, coded, and analyzed along five different variables for each of the thirty-six subjects: (a) response length and sign production rate; (b) lexical and inflectional preservation and change; (c) preservation and sequencing of syntactic constituents; (d) grammatical acceptability; and (e) preservation of sentence meaning.

No significant difference among subjects was found with respect to either response length or sign articulation rate. However, Mayberry found that native signers showed a much higher proportion of preserved lexical stems than any of the other three groups (m=0.72 vs. m=0.56 in each of the other three groups). The native signers also

showed a higher proportion of preserved bound morphemes, a variable on which the late second language learners did better than both the childhood learners and the late first language learners (m=0.42 for the native signers, m=0.34 for the late second language learners, and m=0.26 for the childhood learners and late first language learners).

Mayberry found, too, that age of acquisition significantly affected recall of syntactic constituent structure. Interestingly, the responses of the natives and late second language learners mirrored the syntactic constituent structure of the stimuli to a much higher degree than did those of the childhood or late first language learners (m=0.71 and m=0.69 vs. m=0.53 and m=0.47, respectively). Age of acquisition also affected the extent to which subjects gave grammatically acceptable responses; this finding is due primarily to the late first language learners who gave significantly fewer grammatically acceptable responses than the other groups. Also, the native learners and the late second language learners outperformed the late first language learners and childhood learners in their ability to preserve the general meaning of the stimuli. But the native language learners in Mayberry's study outperformed all groups.

Thus, if a bilingual has acquired a strong command of a second language relatively late in life, there may be reason to suspect that some critical period effects will persist throughout his life. In a careful study of English-French bilinguals who had mastered French as young adults, Coppieters (1987) found that even the best second language speakers differed in important respects from a control group of native speakers on a highly specific test of grammaticality judgments.

Bley-Vroman (1989) has put together several characteristics of first and second language differences in support of his "fundamental difference hypothesis" which purports simply that there is a fundamental difference between a first language, acquired in infancy, and a second language. Bley-Vroman's paper is concerned with second language in early adulthood or beyond.

Bley-Vroman (1989) begins by alluding to "the poverty of the stimulus" in first language development, as classically formulated in Chomsky (1965), the notion that our knowledge of language is vastly underdetermined by our experience. In particular, Bley-Vroman is interested in the independence of first language acquisition from such matters as "intelligence, motivation and emotional state" (Chomsky, 1965: 58), none of which is characteristic of second language learning, according to Bley-Vroman. His argument, then, turns on carefully setting out specific ways in which "the logical problem of second language learning" differs from that of first language development.

First, Bley-Vroman points out that there is a lack of guaranteed success in second language learning. While L_1 learners inevitably achieve "perfect" mastery of a language, (adult) L_2 learners often do not. This is an important observation in the context of the question of the logical problem of second language acquisition, for it was precisely the observation that (all normal) first language learners are guaranteed success in the domain of language learning that led Chomsky and others to posit innate properties of language and learning in the first place.

Second, in addition to a lack of guaranteed success, second language learners also experience general failure. While L_1 learners achieve complete success in learning a

language (a complete grammar, in some sense of "complete"), L_2 learners rarely or never do so, according to Bley-Vroman. This observation was made in Coppieters' (1987) study as well.

Another dimension of difference consists in variation in success, course, and strategy of L_2 learning. While L_1 learners appear to be fairly uniform in degree of success, L_2 learners vary with regard to the degree of attained success in the second language. Also, L_1 acquisition is (basically) uniform in its course of development, but L_2 learners appear to follow very different paths with respect to one another (as shown in Meisel, Clahsen, and Pienemann (1981)). L_2 learners also differ from L_1 learners in terms of learning strategies (mnemonic tricks in vocabulary learning, "avoiding," "guessing," and so on).

Bley-Vroman also believes that second language learners differ from first language learners in terms of variation in goals. Thus, while L_1 learners display generally equal facility with language across a range of topics and purposes, L_2 learners often do not. Some L_2 learners develop a sort of "pidginized" language with only rudimentary grammatical devices but which suffices for communicative purposes. Other L_2 learners may be concerned with acquiring language appropriate to particular speaking situations (waiting on tables, lecturing, so on), while yet others may desire perfect (native-like) grammatical competence.

Fossilization is another special characteristic of second language learning. Unlike L_1 learners, many L_2 learners reach a particular level of success beyond which they apparently cannot move. Such learners, after much practice and drilling, may show signs

of success, only to return later to old habits. These sorts of "learning spurts" were also characteristic of Genie's post-critical period language learning.[34]

Also uncharacteristic of first language learners, second language learners tend to have indeterminate intuitions about the target language they are acquiring. Thus, L_1 learners generally have clear acceptability judgments about sentences in their language (and even vary systematically with respect to the strength of judgments), but even very advanced L_2 learners may lack clear acceptability judgments on sentences of the L_2. This observation is also consistent with Coppieters' findings.

In addition, Bley-Vroman points out, instruction plays an important role in second language learning. It is well attested that L_1 learners do not need anything like direct instruction to acquire a native language; in contrast, many studies cited by Bley-Vroman indicate that instruction aids L_2 learning in crucial respects.

Bley-Vroman also points out that negative evidence is required of second language learners but not of first language learners; that is, L_1 learners achieve mastery of their target language in the complete absence of negative evidence (information regarding which sentences are ill-formed), while L_2 learners are dependent upon information of this type to a large degree (although, according to Bley-Vroman, empirical studies in this area are inconclusive).

Finally, an important role is played by affective factors for second language learners which seem to be far less important for first language learners. Success in L_1

[34]Susan Curtiss (personal communication).

does not depend on matters of personality, socialization, motivation, attitude or the like in any important respect, but it has become a standard belief among language instructors that affective factors like these play a very important role for L_2 learners.

In light of these several characteristics which distinguish second language learning from first language development, Bley-Vroman concludes that "the same language acquisition system which guides children is not available to adults" (49). Hence, "the logical problem of foreign language acquisition then becomes that of explaining the quite high level of competence that is clearly possible in some cases, while permitting the wide range of variation that is also observed" (49-50). For other discussions of critical period effects in L_2, see Johnson and Newport (1989) and Strozer (1991).

Working within a particular tradition in second language research, Bley-Vroman's basic idea is that L_2 is governed by cognitive strategies outside of the language faculty. Thus, a second language learner might obey just those principles of UG that have been instantiated in the L_1, suggesting that the L_2 is parasitic in some way upon the L_1. If Bley-Vroman's basic idea is on the right track, as I will assume here, then it makes sense to speak of "low" or "high" bilingual ability, since one of the languages in question may be highly dependent upon mechanisms external to the language faculty.

Despite the work of Bley-Vroman and others, however, it is important to stress that for many researchers the question of a critical period for language acquisition is still an open one. For a consideration of arguments and evidence which attempt to explain "interference" in L_2 without appealing to a critical period, see papers in Gass and

Schachter (1989) and the excellent discussion in Epstein, Flynn and Martohardjono (1996).

2.1.3 Identifying Proficient Bilinguals

One method of determining bilingual proficiency could involve the use of norm-referenced instruments such as those used in Coppieters' (1987) study.[35] However, another adequate indicator may be a simple case history which considers such factors as (a) language dominance and loss, (b) age of onset of exposure to L_1 and L_2, (b) continued, sustained exposure to both languages, (c) functional specificity for each language, and (d) general verbal fluency.

As mentioned in section 2.1.1, various typologies of bilingualism have been proposed in the literature, ranging from the *ideal bilingual* (or *ambilingual*) to the *marginal* bilingual; depending upon the particular interests of the researcher, these categories are naturally altered and re-focused (for a review, see Skutnabb-Kangas, 1981; Baetens Beardsmore, 1986; Hakuta, 1986; Romaine, 1989). For my purposes here, I will refer to the *native bilingual* or the *proficient bilingual* as one who is relatively evenly dominant in both languages, has actively used both languages since infancy, has had continued, sustained exposure to both languages, and appears to have generally high verbal fluency. These are sufficent, not necessary conditions of "proficient bilingualism"; in other words, while I will take these factors to characterize native or

[35]However, see Valdés and Figueroa (1994) for an excellent discussion of general problems in the use of testing for bilingual assessment.

proficient bilinguals, I recognize that an individual may be highly bilingual (even just as proficient as a native bilingual) without some or perhaps even all of these conditions being met. In addition, I will refer to an *adult second language bilingual* as one who has mastered one of her languages after the critical period (essentially, after puberty).

Naturally, it is desirable that subjects used in research on code switching have a high degree of bilingual ability (ideally, that they be fluent bilinguals), and that they be selected from communities in which code switching is a socially acceptable speech style. I will return to this discussion in the context of my research design in chapter 3.

2.2 Code Switching

Code switching is a speech style in which fluent bilinguals move in and out of two (or conceivably more) languages, as illustrated in the Spanish-English examples in (2) and (3), taken from Belazi, Rubin and Toribio (1994).

(2) This morning *mi hermano y yo fuimos a comprar* some milk[36]
This morning *my brother and I went to buy* some milk

(3) The student brought the homework *para la profesora*
The student brought the homework *for the teacher*

[36]As is conventional in the literature, I will signal code switching boundaries by a change from regular to italicized text or vise versa. However, researchers on code switching sometimes use the notions of "embedded" versus "matrix" language and italicize the embedded strings. Since these notions do not play an important role here, and may at times be difficult to define (as when the structure of an example is not sufficiently understood to allow an uncontroversial structural description, or when an ungrammatical example may have no structural description), I will simply use italics/regular text as a toggle to signal switches between languages. Also, in this section, I will provide glosses and translations (the latter appearing in single quotes) unless the gloss happens to be identical with the translation, in which case the latter will not be presented.

Code switching has been studied from a number of perspectives. While some have focused on social dimensions of language use among code switchers, others have taken an interest in the grammatical properties of code switching. Although the latter aspect is central to my topic here, I will briefly sketch some of the pioneering work on the social dimensions of code switching before looking more closely at the literature on its grammatical aspects.

2.2.1 Social Aspects of Code Switching

Myers-Scotton (1993a) credits Blom and Gumperz (1972) with sparking interest in sociolinguistic dimensions of code switching. Blom and Gumperz (1972) studied code switching between dialects of Norwegian in Hemnesberget, a Norwegian fishing village. Although the topic was actually introduced in previous work (Gumperz and Hernandez-Chavez, 1970), the piece by Blom and Gumperz received considerably more exposure because it was included in Gumperz and Hymes' (1972) edited collection which became a standard textbook in the many new sociolinguistics courses created in universities in the 1970s.

In early work, Gumperz analyzed code switching as "situational" or "metaphorical," adding "conversational" code switching in Gumperz (1982). A situational switch involves a change in participants and/or strategies, whereas metaphorical code switching involves a change in topical emphasis (Gumperz and Hymes, 1972: 409). These switches appear to be motivated by speaker-external factors, according to Gumperz.

Gumperz (1982) proposes six major functions for conversational code switching: (a) quotation; (b) addressee specification; (c) interjection; (d) reiteration; (e) message qualification; and (f) personification vs. objectification. Valdés (1981) posits two additional ones: (g) mitigating the illocutionary effect of speech acts and (h) aggravating the illocutionary effects of speech acts. Hill and Hill (1986) found Gumperz' (1982) and Valdés' (1981) categories for code switching very useful in their analysis of the social aspects of Spanish-Nahuatl code switching in the communities on the western and south-western slopes of the Malinche Volcano in the Valley of Puebla-Tlaxcala. In particular, Hill and Hill (1986) observed high frequencies of code switching regarding topics of religion, money, counting and mathematics.

Gumperz' model focused on micro-level sociolinguistic analysis, using naturally occurring data from small-group interactions as the core data source. Whereas Labov and other sociolinguists tied language use to sociological variables, Gumperz considered language use to be a function of the dynamics of interaction. Thus, for Gumperz, language choice conveys intentional meaning of a sociopragmatic sort:

> Detailed observation of verbal strategies revealed that an individual's choice of speech style has symbolic value and interpretive consequences that cannot be explained simply by correlating the incidence of linguistic variants with independently determined social and contextual categories [Gumperz, 1982: vii].

Thus, Gumperz views language choice as a *discourse strategy*. In Gumperz (1970, 1976), grammatical dimensions of code switching are also considered.

In some respects, the influence of Gumperz's views may be seen in the development of Myers-Scotton's (1993a) Markedness Model, which posits that "speakers have a sense of markedness regarding available linguistic codes for any interaction, but

choose their codes based on the persona and/or relation with others which they wish to have in place" (Myers-Scotton, 1993a: 75). After displaying a transcription of a Swahili-dominated code-switched conversation in Swahili, English and Kikuyu at a parking lot in Nairobi, Myers-Scotton (1993a: 76-79) claims that Swahili is the unmarked choice for conversational interaction for public transactions of this sort.

Modeled after Grice's (1975) cooperative principle, Myers-Scotton (1993a: 113) posits a "negotiation principle" which underlies all code switches:

(4) *The negotiation principle*
Choose the *form* of your conversation contribution such that it indexes the set of rights and obligations [the RO set] which you wish to be in force between speaker and addressee for the current exchange.

The RO set is indexed by features. Speaking English in Nairobi, for instance, "may be indexical of any of a set of attributes, including most prominently 'plus high educational level/socio-economic status', 'plus authority', 'plus formality', and 'plus official'" (86).

In other words, people make assumptions about others based on how they speak, and the choices they make in conversation may signal particular social relationships which they seek to establish. These are interesting facts about language use and are certainly worth noting, but notions such as 'rights and obligations,' 'formality,' 'official' and so on are extremely difficult to factor into a theory of human interaction in terms of the sort of model which Myers-Scotton (1973a) seeks to develop. Essentially, her model translates an observation which is statable in very simple terms into one which uses an unconstrained, proliferating inventory of features which bear on social interaction. (Other interesting work on the social aspects of code switching appear in papers in Durán (1981) and Amastae and Elías-Olivares (1982).)

There is little hope, it seems, that a set of explanatory principles can reliably be developed which predict when and how people will say what, as Chomsky (1959) argued nearly forty years ago. Gumperz also agreed that, while social motivations for code switching can be catalogued and described in interesting and informative ways, theoretical constructs of general or universal use may be quite beyond our reach. This perspective is echoed in Lipski (1978: 261):

> Indeed, since the role of individual idiosyncratic factors seems to be an important aspect of code-switching, in that among groups of approximately equal bilingual abilities, some code-switch more than others, a complete determination of the sufficient conditions for code-switching probably lies beyond the reach of the behavioral sciences. With regard to the linguistic constraints, however, the path toward an eventual model seems more clearly indicated.

2.2.2 Grammatical Aspects of Code Switching

Although Labov (1971) characterized code switching as "the irregular mixture of two distinct systems," more recent work on the topic has shown that the mixture is in fact quite regular. Consider, for instance, the examples shown in (5) and (6).

(5) *I saw *lo*
'I saw it'

(6) *Los estudiantes habían *seen the Italian movie*
'The students have seen the Italian movie'

The fact that (5) and (6) are ill-formed and (2) and (3) well-formed suggests that code switching exhibits grammatical structure.

The remaining question is what the underlying structure is. As the references in Table 1 (page 68) indicate, much has been written on this topic. The earliest proposals regarding the grammatical properties of code switching began to appear in the 1970s with

Gumperz (1970, 1976), Timm (1975), Wentz (1977) and Pfaff (1979). In a study of Spanish-English code switching, Timm (1975) noticed that a code switch may not occur between a subject pronoun and a verb or between a verb and its object pronoun. Pfaff (1979) noticed additional constraints on code switching involving adjectives and nouns. These early studies were concerned with the basic facts of code switching and did not attempt to provide anything approaching an explanation of grammatical phenomena in code switching.

Below I will outline some popular and influential approaches which have attempted to explain code switching behavior; in particular, I will focus on Poplack (1980, 1981); Joshi (1985); Di Sciullo, Muysken and Singh (1986); Mahootian (1993); and Belazi, Rubin and Toribio (1994). I will also briefly discuss "speech-processing" approaches before summarizing the basic findings in the code switching literature.

2.2.2.1 Poplack's (1980, 1981) approach

Poplack (1980, 1981) and Sankoff and Poplack (1981) propose constraints which govern the interaction of the language systems, deemed a "third grammar" approach by Mahootian (1993). Specifically, Poplack proposes the Equivalence Constraint and the Free Morpheme Constraint, defined in (7) and (8).

(7) *The Equivalence Constraint*
Codes will tend to be switched at points where the surface structures of the languages map onto each other.

(8) *The Free Morpheme Constraint*
A switch may occur at any point in the discourse at which it is possible to make a surface constituent cut and still retain a free morpheme.

The idea in (7), given Poplack's examples, is that code switches are allowed within constituents so long as the word order requirements of both languages are met at S-structure; (8), stated differently, tells us that a code switch may not occur at the boundary of a bound morpheme. To illustrate, (7) correctly predicts that the switch in (9) is disallowed, and (8) correctly disallows (10).

(9) *told *le,* le *told,* him *dije,* dije *him* [Poplack, 1981: 176]
told *to-him,* to-him *I-told,* him *I-told,* I-told *him*
'(I) told him'

(10) **estoy* eat-*iendo* [Poplack, 1980: 586]
I-am eat-ing

Like the descriptive accounts, Poplack's constraints do not attempt to explain the facts described in (7) and (8), if indeed they are facts. In addition, because (7) and (8) are taken to be principles of the grammar, this approach suggests that code switching is governed by a sort of "third grammar" which constrains the interaction of the two systems in mixture, a topic I will return to in section 5.1. In addition to these theory-internal difficulties, (7) and (8) do not hold up to empirical tests. For instance, although the constructions in (11) are not disallowed by either of Poplack's constraints, informants for Belazi, Rubin and Toribio (1994: 225) regarded them as unacceptable.

(11a) *The students had *visto la película italiana*
The students had *seen the Italian movie*

(11b) *Los estudiantes habían *seen the Italian movie*
The students had *seen the Italian movie*

Other counter-examples will be presented in section 5.2.1 using data obtained during fieldwork for this dissertation. However, in section 5.3.1.5, I will argue that (7) is an essentially correct generalization.

2.2.2.2 Joshi's (1985) approach

In Joshi's (1985) system, the language which a code-switched construction is judged to be "coming from" is defined as the *matrix language*, while the other language is the *embedded language*. A "control structure" permits shifting from a matrix language to an embedded language but not vice versa. Thus, switches are asymmetrical in this system. Joshi (1985) further proposes the Closed-Class Constraint which stipulates that a code switch is impermissible between a closed-class item and an open-class item, as in (12); however, this constraint applies only to switches *into* the embedded language.

(12) *Constraint on Closed-Class Items*
Closed-class items (e.g., determiners, quantifiers, prepositions, possessives, Aux, Tense, helping verbs) cannot be switched.

For example, in Joshi's (1985) data, a Marathi postposition cannot be switched for the English preposition in (13).

(13) *some chairs-*war* [Joshi, 1985]
some chairs-on
'on some chairs'

Mahootian and Santorini (1996) claim that a French-Tunisian Arabic switch such as the one given in (14) does not actually violate Joshi's (1985) constraint, as Belazi, Rubin and Toribio (1994) claim it does, because the switch into the embedded language does not occur before the preposition, a closed-class item.

(14) J'ai joué avec *il-ku:ra* [Belazi *et al.*, 1994]
I.have played with the-ball
'I have played with the ball'

An unappealing aspect of Joshi's system is the existence of special code switching rules. In addition, (12) fails on empirical grounds, as shown by the Farsi-English

example in (15) (Mahootian, 1993) and the Italian-French example in (16) (Di Sciullo, Muysken and Singh, 1986).

(15) Anyway, I figured *ke* if I worked hard enough, I'd finish in the summer
'Anyway, I figured that if I worked hard enough, I'd finish in the summer

(16) No, *parce que* hanno *donné des cours*
no, because have given of the lectures
'No, because they have given the lectures'

Other counter-examples will be presented in section 5.2.1 using data obtained during fieldwork for this dissertation.[37]

2.2.2.3 Di Sciullo, Muysken and Singh's (1986) approach

Di Sciullo, Muysken and Singh (1986) propose that there is an anti-government requirement on code switching boundaries. Their constraint is shown in (17).

(17) *Government Constraint*[38]
a. If L_q carrier has index q, then Y_q^{max}.
b. In a maximal projection Y^{max}, the L_q carrier is the lexical element that asymmetrically c-commands the other lexical elements or terminal phrase nodes nominated by Y^{max}.

The proposed constraint in (17) has the virtue that it refers to an independently motivated principle of grammar (government), while proposals considered in sections 2.2.2.1 and 2.2.2.2 do not. However, it falls short of the basic requirement of descriptive adequacy. Because government holds between a verb and its object and between a preposition and

[37]As pointed out by Susan Curtiss, the example in (16) also serves as counter-evidence to Poplack's Equivalence Constraint since a switch occurs between French *que* and Italian *hanno*, a juncture where surface structures differ (since Italian is a *pro*-drop language and French is not).

[38]Government and related syntactic relations will be discussed in more detail in section 2.4.2.

its object, (17) predicts that a verb or preposition must be in the language of its complement. This is shown to be incorrect by examples in (2) and (14), repeated here.

(2) This morning *mi hermano y yo fuimos a comprar* some milk
This morning *my brother and I went to buy* some milk

(14) J'ai joué avec *il-ku:ra*
I.have played with the-ball
'I have played with the ball'

Again, other counter-examples will be presented in section 5.2.1 using data obtained during fieldwork for this dissertation.

2.2.2.4 Mahootian's (1993) approach

Mahootian (1993) and Mahootian and Santorini (1995) propose an account which focuses on the complement relation in phrase structure. Thus, as Mahootian (1993) put it, the operative principle is (15).

(15) The language of a head determines the phrase structure position of its complements in code switching just as in monolingual contexts.

Mahootian (1993) used a corpus of Farsi-English code switching data which she collected in naturalistic observations. In Farsi, objects occur before the verb, contrasting with basic word order in English. Mahootian (1993) observed that in code switching contexts the language of the verb determines the placement of the object, as in (16).

(16) You'll buy *xune-ye jaedid*
you'll buy house-POSS new
'You'll buy a new house'

In Mahootian and Santorini (1996), this mechanism is applied to Nartey's (1982) AdãNme-English code switching data to derive the following:

(17) E wo *green dress* ko
he/she PAST tone wear green dress ART
'(S)he wore a green dress'

Since in AdãNme the determiner *ko* is postnominal, it follows the NP *green dress.*

In their commentary on Belazi, Rubin and Toribio's (1994) work, Mahootian and Santorini (1996) slightly modify (16) to focus on more general properties of syntactic heads:[39]

(17) Heads determine the syntactic properties of their complements in code switching and monolingual contexts alike.

While the new formulation continues "focusing on phrase structure position ..., the syntactic properties referred to in ([17]) include other grammatical features such as syntactic category and finiteness" (472).

This approach also has some problems. Mahootian (1993) uses a tree-adjoining grammar (TAG) formalism which she stresses is an implementation of general work in the GB tradition. However, note that (16) and (17) are predicted by (15) or (17) only if the *branching direction* of the complement is encoded in the head. TAG formalisms encode branching direction by positing the existence of "auxiliary trees," partial structures which represent a complement on the left or right of its head, as appropriate to the language under consideration. However, classical GB Theory has long argued against encoding branching direction (Stowell, 1981; Chomsky, 1981), and current work in this

[39]Pandit's (1990) formulation, cited by Mahootian and Santorini (1996), is similar: "Code switching must not violate the grammar of the head of the maximal projection within which it takes place" (43).

tradition posits a universal base in which all complements branch to the left (Chomsky, 1995a).

In addition, there are well-known counter-examples to the formulation in (17). In both English and Spanish, it is generally assumed that Neg(ation) selects a tensed verb to its right. Despite the adherence to (17), the code switches in (18) are strongly deviant.

(18a) *El no *wants to go*
he not want to go
'He doesn't want to go'

(18b) *I never *voy a terminar*
I never AUX PRT finish
'I'll never finish'

Joshi's (1985) example presented above in (13), repeated here, is also a counter-example to (17). Although the complement of the Marathi postposition *war* appears on the right of the phrase, as required, the construction is ill-formed.

(13) *some chairs-*war*
some chairs-on
'on some chairs'

There are a number of other counter-examples to Mahootian and Santorini's system, but such examples are rejected as spurious by these authors because they do not come from naturalistic corpora. The basic argument for doing so relies upon the assumption that code switching is a socially stigmatized behavior, so code switchers may be influenced by this stigma in rendering judgments on sentences (Mahootian, 1993). However, the basic premise here is incorrect. Code switching is not universally stigmatized; indeed, in many cultures it is regarded as a prestigious display of linguistic talent. Conversely, there are individual languages which are extremely stigmatized in

some places (modern Nahuatl in Central Mexico, for instance), but linguists have fruitfully studied such languages using traditional elicitation methods for many years.

In addition, the requirement that only naturally occurring data be used poses many serious problems for linguistic theory. Mahootian (1993: 2) states that the goal of a linguistic theory is to "account for the natural occurrences of the data for which it has been constructed." However, at least within the generative linguistic tradition in which Mahootian works, it is crucially *not* the goal of linguistic theory to account for the set of naturally occurring utterances. First, this is an impossible task, since the set of structures which a syntactic theory accounts for is infinite, due to the recursive properties of the generating function, and could therefore never be finitely enumerable for inspection under natural circumstances. Second, there is a large range of (native) speech errors which occur naturally but which no linguistic theory seeks to represent (slips of the tongue, coughs, false starts, lapse of attention, interruptions, and on and on). Also, constructing a linguistic theory with this arbitrary limitation makes the investigation of *relative* judgments on acceptability utterly impossible, and such relative judgments have proven very useful in the investigation of syntactic phenomena (consider, for example, treatments of strong and weak crossover effects in *wh*-movement). Finally, and perhaps most problematic, naturalistic data do not contain sentences marked as ill-formed at all, so many hypotheses are untestable; specifically, no claim regarding what cannot occur can be properly evaluated.

Other counter-examples to Mahootian's approach will be presented in section 5.2.1 using Spanish-Nahuatl data obtained during fieldwork for this dissertation.

However, having made some specific criticisms, I should point out that (17) captures some basic insights which I will pursue in section 5.1.

2.2.2.5 Belazi, Rubin and Toribio's (1994) approach

Belazi, Rubin and Toribio (1994) propose the Functional Head Constraint, arguing that it emerges from principles independently motivated in the grammar for other phenomena. According to these researchers, the descriptive facts are these:

(19) A code switch may not occur between a functional head and its complement.

To explain the facts in (19), Belazi, Rubin and Toribio (1994) appeal to "feature checking," independently motivated to be at work in numerous other phenomena. However, Belazi, Rubin and Toribio (1994) also add an additional item to the feature stack. According to them, an additional *language feature*, such as [+Spanish] or [+English], is checked together with other features. If the features do not agree (a Spanish functional head with an English complement, or vice versa), then the code switch is blocked. They formulate their constraint as in (20).

(20) *The Functional Head Constraint*
The language feature of the complement f-selected by a functional head, like all other relevant features, must match the corresponding feature of that functional head.

Since (20) applies to f-selected configurations only, switches between lexical heads and their complements are not constrained. (See Abney (1987) for a discussion of f-selection.)

Belazi, Rubin and Toribio (1994) also noticed that modification structures provide a special challenge, since these do not involve functional categories. Thus, they posit the

Word-Grammar Integrity Corollary to account for code switching phenomena in these structures.

(21) *Word-Grammar Integrity Corollary*
A word of language X, with grammar G_X, must obey grammar G_X.

Some of the descriptive facts mentioned in (20) are in dispute, as Mahootian and Santorini (1996) point out, a matter I will return to in section 2.2.2.7. For now I will simply discuss this proposal in terms of its own internal weaknesses, putting aside empirical evidence.

First, the operation of (20) requires a language feature such as [+Spanish] or [+Greek]. Since this proposed "language feature" is not independently motivated for any other linguistic phenomenon, it serves only to re-label the descriptive facts. In addition, linguists take particular grammars to be derivative in nature, not primitive constructs, since primitives are by definition part of universal grammar. A particular language is a set of parameter values over the range of variation permitted by universal grammar, so positing a label for a particular language as a primitive in syntactic theory leads to an ordering paradox.[40]

Also, note that features generally have a relatively small set of discrete values, such as [+finite] or [-past]. By contrast, there are infinitely many particular languages (Keenan and Stabler, 1994), and the dividing lines between them are often quite obscure. Thus, a language feature might be set to [-Greek] as easily as it is set to [+Greek], introducing extreme computational complexity. The feature [+Chinese] would

[40]On the derivative nature of language in linguistic theory, see Lightfoot (1982).

presumably include all the mutually unintelligible languages of China, and [+Norwegian] would exclude Swedish even though Swedish and Norwegian speakers generally understand each other.

As Bickerton (1982) and others have stressed, the rules of grammar are independent of sociolinguistic factors. Chomsky (1995a: 11, n6) has also noted this fact in another connection:

> Thus, what we call "English," "French," "Spanish," and so on, even under idealizations to idiolects in homogeneous speech communities, reflect the Norman Conquest, proximity to Germanic areas, a Basque substratum, and other factors that cannot seriously be regarded as properties of the language faculty.

However, the analysis is greatly improved if we regard [+English] to be a collection of formal features which define "English," as Jacqueline Toribio (personal communication) has suggested. On this view, names for particular languages act as variables for bundles of features which formally characterize them. The ordering paradox disappears, because language features like [+English] or [+Spanish] are no longer taken to be primitives in the theory of syntax.

This now gives the Functional Head Constraint (FHC) in (20) new empirical content. In particular, to evaluate the FHC, particular hypotheses about which features of English, being distinct from features of Spanish, result in a conflict. No such hypotheses are presented or evaluated in Belazi, Rubin and Toribio (1994). In addition, the idea that head-complement configurations are checking domains must be independently motivated. If current approaches are correct in assuming that only head-head and head-spec configurations are checking domains (Sportiche, 1995b; Chomsky, 1995a), then the FHC

could not be correct, even if "the language feature" were given the empirical content it now lacks.

In addition, that (21) should be viewed as a "corollary" to (20) is far from obvious. To be a corollary, it should be incidentally shown in proving (20), or it should be some natural consequence of (20). Since (20) crucially depends upon a language feature in a functional category and (21) involves no functional categories, it is difficult to see how it might count as a corollary.

Some counter-examples to Belazi, Rubin and Toribio's (1994) approach will be presented in section 5.2.1 using Spanish-Nahuatl data from chapter 4, and conflicts between their data and those of other studies will be taken up in section 5.3.

Again, having made some specific criticisms, I nonetheless see a crucial insight in (20) and (21); it is the idea that lexical properties alone, expressed in terms of feature matrices, determine the code switching facts. I will return to this idea in section 5.1.

2.2.2.6 Speech-planning approaches

Finally, I should briefly discuss a class of proposals made within a speech-planning framework, exemplified in work by Azuma (1991, 1993), de Bot (1992) and Myers-Scotton (1993b, 1995). These approaches rely upon work on sentence production by Fromkin (1971) and Garrett (1975) and frequently use Levelt's (1989) *Speaking* model.

According to Azuma (1993) and Myers-Scotton (1993b), the matrix language defines the surface structure positions for content words and functional elements. Thus,

we expect (22a) to be well-formed but not (22b) since in (22b) the determiner *the* is not in the surface position of the matrix language (Azuma, 1993).

(22a) Uchi wa *whole chicken* o kau noyo
we TOPIC *whole chicken* ACC. buy TAG
'We buy a whole chicken'

(22b) *Watashi ga katta *the* hon wa takai
I NOM. bought *the* book TOPIC expensive
'The book I bought is expensive'

As best I can tell, this approach is equivalent to the Equivalence Constraint in (7) and subject to some of the same criticisms. In particular, it is subject to the same counter-examples, such as those presented in (11), repeated here.

(11a) *The students had *visto la película italiana*
The students had *seen the Italian movie*

(11b) *Los estudiantes habían *seen the Italian movie*
The students had *seen the Italian movie*

Notice that (11) is ill-formed even though the matrix language, whether it is English or Spanish in this case, has correctly defined the positions of content words and functional categories. I will present other counter-examples from the Spanish-Nahuatl corpus in section 5.2.1.

On theory-internal considerations, it is not obvious that code switching has the same character as other processing phenomena, such as limitations on center embedding and lengths of sentences. From the examples in (11), it is clear that unacceptability in code switching cannot be accounted for in such terms. While speech processing models invariably assume a uniform mechanism across languages, the examples in (11) do not differ at the surface from monolingual sentences except with respect to the phonetic

shape of some of their constituents. Yet they are ill-formed, a surprising fact if the parser is responsible.

In addition, a much more precise way of talking about processing of such constructions, perhaps along lines explored in Stabler (1994), should be employed if code switching boundaries are to be successfully defined in such terms. Myers-Scotton (1993b: 19) describes the level at which code switching boundaries are defined as "an even more non-representational level in message construction," but the non-representational character of this system is not made precise. Thus, before a theory of code switching can be successfully worked out in terms of a parsing theory, a much more precise theory will be required. (On the limitations of non-representationalist theories, see Fodor (1981), Chomsky (1991), and Wexler (1991).) However, just as parsing considerations are sufficient to rule out *some* monolingual constructions, they may be sufficient to rule out *some* code-switched constructions. I will discuss this matter in connection with some specific examples in section 5.3.1.1.

2.2.2.7 Summary of basic findings in code switching corpora

There is some disagreement regarding code switching boundaries in the literature. Much of Mahootian and Santorini's (1996) criticisms of Belazi, Rubin and Toribio (1994), for instance, are concerned with the accuracy of the data captured in (19). In Table 1 (page 68), I summarize the descriptive facts of the code switching corpora reported in the literature, and indicate if and where these facts have been challenged, as

Table 1: Summary of Basic Findings in Code Switching Corpora

Item ref #	*Descriptive boundaries* (+ = code switch)	*Reported in ...*	*in disagreement with ...*
1a	*because* + CP	Gumperz (1976)	Poplack (1981) Sankoff and Poplack (1981) Mahootian (1993)
1b	conj + CP	Gumperz (1976)	Poplack (1977) McClure (1981)
2	*that* + IP	Belazi, Rubin and Toribio (1994)	Bentahila and Davies (1983) Mahootian (1993)
3a	*have* + VP	Belazi, Rubin and Toribio (1994)	Di Sciullo, Muysken and Singh (1986)
3b	*modal* + VP	Belazi, Rubin and Toribio (1994)	Di Sciullo, Muysken and Singh (1986)
3c	*to* + V	Timm (1975)	Lipski (1978) Poplack (1981) McClure (1981)
3d	Aux + V	Timm (1975)	Lipski (1978) Poplack (1981) McClure (1981) Mahootian (1993)
3e	Neg + V	Timm (1975)	undisputed
4a	Q + NP	Belazi, Rubin and Toribio (1994)	Bentahila and Davies (1992) Mahootian (1993)
4b	Demonstrative + NP	Belazi, Rubin and Toribio (1994)	Nishimura (1985) Bentahila and Davies (1992) Mahootian (1993)
4c	Article + NP	Belazi, Rubin and Toribio (1994)	Brown (1986) Bentahila and Davies (1992) Mahootian (1993)
4d	Complex D + NP	Wentz (1977)	Poplack (1981)
5a	N + Adj (Adj from Adj-N language, N from N-Adj language)	Gumperz (1976) Lipski (1978) Belazi, Rubin and Toribio (1994)	Bokamba (1989) Mahootian and Santorini (1996)
5b	Adj + N (Adj from N-Adj language, N from Adj-N language)	Belazi, Rubin and Toribio (1994)	Poplack (1981)
6a	Subject pronoun + V	Timm (1975) Gumperz (1976) Lipski (1978)	Poplack (1981) Mahootian (1993) Bentahila and Davies (1983)
6b	V + object pronoun	Timm (1975) Gumperz (1976) Lipski (1978)	Poplack (1981) Mahootian (1993)
6c	clitic + V *or* V + clitic	Timm (1975)	undisputed
6d	Gapping constructions with Aux second V switched (marginal)	Gumperz (1976)	Poplack (1981)
7	A switch involving a bound morpheme	Poplack (1981) Sankoff and Poplack (1981)	Nishimura (1985) Mahootian (1993) Myers-Scotton (1993b)

best I know.[41] As will be discussed in chapter 3, Table 1 provided considerable guidance for my own fieldwork.

Despite problems in the proposals I have reviewed in this section, considerable progress has been made in our understanding of the descriptive facts of code switching, and a number of excellent attempts have been made to explain the distribution of these facts. In my view, the general approaches pursued by Belazi, Rubin and Toribio (1994) and Mahootian (1993) are the most promising avenues for success. Although both proposals are overly concerned with the complement relation in phrase structure, there are insights in both regarding the important role of syntactic heads, the lexicalization of parametric variation, and (in the former) feature checking theory. I will argue, too, that Poplack's Free Morpheme Constraint is essentially descriptively correct, but it must be refined and shown to relate to more primitive mechanisms of grammar (some suggestions are made in section 5.3.1.7).

After a discussion of important language contact phenomena, intended to aid in the development of a proper understanding of code switching data, I will review relevant topics in the theory of syntax as a way of pointing to the particular theoretical direction that the present study will take.

2.3 Language Contact Phenomena

Code switching is most likely to occur in bilingual communities -- that is, in communities in which languages have come into contact with one another. It is

[41]I am grateful to Jacqueline Toribio and Shahrzad Mahootian for comments on a draft of Table 1.

important in research on code switching to carefully distinguish code switching, in which two distinct language systems interact, from language contact phenomena in which one language influences the lexicon and perhaps the grammar of another.

2.3.1 Borrowings and Calques

Borrowing often occurs wherever languages come into contact; however, the degree to which speakers are aware of the non-native character of borrowed words may differ with each borrowed item. For instance, a monolingual English speaker might use the term *pork* without the slightest awareness that it was borrowed from French during the Norman Conquest. On the other hand, a speaker might use the expression *tour de force* fully aware that the expression is of French origin. In this latter case, the English speaker may have some grasp of the internal structure of the phrase without having a fully productive knowledge of the internal structure of French DPs. An English speaker who encounters the French word *genre* may also have difficulty pronouncing the word with a palatal fricative in initial position, since English phonology does not permit this.

Thus, it may be useful to think of lexical borrowing as a matter of degree. For most speakers, the French word *porc* is fully incorporated into English (spelled *pork*). Its only native properties are its phonetic features, its syntactic category, and aspects of its original meaning. In contrast, *genre* has retained all of these properties in addition to its peculiar phonology (for some English speakers). The phrase *tour de force* may have even kept some of the original syntactic properties spelled out within its maximal projection. However, since the monolingual speaker imagined here does not have a fully

productive linguistic system for French, these aspects of borrowing should not be regarded as cases of code switching.

Indeed, most researchers on code switching have emphasized that borrowing must be carefully distinguished from code switching (Pfaff, 1979; Sankoff and Poplack, 1981; Belazi, Rubin and Toribio, 1994). However, Hill and Hill (1986: 345) note that code switching and borrowing are often difficult to separate:

> In practice, it is quite difficult for linguists to distinguish between cases of borrowing and cases of code-switching. Since code-switching occurs in situations where two languages are in contact, foreign material in the usage of bilinguals can be of both types. It is not possible to divine the nature of speaker self-consciousness about foreign materials, so linguists have generally tried to distinguish between borrowing and code-switching on strictly linguistic grounds.

Borrowed words are often assumed to be marked by *morphological nativization*. Spanish words borrowed into Nahuatl, for instance, are marked with the thematic suffixes *-oa* (transitives), *-(i)hui* (intransitives), and *-lia* (applicatives). Hill and Hill (1986: 158) provide examples:

(23) quin-costar-oa trabajo
them cost-TRNS work
'It costs them work'

(24) costar-ihui in nêca trabajo
cost-INTRNS the that work
'That work is costly'

Unlike verbs that are borrowed from Spanish, nouns are not morphologically marked in any similar way, as (23) and (24) show (*trabajo* in these examples has the same morphological shape as the Spanish word *trabajo* 'work'). However, the Nahuatl plural suffix *-tin* or *-meh* is often attached to Spanish nouns. Notice, for instance, that the

speaker uses both Nahuatl (*-tin*) and Spanish (*-s*) morphology in a single utterance in (25) to pluralize the Spanish(-origin) word *persona* ('person').

(25) Pues, cateh *persónahtin,* âquin cmatih tlahtôzqueh, *personas* cpiah ocachi edad. [Hill and Hill, 1986: 165]
'Well, there are people who know how to speak, people who are older'

In (25), both *personas* and *persónahtin* may be word-level code switches, or they may suggest that *persona* has been borrowed into Nahuatl. This case illustrates the difficulty of distinguishing between borrowing and word-level code switches in bilingual speech communities. While Spanish *trabajo* appears to be fully incorporated into Nahuatl in (23) and (24), *persona* in (25) is in one instance morphologically coded for Spanish and in another for Nahuatl.

It is also possible to borrow only pragmatic or morphosyntactic properties while using the phonetic material of the native language; this is the case of calques, also called "loan translations." These are special instances of borrowing in which the phonetic properties of words from one language are used in combination with pragmatic or morphological properties of words from another. For instance, (26a) is the Nahuatl expression for 'I am hungry,' used universally in the towns around the Malinche Volcano and attested in dictionaries of the classical language from the earliest period, according to Hill and Hill (1986). Nonetheless, Hill and Hill (1986: 140) report that one speaker insisted on (26b), a reflection of Spanish usage, as the "legitimate" and "correct" form.

(26a) nimayana
I am hungry

(26b) nicpia apiztli
I have hunger

While (26b) is well formed from the point of view of grammar (just as English "I have hunger" is well-formed), we might think of this loan translation as *pragmatically,* or maybe *stylistically* dispreferred.

Hakuta (1986: 55) reports a similar case of loan translation from Leopold's notes of his German-English bilingual child who uttered (27), an expression composed of English words (except for German *man* 'one'), but which uses German syntax in the subordinate clause.

(27) Which grade is *man* in when *man* nine years old is?

However, although (27) is characterized as a loan translation, it is not the clear case that (26b) should be. It may be, on some analysis of (27), that *man* introduces features into the derivation which can only be satisfied with the V-final construction observed here, in which case (27) would be an example of code switching and its peculiar syntax would be explained. Analyses in this vein will be presented for Spanish-Nahuatl code switching in section 5.2.2.

Finally, Suárez (1977) presents twenty common words used in nearly all modern dialects of Nahuatl which have been borrowed from Spanish. These are given below in (28) (where the glosses are my English renditions based on Suárez' explication). In addition to these, Nahuatl of southeast Puebla, where I collected data, also productively uses *kuando* 'when,' from Spanish *cuando* (Alva Hernández, 1996; see also section 4.2.12).

The forms listed in (26) have replaced their classical equivalents in nearly all modern dialects of Nahuatl. There are also numerous "content" words borrowed into

Nahuatl from Spanish (perhaps as many words have been borrowed from Nahuatl into Spanish) (Hill and Hill, 1986). The extent of these content-word borrowings varies considerably from dialect to dialect.

(28)

Spanish word	*Nahuatl spelling*	*English gloss*
antes	antes	before
cada	kada	each
como	komo	like, as
de	de	of, from
desde	desde	(ever) since
después	despues	after
entonces	entonses	so, then
hasta	asta	until
(lo) que	loke	that which, the one who
más	mas	more (comparative)
más que	maske	although
mientras	mientras	while
ni	ni	neither
o	o	or
para	para	in order to
pero	pero	but
por	por	because (of)
porque	porke	because
pues	pues	well (interjection)
si	si	if

While these facts serve as very interesting data for the study of the bilingual lexicon, in the present context they suggest a challenging project of data collection. Nevertheless, despite the great influence of Spanish on Nahuatl, the languages are still sufficiently different to distinguish genuine cases of code switching. I will return to particular ways in which I will control for these factors in chapter 3.

Shaffer (1978) insists that all cases of language mixture involving the use of an individual lexical item should be regarded as cases of borrowing in the study of code switching, if only as a methodological precaution. The judgment as to whether a

particular piece of data represents borrowing or code switching will sometimes be a very complex matter, involving consideration of many different factors. As mentioned in section 2.3.1, phonological and morphological incorporation, as well as a comparison with monolingual speech, will be some of the factors which aid in making this determination. In section 5.3.1.7, I will attempt to make sense of this idea in terms of the architecture of the bilingual language faculty. In chapters 4 and 5, I will refer to words that have been lexically, syntactically, morphologically and phonologically incorporated into a host language as "fully borrowed" items, distinguishing them from "partial borrowings" (like *tour de force* and *genre* in English) and calques.

2.3.2 Creoles and Pidgins

Under particular social conditions, creoles and pidgins emerge when languages come into contact with one another. A pidgin is a simplifed language mostly made up of the substantive categories (nouns, verbs, adjectives, adverbs) from two languages which come into contact among adults who do not know each other's languages. In many situations, these pidgins become first languages for children raised in environments where they are spoken, resulting in the formation of *creole* languages, significantly different from the pidgins to which their speakers were first exposed.

Pidgins are characterized by a severely reduced lexicon, a sharp reduction in grammatical complexity, including the dropping of inflectional morphology, and a minimal syntax which is linearly ordered (Holm, 1988). By contrast, creoles exhibit complex structure, and, to an astonishing degree, a vast range of similarities may be found among the world's creoles, geographically separated by thousands of miles.

Consider, for instance, the similarities between Guyanese Créole, a French-based creole, and Krio, an English-based creole, gleaned from Todd (1985: 24):

French	*Guyanese*	*Krio*	*English*
Mangez	Mãòe	Chüp	Eat
J'ai mangé	Mo mãòe	A chüp	I ate
Il/Elle a mangé	Li mãòe	I chüp	He/She ate
Je mange/Je suis en train de manger	Mo ka mãòe	A de chüp	I am eating
J'avais mangé	Mo te mãòe	A bin chüp	I ate/had eaten
Je mangeais	Mo te ka mãòe	A bin de chüp	I was eating
Je mangerai	Mo ke mãòe	A go chüp	I shall eat
Il/Elle est plus grand que vous	Li gros pas u	I big pas yu	He/She/It is bigger than you

These and many other similarities among the world's creoles provide strong evidence for the existence of Universal Grammar, the innate "bioprogram" of which Bickerton (1981) wrote, a view also reflected in Naro (1971), Kay and Sankoff (1974) and many others. Macedo (1986) suggests that the similarities between the world's creoles reflect the unmarked values of core grammar, set to their default values in the presence of unstable and highly limited pidgin data.

In a study of thirteen creoles of various influence (French-based, English-based and Iberian), Taylor (1971) found the common properties listed in (29).

(29) Taylor's (1971) list of common features of creoles he studied
   a. The third person plural pronoun serves as nominal pluralizer.
   b. The combination of the markers of past and future expresses the conditional.
   c. The word for 'give' also functions as the dative preposition 'to' or 'for.'
   d. Plural 'which thing/person/time/place?' are employed to express 'what?', 'who?', 'when?', 'where?'
   e. A prepositional phrase is employed to express the possessive absolute ('mine', 'ours', 'the man's', etc.)
   f. The demonstrative pronoun is postponed to its referent ('house this').
   e. The definite article is postponed to its referent ('house the').
   g. The pronominal determinant is postponed to its referent ('house my').
   h. '(My) body' serves to express '(my)self.'
   i. The iterative (habitual) function is merged with the completive, the progressive, and the future.
   j. *na* is a locative preposition ('at, by, from, in, on, to').
   k. *ma* is 'but.'

Hancock's (1971) Map of Creoles and Pidgins notes the existence of a "creolized Nahuatl-Spanish" used in Nicaragua in the sixteenth century, now extinct. However, from the discussion in Hancock's source (Elliott, 1884), it is far from clear that this Nahuatl-Spanish hybrid was either a creole or a pidgin. Unlike a pidgin, Elliott's Nicaraguan variety had a complex system of inflectional and derivational morphology, including diminutives and patronymics. Also, unlike a creole, which would generally have numerous properties which neither donor language possessed, it exhibits in its syntax "only a few phrases of construction that separate this dialect from the present Castilian." Eliott (1884: 66) concludes,

> In the majority of cases ... the Spanish construction has been adhered to; in truth, this has been carried out so faithfully in so many instances, and the sequence of tenses so scrupulously maintained, the strikingly idiomatic expressions of Spanish used so naturally and fluently, that one might almost be led to believe it the language of a Spaniard himself.

What marks this Nicaraguan dialect as unusual is its heavy incorporation of Nahuatl vocabulary. As Elliott suggests, this Nicaraguan dialect is a variety of Spanish which has been heavily influenced by the native vocabulary of the Aztecs, not a creole at all.

Code switching may be safely distinguished from pidgins and creoles in at least two ways. First, the linguistic system under analysis will exhibit properties in common with creoles if it is a creole, and properties in common with pidgins if it is a pidgin. Code switching does not exhibit the functional reduction of a pidgin, nor does it evince the well-studied properties of a creole. Rather, in code switching, a structure emerges which has properties of *both* language systems. As the data in chapter 4 show, sometimes code switching leads to the formation of structures which neither language, taken alone, would permit; however, these constructions do not have the properties of pidgins or creoles discussed in this section.

Secondly, unlike the communities in which code switching occurs, communities in which creoles are spoken are usually home to monolinguals. If some set of constructions taken to be code switching were used by monolinguals, then there would be a possibility that a creole or a heavily lexically influenced dialect had been mistaken for language mixture. However, 2.37% of the Mexican population is monolingual in Nahuatl (INEGI, 1993), and although there is heavy Spanish influence on its lexicon, the constructions produced as code switching between these languages have nothing in common with creole or pidgin languages, and do not form a subset of either monolingual Nahuatl or monolingual Spanish (as shown in chapter 4 of this dissertation; see also Hill and Hill (1986)).

Because creoles and pidgins typically result from particular social conditions, and because they have unique linguistic properties as well, it should be relatively easy to distinguish them from code switching. In particular, consultants used in my study should be able to speak a local variety of Spanish to monolingual speakers of Spanish and a local variety of Nahuatl to monolingual speakers of Nahuatl. If either of these local varieties had the social or linguistic character of a creole or a pidgin, there would be grounds for concern. However, in fieldwork in Southeast Puebla, no such characteristics were observed.

2.4 The Theory of Syntax

Below I present the theoretical framework for my study, first considering some general advantages of linguistic formalism. The minimalist program, discussed in 2.4.3, comprises the particular theoretical framework I will use in chapter 5.

2.4.1 Some Advantages of Formalism in the Study of Grammar

Efforts to construct formal theories of grammar have a long history in linguistics and are rooted in general principles of science developed during the Enlightenment. As Chomsky pointed out in the Preface to *Syntactic Structures* (1957: 5), the use of formalism in linguistics has a number of concrete advantages:

> The search for rigorous formulation in linguistics has a much more serious motivation than mere concern for logical niceties or the desire to purify well-established methods of linguistic analysis. Precisely constructed models for linguistic structure can play an important role, both negative and positive, in the process of discovery itself. By pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed.

In addition, theories which lack the precision possible through formalism may be too vague to properly evaluate. In *Discourse on Bodies of Water* (1612), for instance, Galileo refuted an anonymous Jesuit opponent in a discussion of the nature of sunspots by pointing out that his arguments were so vague as to be true in all conceivable situations (tautological), hence not falsifiable. Galileo turned to the formal language of mathematics in order to express his claims in explicit terms, thus establishing an account of physical phenomena in a precisely formulated way. In linguistics too, "obscure and intuition-bound notions can neither lead to absurd conclusions nor provide new and correct ones, and hence they fail to be useful in two important respects" (Chomsky, 1957: 5).

Of course, there are lots of important and interesting academic pursuits which may not be amenable to the language of science. Such matters as art, politics, the human will, and certain aspects of language use may be quite beyond the grasp of a scientific model. This observation is only demeaning if we assume that scientific modes of understanding are the only valuable modes, a view which has not been seriously entertained since the end of Logical Positivism (Morrow and Brown, 1994).

Thus, in the present study, the theoretical framework will consist of generative grammar, broadly construed to include any precisely formulated model of language structure. In particular, the project of Chomsky (1957) onward, culminating in Chomsky (1995b), will form the basis of analysis for data presented here.

2.4.2 Generative Grammar Before the Minimalist Program

Chomsky's (1955, 1956, 1957) early work argued that the syntax of human language could not be properly modeled by either a generative context free grammar (such as those which underlie the languages of elementary arithmetic and formal logic) or a context sensitive grammar (one which makes reference to constituents in strings). Hence, he proposed a hybrid generative-transformational grammar as a plausible model of human language. Because his model took advantage of recent developments in the study of the foundations of mathematics (recursive function theory), Chomsky was able to make sense of Humboldt's (1836) characterization of language as "the infinite use of finite means."

Later work on syntax focused almost entirely on the transformational component of the grammar, the component that was least constrained. However, Chomsky (1970) argued that work in syntactic theory should focus on phrase structure rather than transformations, a proposal which resulted in the formulation of X-bar Theory (elaborated in Jackendoff (1977)). Eventually linguists proposed that the transformational component of the grammar be essentially eliminated, reduced to a single movement rule (called Move-α).

Shortly after the elimination of transformations as such, Stowell (1981) noticed that phrase structure rules were redundant with subcategorization specifications in the lexicon which had been proposed in Chomsky (1965). Stowell (1981), Chomsky (1981) and others thus proposed that the phrase structure grammar also be eliminated, replaced by a rich set of interactive modules or subtheories which serve to constrain movement and phrase structure.

$\overline{X}$ Theory played an important role in this collection of subtheories, capturing aspects of the internal constituent structure of sentences. In $\overline{X}$ Theory, lexical and functional categories are abbreviated as N for noun, V for verb, A for adjective, P for preposition, Adv for adverb, I for inflection, and C for complementizer. X (or Y or Z . . .) is a variable over these categorial constants (such that X is either N, V, A, P, Adv, I, or C). In a tree diagram, a word, which specifies its category as part of its meaning, expands to an appropriate X; X expands to $\overline{X}$, the intermediate level, which in turn expands to XP, the phrasal level. Spec(ifier) may be filled by any XP (or maximal projection). In analogy to a genealogical tree, the $\overline{X}$ schema specifies a tree structure in the way expressed in (30).

(30) *The $\overline{X}$ Theory*
XP is the mother of Spec, the sister of $\overline{X}$
$\overline{X}$ is the mother of $\overline{X}$, sister of YP
$\overline{X}$ is the mother of X, sister of YP

The schema expressed in (30) was believed to be part of the learner's innate mental capacity in early GB Theory. The particular order of elements dominated by any node, a factor which falls out from the interaction of other modules of the grammar, is not

intended by the schema in (30). Rather, (30) simply defines the hierarchical arrangement of nodes in a tree (Stowell, 1981).

In addition, in the version of generative grammar which emerged in the 1980s, grammatical relations such as "government" and "binding" played very important roles, leading to the new collection of developments being dubbed "government-binding" or GB theory. Constraints were eventually reformulated in terms of grammatical relations defined on syntactic trees. It was discovered, for example, that many of the constraints individuated for various constructions could be usefully explained by appealing to a simple c-command relation,[42] defined in (31) in terms of dominance of nodes in trees.

(31) c-command
X c-commands Y if and only if
(a) X does not dominate Y and Y does not dominate X; and
(b) the first branching node dominating X also dominates Y.

This notion is re-deployed in defining the relation of *government* in (32).

(32) Government
X governs Y if and only if
(a) X is a governor; and
(b) X c-commands Y and Y c-commands X.[43]

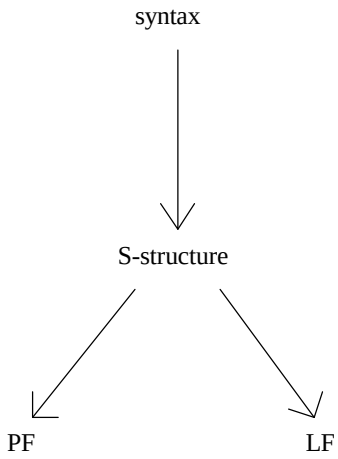
In (31) and (32), X and Y correspond to any phrasal head (N, V, P, A, . . .). More recent proposals differentiate between lexical heads (N, V, P, A, Adv) and functional heads, as will be discussed in section 2.4.3.

[42]Originally constraints on rules were presented as a list of descriptive facts. Ross (1967) is credited with discovering the central corpus of constraints which defined the research program in generative grammar into the 1980s.

[43]For examples of these principles at work, and for a definition of governor, used in (32), see Haegeman (1991) or Webelhuth (1995).

The classic GB framework may be illustrated in Figure 1, from Chomsky (1981: 17). The rules of the syntax generate S-structures (surface structures). Interpretive components then map S-structure to PF (phonetic form) and to LF (logical form).

Figure 1: The GB Framework



In this system, Universal Grammar (UG) was taken to be comprised of principles, invariant for all languages, and of parameters, switches which define the hypothesis space for children acquiring their language. In GB theory, parameters were assumed to be options within the computational component of the grammar. For instance, in discussing the *pro*-drop parameter, Hyams (1986) proposed that children assume their language to be a *pro*-drop language (that is, to allow sentences with phonologically null subjects, like Spanish and Italian) until positive evidence to the contrary is obtained, as when English-speaking children observe the use of pleonastic subjects like *it* and *there*, absent from *pro*-drop languages. The acquired parameter setting applies to the computational system, determining whether or not a subject may be null.

As linguists struggled to further remove highly redundant and unconstrained machinery from the theory of syntax, this conception of parameters was significantly revised. An idea was developing that all variation, and hence all learning, was associated with abstract and concrete morphological properties of the lexicon.

2.4.3 The Minimalist Program

Chomsky (1991: 23) commented on the promise of a syntactic theory in which parameters are restricted to the lexicon:

> If there were only one human language, the story would essentially end there. But we know that this is false, a rather surprising fact. The general principles of the initial state evidently allow a range of variation. Associated with many principles there are parameters with a few--perhaps just two--values. Possibly, as proposed by Hagit Borer, the parameters are actually restricted to the lexicon, which would mean that the rest of the I-language is fixed and invariant, a far-reaching idea that has proven quite productive.

Restricting parameters to the lexicon means that linguistic variation falls out of just the morphological properties (abstract and concrete) of the lexicon (Borer, 1983). In this model, then, there are two central components: C_{HL}, a computational system for human language, which is presumed to be invariant across languages, and a lexicon, to which the idiosyncratic differences observed across languages are attributed. In addition, note that the suggestion that the I-language is fixed and invariant in this way introduces a version of the Universal Base Hypothesis, the idea that phrase structure does not vary across languages; surface differences in word order relate only to the re-arrangement of elements in the syntactic tree as the result of movement operations, triggered by lexically encoded morphological features.

Even phrase structure is derived from the lexicon in the minimalist program. An operation, which may be called *Select,* picks lexical items from the lexicon and introduces them into the numeration, an assembled subset of the lexicon used to construct a derivation. Another operation, *Merge*, takes items from the numeration and forms new, hierarchically arranged syntactic objects. The operation *Move* applies to syntactic objects formed by *Merge* to build new structures; it forms Δ from κ and α (κ the target of movement and α the element affected by movement) by replacing κ with $\{\Gamma, \{\alpha, \kappa\}\}$ ($=\Delta$) (Chomsky, 1995b) (see Stabler (1997b) for a slightly different minimalist model). Hence, in the Minimalist Program, phrase structure trees are built derivationally by the application of the three operations *Select, Merge* and *Move,* constrained only by the condition that lexically encoded features match in the course of a derivation. Phrase structure, along with configurationally defined intermediate and maximal projections, therefore has no independent status in C_{HL}, eliminating (30) of section 2.4.2.

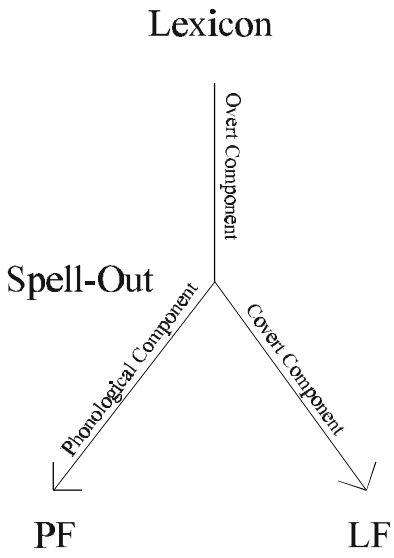
Movements are driven by feature checking, and may be of two types: A head may undergo *head movement* and adjoin to another head, or a maximal projection may move to the specific position of a head; in either case, the element moves for the purpose of checking morphological features of case, number, person, and gender. In addition, its movement may be *overt* or *covert*. Overt movements are driven by *strong* features and are visible at PF (*phonetic form*, formally known as "the surface structure") and LF (*logical form*, the interpretive level). Covert movements, driven by *weak* features, are visible only at LF.

Principles of Economy select among convergent derivations. One such principle, Full Interpretation (FI), requires that no symbol lacking a sensorimotor interpretation be admitted at PF; applied at LF, FI entails that "every element of the representation have a (language-independent) interpretation" (Chomsky, 1995a: 27). Thus, uninterpretable features (denoted -Interpretable) must be checked and (in some proposals) deleted by LF. Such features include case, person, number and gender. Other Economy Principles will be discussed in the exposition in chapter 5.

A derivation is said to *converge at an interface level* (PF or LF) if it satisfies FI at that level; it *converges* if FI is satisfied at both levels. A derivation that does not converge is also referred to as one that *crashes*. If features are not checked, the derivation *crashes*; if they *mismatch*, the derivation is canceled (that is, a different convergent derivation may not be constructed).

At some point in the derivation, an operation Spell-Out applies to strip away from the derivation those elements relevant only to PF; what remains is mapped to LF by a subsystem of C_{HL} called the *covert component*. The elements relevant only to PF are mapped to PF by operations unlike the covert component, operations which comprise the *phonological component.* The phonological component is also regarded as a subsystem of C_{HL}. The subsystem of C_{HL} which maps the lexicon to Spell-Out is the *overt component* (often called overt syntax). Figure 2 outlines the relationships between these components in the Minimalist Program, in analogy to the schematic of GB Theory presented in Figure 1 (page 84). Note that the various components (overt, covert, phonological) are all part of C_{HL}, the computational system for human language.

Figure 2: The Minimalist Framework



The Minimalist Program, although extremely promising in many respects, should be regarded as a developing research program. As such, a number of proposals have been made in which some of the details of the system proposed in Chomsky (1995a) are reworked. Of particular interest are Lasnik (1995a, 1995b, 1995c), papers in Abraham, Epstein, Thráinsson and Zwart (1996), Ferguson (1997), Ura (1997), Stabler (1997b), among others. Radford (1997) provides an excellent introduction to core topics.

Of course, just as in other milestones in linguistic research, the Minimalist Program is based upon a long history of work in the theory of syntax, much of which has served as an important transition to the new framework. In chapter 5, I will discuss proposals relevant to the data of chapter 4, focusing on minimalist approaches and, where possible, testing competing theories with the newly discovered data.

A very important aspect of minimalism is that all learning is lexical, and all parameters are microparameters associated with individual lexical items. This makes a

rather different conception of bilingualism possible, since it is no longer necessary to regard grammars as compartmentalized in some way in the language faculty. In the minimalist framework, C_{HL} is invariant across languages, and the Lexicon does not need to be privy to sociopolitical distinctions like Spanish, Nahuatl, and Chinese. The difference between an SVO language like English and an SOV language like Korean, for instance, is defined in terms of the strength of features in the object DP (weak in English, strong in Korean); apart from this sort of parametric variation, there are no differences in the rules of syntax, allowing a great simplification in our conception of bilingualism.

But the Phonological Component, responsible for mapping the numeration to PF, is of a very different character. These rules build structure in a way that syntax does not, and in doing so they often refer to specific morphological material with its phonetic content. The impact of these facts will be explored extensively in chapter 5.

2.5 Nahuatl and Spanish

It will also be of some value to gain a preliminary sense of what resources are available for the grammatical study of Spanish and Nahuatl. Here I will also address the historical and genetic relationships between the two languages before reviewing the linguistic literature relevant to each.

2.5.1 Genetic and Typological Relationships

After applauding the development of "a substantial data base with respect to both the formal and functional aspects of codeswitching," McClure (1981) notes an important concern:

> This data base is affected by one problem. It is heavily skewed toward studies involving alternate use of Spanish and English in the United States. Studies involving other languages, particularly non-Indo-European ones are necessary in order to test the cross-linguistic validity of the constraints which have been formulated [69, n1].

The Spanish-Nahuatl corpus presented in chapter 4 represents code switching data involving a genetically and typologically unrelated pair of languages, and may indeed provide additional insights regarding the theory of code switching.

Linguists have classified languages according to their genetic (or *historical*) closeness as well as their typological similarities.[44] The genetic method is concerned with historical relationships between languages, while typologists are interested in similarities between language groups regardless of their historical connections. For instance, English is historically close to modern German, but typologically it may be much closer to other analytic languages like Chinese or Samoan with which it has no known historical relationship. A look at the historical and typological relationships between Spanish and Nahuatl may be useful in light of McClure's comments quoted above.

According to Lamb (1964), Nahuatl is a member of the Uto-Aztecan family, which extends geographically over a vast area of Mexico and the United States, and contains the subfamilies of Numic, Tubatulabal, Hopi, Takic, Pimic, Taracahitic, Corachol, and Aztecan. The northernmost Uto-Aztecan language, Paiute, may be found

[44]Areal classifications group languages according to their similarities established within geographical areas, such as "the Scandinavian languages" or the "London-influenced dialects." This system of classification, preferred by some, intersects with both the historical and typological systems which I will focus on here. (See Comrie (1981) and Crystal (1987) for discussion.)

as far north as Idaho and Oregon, and members of the Aztecan subfamily are spoken in Central Mexico and as far south as Nicaragua. The Aztecan subfamily contains Pochutla and Classical Nahuatl. It is estimated that Proto-Uto-Aztecan has a time depth of about 5000 years, about the same time depth assumed for Proto-Indo-European (Langacker, 1977).

Although it is generally thought that Uto-Aztecan is a distant relative to the Kiowa-Tanoan family, Langacker (1977) reports that other connections are regarded as highly speculative. As is well known, Spanish is a member of the Romance subfamily of Indo-European in the Italic branch. Thus, there is no known historical relationship between Spanish and Nahuatl, although these languages have influenced each other's vocabulary following the Conquest; I will return to this topic in section 2.5.3.1.

Typologically, Spanish is usually classified as a synthetic language, one in which such features as case and agreement are marked morphologically. Nahuatl, on the other hand, is a polysynthetic or incorporating language; in languages of this type, a single word may express the subject, verb and object of a proposition. Sapir (1921), who did pioneering work on polysynthetic languages in America, described them as a class of languages in which

> The elaboration of the word is extreme. Concepts which we should never dream of treating in a subordinate fashion are symbolized by derivational affixes or "symbolic" changes in the radical element [the stem], while the most abstract notions, including the syntactic relations, may also be conveyed by the word [128].

Baker (1996) characterizes a polysynthetic language as one with productive and full noun incorporation (NI) *and* full and obligatory agreement paradigms for both

subject and object. These properties are built into his Polysynthesis Parameter, a topic I will touch upon in chapter 5. Nahuatl, as Baker (1996) points out, is a clear member of the class of languages he defines as polysynthetic. So far as I know, all of the modern dialects of Nahuatl, as well as the classical language, share these properties (see section 2.5.3.1).

The degree to which languages are viewed as typologically distinct is, of course, very much a function of the analysis of their morphological and syntactic structure. On Baker's (1996) approach, polysynthetic languages are radically different in structure from Romance languages, particularly on the assumption that all NPs are adjuncts in the polysynthetic group. However, taking the rich verbal morphology of polysynthetic languages to be reflexes of clitics and/or agreements, the peculiar properties of polysynthetic languages significantly diminish in number.

Table 2, which is based on some of the properties listed in Baker (1996: 498-499) for polysynthetic languages, presents some descriptive characteristics of Spanish and Nahuatl typology. The intention here is to provide a sense of important language differences and similarities which will play a role in the discussion in chapter 5.

Table 2: Some Typological Characteristics of Spanish and Nahuatl[45]

Property	*Spanish*	*Nahuatl*
syntactic noun incorporation	no	yes
object agreement obligatory	no	yes
subject agreement obligatory	yes	yes
free pro-drop	yes	yes
SVO word order	yes	yes
NP reflexive	yes	yes
true quantifiers	yes	yes
obligatory wh-movement	yes	yes
true determiners	yes	yes
infinitives	yes	no

Notice that, with regard to the properties listed, Spanish and Nahuatl only differ in terms of productive noun incorporation, object agreement morphology, and the availability of infinitival constructions. There are certainly other differences, but these are fairly basic.

2.5.2 The Spanish Language

Spanish is one of the most extensively studied languages in the world. Claudia Parodi, for instance, has written two full volumes surveying recent studies in Spanish linguistics by Mexican linguists alone (1981, in press), and Linguistic Abstracts lists

[45]Although a detailed discussion is beyond the scope of this dissertation, I disagree with many of Baker's (1996) views regarding the descriptive typological characteristics of Nahuatl. For instance, although Baker classifies Nahuatl as favoring V-initial word order, my experience with Guerrero and Southeast Puebla Nahuatl suggests a strong SVO preference with considerable freedom for dislocation (see section 2.5.3). Also, virtually every modern variety of Nahuatl has fully borrowed the Spanish quantifier *cada* 'each' (see section 2.3.1), so it is incorrect, at least for the modern varieties, to claim that Nahuatl has no true quantifiers. The difference in views may in part be due to dialectal differences, particularly between the modern and classical varieties, or it may follow from disagreements in analysis among the descriptive grammarians whom Baker relied upon in his study. In any event, the typological characteristics in Table 2 are at odds in some respects with Baker's Table 11-1.

more than 350 articles on Spanish written just since 1985.[46] In chapter 5, I will draw upon these and other resources as necessary in my discussion of the data.

2.5.3 The Nahuatl Language

Hill and Hill (1986) characterize Nahuatl as the most extensively studied indigenous language of Mexico. Southeast Puebla Nahuatl has productive noun incorporation; the verb takes an object prefix and a subject prefix, and tense is marked variously with a prefix (*o-*, PAST, for example) or suffix (*-s*, FUTURE). Both the subject and object may be dropped. Agreement and tense are marked on complement clauses as well as matrix clauses, distinguishing Nahuatl and other polysynthetic languages from languages which use tenseless (infinitival) verb forms in embedded IP complements.

Consider, for instance, the Nahuatl expression in (33a).

(33a) Nikneki nikoas tlakemetl
ni-k-neki ni-k-koa-s tlake-me-tl
1S-3Os-want 1S-3Os-buy-FUT garment-PL-NSF
'I want to buy some clothes'

Subject agreement is marked on the matrix verb as well as the embedded verb with *ni-*. Transitive verbs also require object agreement, so the prefix *k-* is used for both verbs ('want' and 'buy'). Here the idea is that the act of buying will take place in the future, so future tense is marked on the verb of the IP complement. Nouns, which are not marked for case, often take an absolutive (or default) suffix *-tl*, denoted NSF here (for 'noun suffix,' following Baker (1996)).

[46]The electronic version of Linguistics Abstracts is available from the Blackwell web page at http://www.blackwellpublishers.co.uk/labs/.

Nahuatl also productively uses noun incorporation (NI), like many polysynthetic languages. In (33b), for instance, the noun stem *tlake* 'garment' is incorporated into the verb, displacing the object agreement prefix *k*-.

(33b) Nikneki nitlakekoas
ni-k-neki ni-tlake-koa-s
1S-3Os-want 1S-garment-buy-FUT
'I want to buy some clothes'

A comparison of the constructions in (33) indicates that NI is optional in Nahuatl. (These examples are from my data; see Merlan (1976) and Hill and Hill (1986) for others.)

In addition, as (33a) suggests, Southeast Puebla Nahuatl is an SVO language; however, VSO word order also sometimes occurs, as in Spanish. Tuggy (1979), Brockway (1979), and Sischo (1979) characterize Nahuatl as an SVO language too, and Brockway adds that the "order of major constituents is relatively fixed ..." (146). Beller and Beller (1979), on the other hand, classify Huasteca Nahuatl as "most commonly" VSO with flexibility for SVO and even VOS. Launey (1992: 36-37) similarly regards classical Nahuatl as VSO, adding that topicalization of the subject (SVO), object (OVS) or both (SOV) is also possible. Hill and Hill (1986) view Malinche Nahuatl as VSO; however, SVO, OVS and SOV also occur in their data (VSO = 5, SVO = 6, OVS = 3, and SOV = 1).

I studied five short essays written by residents of San Sebastián in order to gain a clearer sense of Southeast Puebla Nahuatl basic word order; these findings are presented in Table 3.[47]

Table 3: Frequencies of Nahuatl Basic Word Orders in Five Texts from San Sebastián Zinacatepec, Puebla

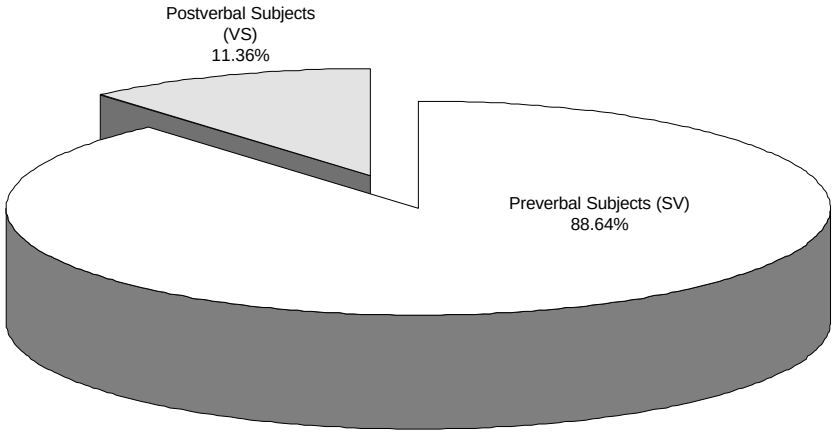
Word Orders	***Text #1***	***Text #2***	***Text #3***	***Text #4***	***Text #5***	***Totals***	***Relative Frequencies***
V	17	28	27	32	9	***113***	***42.80%***
VO	7	11	19	19	7	***63***	***23.86%***
SV	1	6	24	4	1	***36***	***13.64%***
SVO	0	10	23	5	4	***42***	***15.91%***
VS	5	1	1	0	2	***9***	***3.41%***
VSO	0	0	1	0	0	***1***	***0.38%***
Totals	***30***	***56***	***95***	***60***	***23***	***264***	***100.00%***

Texts used: (1) "Sej Welta Onia Acapulco," by Lidia Cedillo Hernández; (2) "Sej Welta Onechmak Tokes," by Jaime Pablo Ignacio; (3) "David iwan Galiat," by Antonia Cortés Hernández; (4) "Pancho Loko," by Arnulfo Prado Bernardo; (5) "Chokotzi Non Okinekia Tekitis," by Isaías Feliciano Santiago.

As in Hill and Hill (1986), Vs occurred very frequently without an overt subject or object, making it difficult to detect a basic word order. However, subjects occurred with verbs in SV, SVO, VS and VSO constructions preverbally nearly 90% of the time and postverbally only about 10% of the time, as shown in Figure 3. Objects *never* occurred before verbs. When subject, verb and object were all overtly present, SVO occurred about 98% of the time and VSO about 2% of the time (Figure 4).

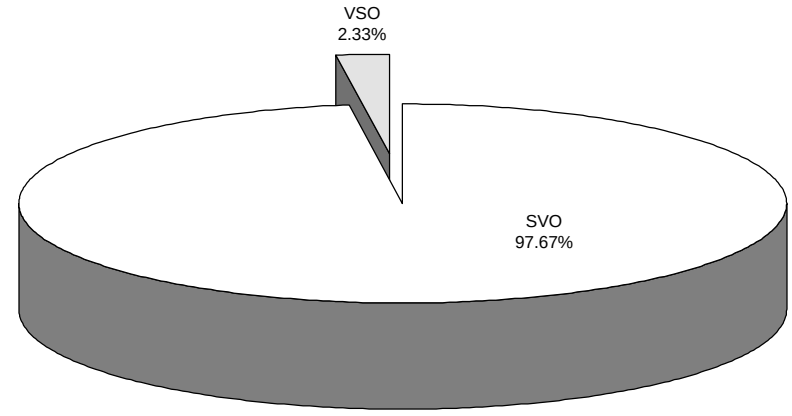
[47]Constructions involving *wh*-elements, negative QPs, passives and *be* were excluded since these exhibit syntactic peculiarities. Complement clauses and quoted speech were counted as objects of their verbs.

Figure 3: The Relative Frequencies of Word Orders for Subjects When Overt in Five Nahuatl Texts from San Sebastián Zinacatepec



n = 79.
Texts used: (1) "Sej Welta Onia Acapulco," by Lidia Cedillo Hernández; (2) "Sej Welta Onechmak Tokes," by Jaime Pablo Ignacio; (3) "David iwan Galiat," by Antonia Cortés Hernández; (4) "Pancho Loko," by Arnulfo Prado Bernardo; (5) "Chokotzi Non Okinekia Tekitis," by Isaías Feliciano Santiago.

Figure 4: The Relative Frequencies of Word Orders for Subject, Verb and Object When Overt in Five Nahuatl Texts from San Sebastián Zinacatepec



n = 38.
Texts used: (1) "Sej Welta Onia Acapulco," by Lidia Cedillo Hernández; (2) "Sej Welta Onechmak Tokes," by Jaime Pablo Ignacio; (3) "David iwan Galiat," by Antonia Cortés Hernández; (4) "Pancho Loko," by Arnulfo Prado Bernardo; (5) "Chokotzi Non Okinekia Tekitis," by Isaías Feliciano Santiago.

Native speakers of Southeast Puebla Nahuatl also confirm this basic idea. The SVO order in (34a) is said to be "most natural," while VSO in (34b) and SOV in (34c) were judged to be possible for focus or contrast only. OVS in (34d) is unacceptable.

(34a) Ne niktlasojtla in Maria
ne ni-k-tlasojtla in Maria
I 1S-3Os-love IN Maria
'I love Maria'

(34b) Niktlasojtla ne in Maria
ni-k-tlasojtla ne in Maria
1S-3Os-love I IN Maria
'I love Maria'

(34c) Ne in Maria niktlasojtla
ne ni-k-tlasojtla in Maria
I 1S-3Os-love IN Maria
'I love Maria'

(34d) *In Maria niktlasojtla ne
ne ni-k-tlasojtla in Maria
I 1S-3Os-love IN Maria
'I love Maria'

I will therefore follow Tuggy (1979), Brockway (1979), and Sischo (1979) and take modern Nahuatl to have a basic SVO word order but also to allow a postverbal subject, like Spanish and other European languages, and a contextually restricted use of object focus, as in (34c). Of course, there may also be dialectal variation on this point, but all of the data in Table 3 are consistent with this view, so it appears quite reasonable at least with respect to Southeast Puebla Nahuatl.

This picture of Nahuatl word order is consistent with a conventional analysis of its phrase structure, departing from proposals in Jelinek (1984) and Baker (1996). There, Nahuatl is presumed to have non-configurational properties (free word order) which

might be easily explained if its NPs are assumed always to occupy adjunct positions. Since Nahuatl word order is in fact relatively fixed, the core motivation for this assumption is lost. In this dissertation, then, I will presume Nahuatl to have phrase structure that is minimally different from Spanish, English and other well-studied languages.

2.5.3.1 Varieties of Nahuatl

Nahuatl has undergone a process of dialectalization throughout Central Mexico, to the extent that many varieties are only marginally mutually intelligible (Egland, Bartholomew and Cruz Ramos, 1978). Relying upon common characteristics of Nahuatl dialects, Yolanda Lastra de Suárez (1986) tentatively suggests an area typology of four major dialects. In her system, the West Periphery consists of the subareas of the West Coast, Western Mexico State, and Durango-Nayarit. The East Periphery consists of the Puebla Mountains, Istmo, and Pipil. Huasteca, regarded by Una Canger (1980) to be part of the Central dialect, is a discrete dialect area with no subareas in Lastra de Suárez' (1986) system. The Central dialect consists of the "subárea nuclear," Puebla-Tlaxcala, Xochiltepec-Huatlatlauca, Southeast Puebla, Central Guerrero, and Southern Guerrero.[48]

Differences between dialects are often present even in very basic vocabulary. In Central Guerrero, for instance, the singular personal pronouns have the form *nehua* 'I', *tehua* 'you' and *yehua* 'he/she,' while Southeast Puebla Nahuatl uses the reductions *ne*, *te*

[48]The labels used here for dialect areas are my translations of terms used in Lastra de Suárez' Spanish-language discussion. See her work for the original text.

and *ye*, respectively. García Escamilla (1993-1995), incidentally, lists these forms, together with *nehuatl/tehuatl/yehuatl*, as variants used in classical times as well. Flores Farfán (1992) provides an interesting discussion showing surprising differences in the pronunciation of basic vocabulary even among the geographically close varieties of the Alto Balsas region.

There are some small syntactic differences as well. The Central and Southeast Puebla dialects use *amo* for negation, placed before the verb, but in Central Guerrero negation is expressed using a verbal prefix *x-* (also available as *mach-* in Southeast Puebla). Also, as pointed out to me by David Tuggy (personal communication), Tetelcingo and Orizaba Nahuatl have different preferences for ordering of a possessed noun and a nominal possessor (so, *ikni Juan* or *Juan ikni* is "John's brother" in Tetelcingo Nahuatl, but in Orizaba Nahuatl *ikni Juan* means "his brother John" and *Juan ikni* means "John's brother"). Tuggy also provides the example of relative clauses; in Tetelcingo Nahuatl, these usually involve an overt subordinator *ke*, but *tli(n), aaki* or *lo ke tli(n)* are also available; Orizaba Nahuatl can use *tlen* or *akin*, but also has a range of clauses unmarked or only marked with *n* (reduced from *in*). Other kinds of subordinate clauses are also often unmarked or marked differently. The preferred order of some of these clauses with respect to the main clause also may be different.

While these differences are of some interest, they do not indicate far-reaching typological differences across dialects. In all modern varieties of Nahuatl, subject agreement morphemes are obligatorily marked on verbs, object agreement morphemes on transitive verbs, and NI is fully productive for the special subset of nouns known to

undergo NI in other polysynthetic languages. As mentioned in the previous section, there is a small amount of evidence that there may be differences in basic word order across dialects, but this claim is so far unsubstantiated and I suspect false.

The influence of Spanish on Nahuatl is also worth discussing. Hill and Hill (1986) argue that Nahuatl has undergone a "syncretism" with Spanish. Specifically, they suggest that, where there are options allowed by its morphology and syntax, Nahuatl has gone the way of Spanish, favoring the Spanish construction type whenever possible. A Spanish "way of talking" has emerged for many Nahuatl bilinguals, according to the Hills, but this different style does not reflect morphological or syntactic changes in Nahuatl grammar per se; so, for them, Nahuatl *speech* has become like Spanish, but the language itself has not changed in terms of its morphology or syntax.

The Hills offer noun incorporation (NI) as an example of an optional construction which now favors the Spanish pattern in Malinche Nahuatl (that is, according to the Hills, these speakers now tend to use unincorporated structures, as in Spanish). In the Hills' data, only 126 verbs exhibited NI in a total of 16,025 verbs in their corpus. However, NI in polysynthetic languages is known to be severely restricted, as the Hills note; in particular, only non-specific, inanimate nouns in object position may undergo NI (Baker, 1988). Therefore, a more telling measure of NI productivity would be the ratio of NI to non-NI constructions which involve non-specific, inanimate object nouns. Thus, it is not clear that Nahuatl has changed even in this stylistic way. NI is still allowed by every variety of Nahuatl; moreover, if the language did begin to use only unincorporated, SVO constructions, we could not safely conclude that the change has resulted from contact

with Spanish since polysynthetic languages tend to become analytic over time regardless of language contact phenomena (Dressler, 1988).

Suárez (1977), on the other hand, in the spirit of Boas (1966) and others, has studied "the effect that the borrowing of grammatical words has produced in the structure of the [Nahuatl] language" (115; translation mine). He examines the structural effects of twenty Spanish functional elements, borrowed into Nahuatl.[49] The borrowed functional items, however, generally reflect phonological changes as the result of contact with Spanish. For instance, while precolonial Nahuatl used *in* to mark complement clauses, as in (35), many modern varieties use Spanish *ke*, as in (36) (examples from Suárez (1977: 146-147; translations mine)).

(35) Okittak *in* ye itatla
'He saw *that* it was burning'

(36) Kineltokakeh *ke* inon se toteko
'They believed *that* he was a god'

However, modern varieties of Nahuatl allow *ke* to be omitted, much as English does with *that*; this is also like precolonial Nahuatl *in* and *unlike* Spanish *que*, indicating that Nahuatl borrowed the phonetic shape of the Spanish complementizer but not its morphosyntactic properties. Thus, while many of Suárez' very interesting examples show that Nahuatl obtained new lexical resources as a result of its Spanish borrowings, they do not seem to suggest far-reaching changes in the syntactic and morphological systems of the language.

[49]Suárez' (1977) list of items is presented in section 2.3.1.

Other works on Nahuatl dialectology, involving a detailed and fascinating analysis which goes beyond the scope of this dissertation, include Dakin (1975, 1976) and Valiñas (1979, 1982, 1997). On historical changes in Nahuatl, especially its phonology, see Dakin's extensive studies (1982, 1995).

2.5.3.2 Nahuatl Courses and Linguistic Studies

There have been many recent publications of Nahuatl courses, grammatical sketches, and various other linguistic studies. Courses focusing on various modern dialects include Beller and Beller (1976, 1979), Horcasitas (1977), Calarza and López Avila (1987), and Tuggy (1990). Grammatical sketches of some modern varieties have been provided by Whorf (1946), Tuggy (1977), Beller and Beller (1977), and Wolgemuth (1981). Andrews (1975), Launey (1981)--translated into Spanish as Launey (1992)--Campbell and Karttunen (1989), and García Escamilla (1993-1995) are introductions to the classical language. The latter is designed essentially for children and consists of fifteen small booklets.

Siméon (1996 [1885]) and Karttunen (1983) are excellent dictionaries of classical Nahuatl. Brewer and Brewer (1971) and Matias Alonso and Medina Lima (1996) have put together *vocabularios* of Tetelcingo and Guerrero Nahuatl respectively. Alva Hernández (1996) has constructed a dictionary of Coxcatlán Nahuatl, the only work known to me which deals specifically with Southeast Puebla Nahuatl, the dialect presented in this dissertation. In addition, the Summer Institute of Linguistics has produced several translations of the New Testament into a number of modern varieties; apart from these, there are few texts of the contemporary language available.

León-Portilla (1988) has put together an impressive annotated bibliography of nearly everything written on Nahuatl language and culture. Volume 1 of his work is a historical discussion of scholarly work on Nahuatl, and volume 2 consists of 2,961 annotated references on various aspects of Aztec life and Nahuatl language.

More narrowly focused discussions of various aspects of Nahuatl grammar may be found in Merlan (1976) and Baker (1996). Launey (1994) advances the notion that the key typological distinction between Nahuatl and other languages is that its phrases are all predicates. These discussions rely heavily upon the more extensive, general literature on polysynthetic languages which I will discuss, together with other relevant proposals, in chapter 5.

2.5.3.3 Nahuatl Orthography

As one might expect, there is no uniform convention for writing modern Nahautl. Originally the Aztecs wrote their language in hieroglyphics, where each hieroglyph represented a full word, and had amassed an extensive collection of historical and administration literature by the time of the Conquest, most of which was destroyed by Spanish clergymen for its "satanic" content. One example, the Codex Mendoza, survives in the Bodleian Library in Oxford. See Coulmas (1996: 29-33) and León-Portilla (1988) for additional information and examples of Aztec hieroglyphics.

Contemporary Nahuatl is written within two distinct traditions, with some communities mixing these. Many of the classical scholars use fifteenth-century Spanish orthography, while others have been influenced by the International Phonetic Alphabet (IPA) used by fieldworkers and/or a desire for distance from Spanish culture. As an

example of the difference in orthography, the letter *c* is often used for /k/ and /s/ by those who prefer the Spanish system, while many others simply use *k* and *s*. However, even among the Spanish-influenced orthographies many Nahuatl borrowings are spelled with *k* (*ke* < *que* 'that,' *kuando* < *cuando* 'when,' *porke* < *porque* 'because,' and others; see Suárez (1977)). For an example of the traditional, Spanish-influenced system, see García Escamilla (1993-1995); for an example of the IPA-like system, see Beller and Beller (1976, 1979). In section 3.3.3, I will discuss the conventions used by my consultants, preserved in the transcriptions of data in chapter 4.

2.6 Spanish and Nahuatl in Central Mexico

Although Nahuatl is now a minority language throughout Mexico, it was once the *lingua franca* of a vast empire. In contemporary Mexico, the great achievements of art and architecture of the precolonial world are publicly displayed in federal museums, while the indigenous people themselves inhabit the poorest and least developed parts of the country. A review of the historical and political circumstances leading up to Spanish-Nahuatl bilingualism today appears in section 2.6.1, followed by a look at some of the relevant demographics of Spanish and Nahuatl in contemporary Mexico in section 2.6.2.

2.6.1 The Aztecs and Hernán Cortés[50]

Nahuatl, also known as Mexicano or Aztec, was the language of the Aztec Empire, which dominated Central and Southern Mexico from the fourteenth to the

[50]The historical sketch presented here derives from Turner (1911), Katz (1981), Fuentes (1992), Casío Villegas (1996), Zinn (1995), and Chomsky (1993b).

sixteenth century. Following the fall of the Toltec civilization, waves of northern immigrants from Aztlan (possibly located in modern-day Arizona), whose tribes were called Nahuatlaca or Mexica,[51] flooded into Mexico's central plateau area around Lake Texcoco. As the Aztecs grew in number, they established superior military and civil organizations. Finally, by 1325, they founded the city of Tenochtitlán, now called *México* or Mexico City.

The Aztecs formed military alliances with other tribes, creating an empire that extended from Central Mexico to the Guatemalan border. In the early fifteenth century Tenochtitlán jointly governed with the city-states of Tlatelolco and Texcoco. However, within a century of the triple alliance, the Aztecs seized complete control; while kingships remained in the other city-states, these eventually became mere honorary titles.

The Aztec Empire was theocratic, with the high priest functioning as the seat of absolute authority. In ritual religious ceremonies, prisoners of war (and volunteers!) would climb the steps of the pyramid where priests stretched them across a convex stone and sliced their hearts out with a sharp knife. Children, too, especially twins, were sacrificed to the gods, often in the cruelest manner imaginable. On one occasion, in an effort to stave off the end of the world, sixty-thousand prisoners were brutally sacrificed to the Sun in a single day; the blood is said to have filled the streets of the city.

The Aztecs built enormous constructions with stone tools and human labor, developed a pictographic writing system inscribed on paper and animal hides, a calendar

[51][mɛšika]

modeled on Mayan achievements, an irrigation system, and developed a number of important mathematical concepts. However, near the end of the empire, the Aztec leaders, whose people's language was used throughout their vast conquered territories as a *lingua franca*, began to find it extremely difficult to suppress local rebellions.

When a bearded white man and his entourage arrived in large sea vessels, in 1519, clad in iron and riding strange, mysterious beasts (horses), Montezuma II, the Aztec emperor, mistook him for Quetzalcóatl, the legendary man-god who had died three hundred years before with the promise to return.

In fact, this man was Hernán Cortés. He had come from Spain on an expedition financed with the capital of merchants and landowners in search of unprotected gold and biddable slaves. Starting in Veracruz, he began on a death march, moving from town to town, "killing with the kind of deliberateness that accompanies a strategy--to paralyze the will of the population by a sudden frightful need" (Zinn, 1995: 11).

In addition to spoils, Cortés also took many prisoners and slaves along the way. Among them was Malinche, whom Cortés made his lover and spokeswoman. As he overtook native populations on his way to Tenochtitlán, their armies were conscripted by the Spanish. Of these the Tlaxcalans, enemies of the Aztecs, became the most important conquest.

Montezuma pursued an irresolute course of action during Cortés' massacre of neighboring civilian populations under his control. Finally, he decided not to oppose the Spanish invaders but to wait for them at the Aztec capital and learn more about their purposes. On November 8, 1519, Cortés and his small force, including around six-

hundred Spanish conscripted soldiers, entered the city and set up headquarters. Received with the honor due Quetzalcóatl, Cortés' soldiers were allowed to roam the city freely, finding gold and many other treasures in the storehouses.

Cortés then seized Montezuma as hostage and forced him to swear allegiance to the king of Spain and to provide an enormous ransom in gold and jewels. Meanwhile, back in Cuba, the Spanish conqueror Velázquez had become increasingly concerned about Cortés' loyalty and personal ambitions, so he sent a Spanish expedition to investigate. When Cortés received word that challengers were arriving, he left two hundred men at Tenochtitlán under Pedro de Alvarado, and marched with a small force to the coast, attacking the Spanish camp at night and inducing the majority of the soldiers to join his forces.

Meanwhile, the harshness of Alvarado's treatment inspired a popular revolt against the Spaniards and the imprisoned ruler Montezuma. Upon re-entering Tenochtitlán, Cortés and his men were surrounded and attacked. Cortés subsequently persuaded Montezuma to address the people in an effort to quiet the rebellion; however, the angry crowd stoned Montezuma, and he died three days later. On June 30, 1520, the Spanish and their conscripts were driven out of the city by a group of Aztecs, led by Montezuma's nephew Cuauhtémoc, on the *Noche Triste* ("Sad Night"), as the Spanish historians would later call it. The Aztecs pursued the retreating Spanish troops, many of whom drowned in the waters surrounding Tenochtitlán, too heavily weighted down with gold and silver plunder to escape.

On July 7, 1520, after defeating a very large force of Aztecs, Cortés managed to reach Tlaxcala. The Tlaxcalan territory was surrounded by the Aztec Empire but was never actually conquered by its rulers. Cortés convinced the Tlaxcalans to form an alliance with him against the Aztecs. He was thus able to reorganize his army, also benefiting from some reinforcements and equipment from Veracruz. He then returned to the capital, capturing Aztec outposts on the way. On August 13, 1521, after a desperate siege that lasted three months, Cuauhtémoc, the new emperor, was captured, and Tenochtitlán fell.

Cortés ordered that Tenochtitlán be leveled to the ground, and upon its ruins he built Mexico City, capital of New Spain. The Spanish destroyed pyramids and other religious edifices of all types, erecting churches for their new flocks to attend. Virtually every artifact of Aztec culture was assaulted, often eradicated.

For the next four hundred years the indigenous people would live as serfs and wage slaves under the rule of the Spanish and the *criollos*, the white Mexicans of pure European descent who led the War of Independence against Spain. Between 1877 and 1911, under the dictatorship of Porfirio Díaz, large communal lands which the Natives used for agriculture were granted over to wealthy plantation owners. When the Natives complained, they were sold into peonage.

Meanwhile, Emiliano Zapata, Pancho Villa and others were engaged in armed revolts. By 1917, a Constitution was drafted which provided for a labor code, prohibited a president from serving consecutive terms, expropriated all property of religious orders, restored communal lands to the Natives, and curbed foreign ownership of mineral

properties and land. A period of instability followed, however, as the United States and Europe waged a "secret war in Mexico," as Friedrich Katz (1981) has called it, to guarantee access to wealthy businessmen for the purpose of plundering Mexico's natural resources for their gain.

Finally, in 1932, the Partido Nacional Revolucionario (PNR) came to power on the promise that it would develop a cooperative economic system "tending toward socialism." In 1937, the Mexican railway system was nationalized, as were the subsoil rights of the oil companies. In 1936, an expropriation law was passed enabling the government to seize private property whenever necessary for public or social welfare, and when foreign-owned oil companies refused to pay higher wages to striking workers, the Mexican government expropriated the oil properties. A government agency called Petróleos Mexicanos, or Pemex, was created to administer the nationalized industry.

This socialist spirit of the early PNR, however, soon began to disappear. In recent examples, Carlos Salinas and the PNR, which had changed its name to the Partido Revolucionario Institucional (PRI) in 1949, accepted the North American Free Trade Agreement (NAFTA) which effectively assigns Mexico the role of providing the U.S. and Canada with cheap labor for assembly plants, where harsh working conditions, low wages, and the absence of environmental controls offer exceptionally profitable conditions for foreign investors, and little chance for workers to organize (Chomsky, 1993b: 188-189). Similarly, under pressure from the World Bank, Mexico has recently begun to privatize its public holdings, selling off public television stations and setting the stage once again for foreign takeovers of its petroleum resources.

Despite noble beginnings, the PRI has tended to implement only as much social reform as necessary to gain support from the poor and to impede peasant revolts. In 1968, in the presence of CIA "observers," the PRI government ordered the massacre of thousands of peaceful demonstrators, citizens of the republic, just days before the Olympic Games. This was carried out in the middle of Mexico City, and the event was completely ignored by U.S. and Mexican media (although the Games, of course, were fully covered). Today, similarly, under the PRI government, the Mexican White Guard conducts terrorist attacks against civilian peasants in Chiapas, presumed to be part of Marcos' "rabble," who are once again demanding agrarian reforms. Notwithstanding widespread discontent, the PRI has never lost a presidential election in Mexico. Power is maintained by massive vote fraud, intimidation, bribery, and murder. Thus, despite the facade presented in mainstream U.S. and Mexican media, modern Mexico, much less the PRI, is not a genuinely democratic entity, and no trace of its radical origins is now evident.

Recently, on July 6, 1997, Cuauhtémoc Cárdenas, leader of the Partido Revolucionario Democrático (PRD), was elected as Mayor of Mexico City, the second most important political post in the country and a platform from which to campaign for the next presidential election. Although Cuauhtémoc is widely expected to implement pro-social, democratic reforms, some have claimed that his sympathy for workers and

ordinary people is no greater than that of his former colleagues in the PRI.[52] Only time will tell.

The native people of Central Mexico have lived a long history of oppression, first at the hands of the Aztec rulers and later by the Spanish conquerors. Today, the indigenous peoples of Mexico are among the poorest and most marginalized in their country, and working wages continue to fall, increasingly so since NAFTA ($1.38/hour in 1982, $0.45/hour in 1990).[53]

Since the arrival of Hernán Cortés, the *indígenas* and their cultures, stigmatized as the lowest caste of society, have become increasingly outnumbered by their "minority of exploiters," to use Bakunin's (1970 [1883]) phrase. For instance, by 1895, the year of the first general census, just 16.27% of the Mexican population spoke an indigenous language; by 1990, only 6.5% did (INEGI, 1994a). Of the seventy known indigenous languages of Mexico, two--Cuitlateca and Chiapaneco--have now become extinct, and others are nearing extinction (Castañeda, 1990). In section 2.6.2, the contemporary language situation of Spanish-Nahuatl bilinguals will be discussed.

2.6.2 Spanish and Nahuatl in Contemporary Mexico

Nahuatl, like Spanish, is spoken in every state in contemporary Mexico. As indicated in Table 4, Puebla has more speakers of Nahuatl than any other state, with

[52]See "Mexico: For Workers Revolution" in *Workers Vanguard*, number 672, 8 August 1997.

[53]Data cited in Chomsky's (1993b: 188) excellent history of U.S. aggression in Latin America and other regions. Compare INEGI (1994a: 195-337) data on workers' earnings since 1877.

nearly one-third of the national total. The cultural/geographical area known as Huasteca, which consists of the states of Puebla, Veracruz, Hidalgo, Tamaulipas and San Luis Potosí, is home to 79.6% of the Nahuatl speakers in modern Mexico, with the state of Guerrero hosting 9.8%. The remaining 10.6% is scattered throughout the republic with concentrations above 10,000 in the state of Mexico, the Federal District, Tlaxcala, and Morelos. (In this section, the term "federal entities" refers to the Mexican states plus the Federal District, as intended in Mexican goverment publications.)

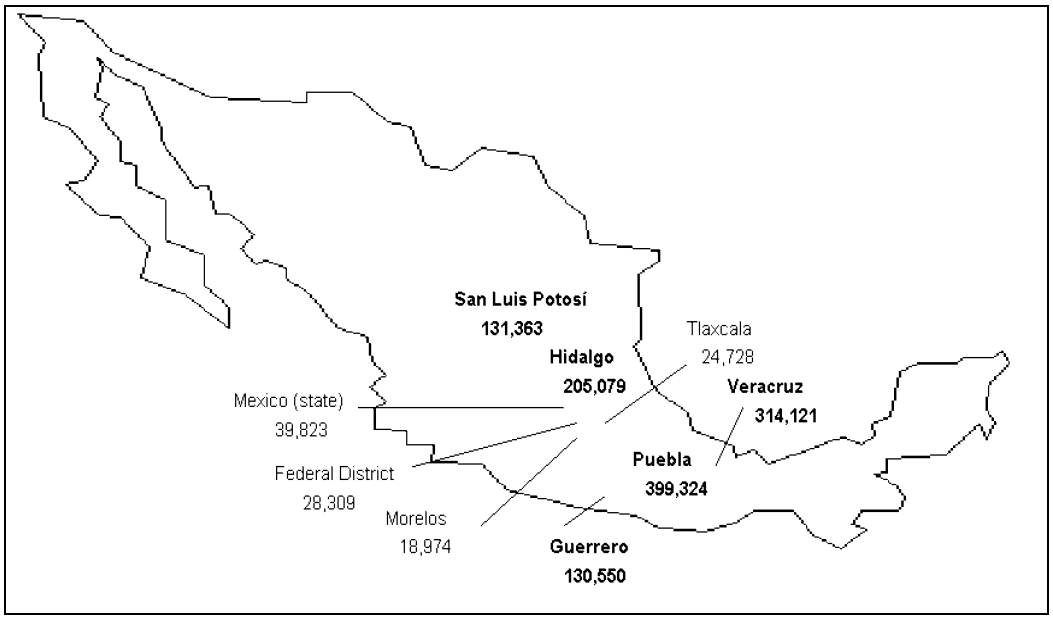
Thus, while the *lingua franca* of the Aztec world may be found throughout Mexico, it is strongly concentrated in Central Mexico, running from the northern coastal city of Veracruz through Mexico City, up to San Luis Potosí and down to the Sierra Madre del Sur in Guerrero. It is by far the most populous indigenous language in Mexico, its speakers making up nearly 23% of the population of speakers of indigenous languages in 1990 (INEGI, 1994a: 139). See the map in Figure 5.

Table 4: Nahuatl Speakers by Federal Entity in Mexico (Ranked by Population), 1995

Mexican Federal Entity	*Nahuatl Speakers*	*Mexican Federal Entity*	*Nahuatl Speakers*
Puebla	399,324	Tabasco	719
Veracruz	314,121	Durando	691
Hidalgo	205,079	Colima	686
San Luis Potosí	131,363	Querétaro	655
Guerrero	130,550	Quintana Roo	639
Mexico (state)	39,823	Guanajuato	492
Federal District	28,309	Nayarit	478
Tlaxcala	24,728	Coahuila	380
Morelos	18,974	Sonora	364
Oaxaca	9,158	Chiapas	355
Tamaulipas	5,072	Campeche	314
Jalisco	3,580	Baja California Sur	288
Nuevo León	3,560	Chihuahua	265
Michoacán	2,785	Zacatecas	173
Sinaloa	1,129	Aguascalientes	153
Baja California Norte	1,106	Yucatán	127
Total speakers (age 5 and older)			*1,325,440*

Mexican census data (INEGI, 1997: 154-171).

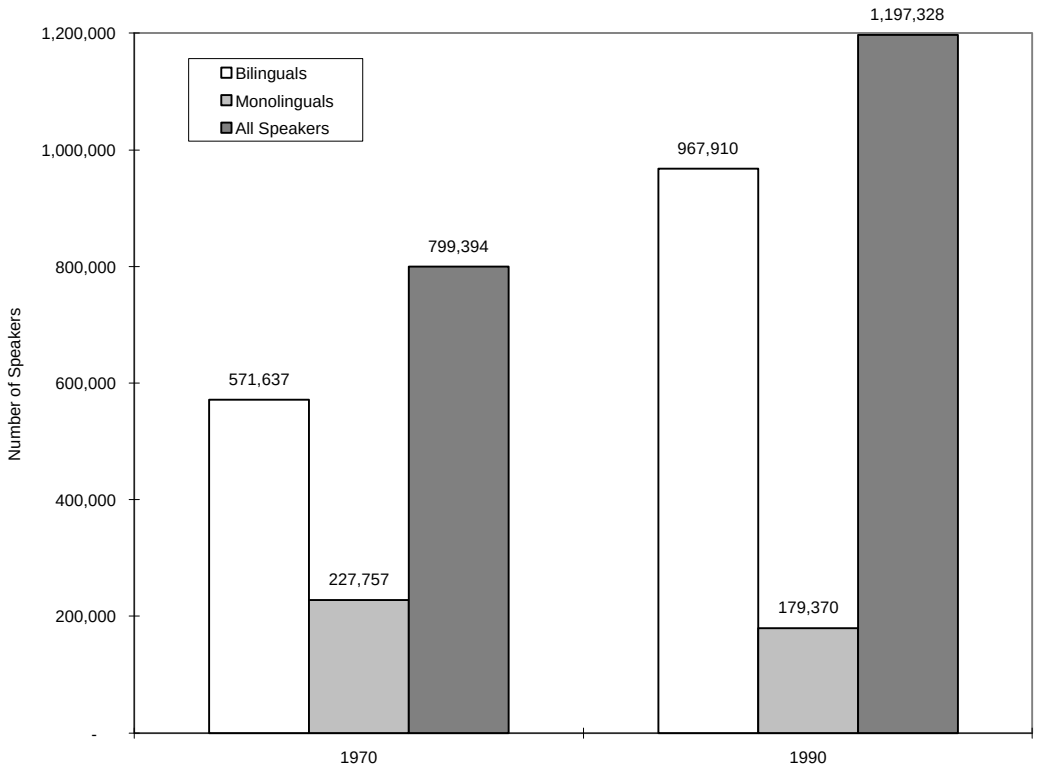
Figure 5: Map of Mexico Showing Federal Entities with 10,000+ Nahuatl Speakers, 1995



Mexican census data (INEGI, 1997: 154-171) (compare Table 4).

In fact, viewed just in terms of raw numbers, Nahuatl speakers in Mexico have been steadily increasing, with bilingualism particularly on the rise, as shown in Figure 6. In 1970, the census office reported that there were 799,394 Nahuatl speakers nationwide, while the most recent census, in 1990, found 1,197,328. Of these, 71.51% were bilingual in Spanish and Nahuatl in 1970, increasing to 80.84% bilingualism by 1990.

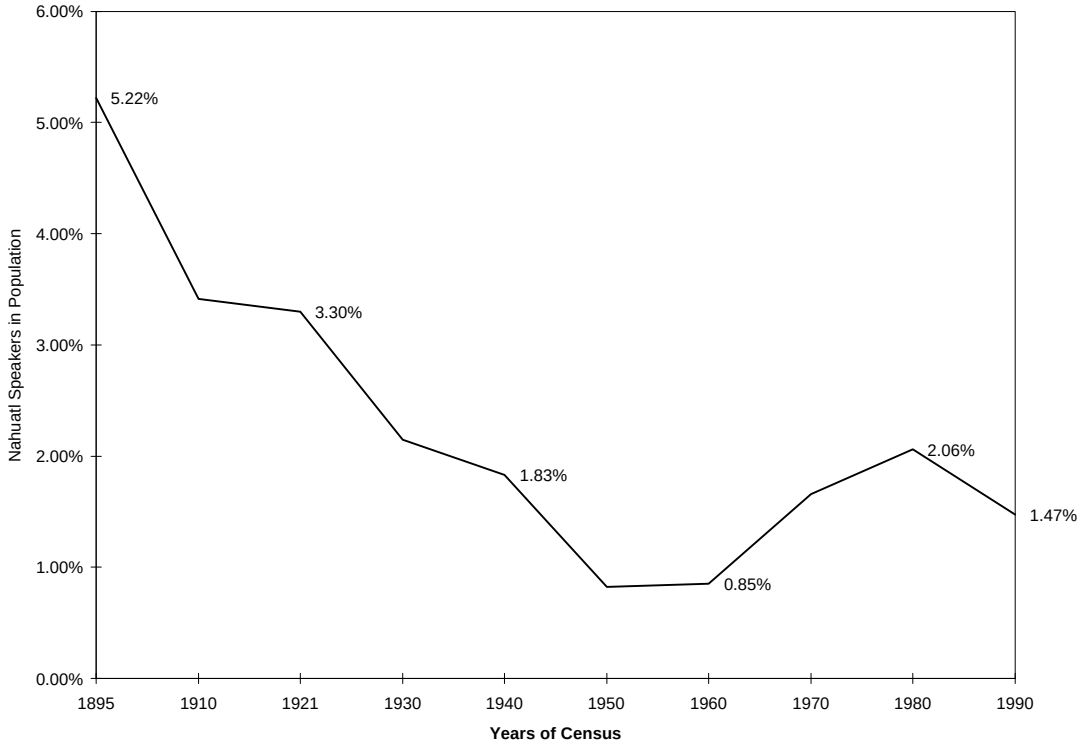
Figure 6: Nahuatl Monolingualism and Spanish-Nahuatl Bilingualism, 1970 and 1990



Mexican census data (INEGI, 1993: 17).
(*Note*: 50,048 Nahuatl speakers did not specify whether they knew Spanish or not in 1990 and are therefore not included in the 1990 monolingual/bilingual breakdown.)

However, the *proportion* of Nahuatl speakers in the entire Mexican population has been steadily declining. Census data, available from 1895 to 1990, indicate that the percentage of Nahuatl speakers in the Mexican population went from 5.22% in 1895 to 1.47% in 1990, as shown in Figure 7.

Figure 7: Percentage of Nahuatl Speakers in Mexican Population, 1895-1990



Mexican census data (INEGI, 1994a: 13, 136-139).

Of course, these figures alone do not suggest *causes* for the decrease. Because populations grow exponentially, it is to be expected that over time the relatively small proportion of Nahuatl speakers would decrease, even if nobody in their communities ever moved away or stopped speaking Nahuatl. A question arises, then, as to whether these data suggest that Nahuatl may one day become extinct, like Cuitlateca and Chiapaneco, or simply proportionally decrease in number of speakers as the general population grows at a faster rate.

Conklin and Lourie (1983) list numerous factors which contribute to language loss cross-culturally; these include occupational shifts requiring migration from rural to

urban areas, denial of ethnic identity to achieve vocational mobility, and lack of mother tongue institutions (mass media, schools, leisure activities). Although their work deals with immigrant rather than aboriginal languages, much of their discussion applies to both situations.

Although many rural Nahuatl speakers are still able to make a living on hereditary plots of land, they face a crisis of shrinking agricultural resources. There is relatively high emigration from Nahuatl areas, and relatively high immigration into the large urban and industrial centers in Mexico. The five states with the greatest number of Nahuatl speakers, for instance, lost a total of 1,513,873 people between 1970 and 1980, and 1,757,682 between 1980 and 1990, due to internal migration (see Table 5). These population shifts, presumably caused by a need for work and/or higher education, cause the sort of population shift which Conklin and Lourie (1983) believe contributes to language loss.

Table 5: Internal Emigration from the Five Most Populous Nahuatl States, 1970-1990

Five Nahuatl States	*1970-1980 Net Population Loss*	*1980-1990 Net Population Loss*
Guerrero	-307,187	-392,755
Puebla	-376,961	-384,812
San Luis Potosí	-352,689	-353,675
Hidaldo	-348,222	-336,285
Veracruz	-128,815	-290,155
Total	-1,513,874	-1,757,682

Mexican census data (INEGI, 1994a: 50-51).

Similarly, there may be something of a denial of ethnic identity in order to achieve a degree of social mobility, Conklin and Lourie's (1983) terms. My consultants reported to me that there is great social stigma attached to speaking Nahuatl in the city, where many monolingual Spanish-speakers live. Like some speakers of other indigenous

languages, many Nahuatl speakers are ashamed to speak their language in public places. indigenous languages are commonly referred to as *dialectos* in Mexico, regarded as inferior to Spanish and other European languages. To this day, speaking Spanish is called "speaking Christian," in contrast to the heathen tongues used in indigenous cultures, leaving these indigenous languages with an additional religious stigma in their predominantly Catholic country. In addition, my consultants had no idea that Montezuma and the Aztec Empire spoke their very same language, although they knew all about the great Spanish Conquest and the triumph of Christian religion. In order to survive in the culture of the Conquistadores, the *indígenas* have assimilated and come to see an advantage to hiding their indigenous roots in many social and economic circumstances, and intensive propaganda in the schools and mass media have robbed them of the linguistic and ethnic heritage of their communities.

There is also a great lack of mother tongue institutions for Nahuatl speakers, another factor which Conklin and Lourie (1983) claim may lead to language loss.

In towns I visited, it was scarely known among the indigenous people that their language could be written, just like Spanish. There are virtually no texts available in Nahuatl, apart from the translations of the New Testament provided by missionaries of the Summer Institute of Linguistics.

Partly due to recent peasant uprising--in Chiapas, in particular--the Mexican government has begun to implement a bilingual education program reputedly aimed at maintaining indigenous languages while introducing the Spanish language among monolinguals. Even in places where the hereditary language has died out, the new

maintenance programs advocate teaching it along with Spanish in public schools. For this purpose, the government has begun to produce textbooks for indigenous children in their native languages, generally focused on the development of basic skills in reading and writing.

These programs may indeed help the indigenous languages to regain some of their social prestige and hence also promote the continued use of the languages in communities where they have nearly died out. However, unless Nahuatl speakers are allowed to regain productive control of their lives within their communities and schools, revitalizing Nahuatl cultural expression and invention, these efforts may have little effect.[54]

In addition, there are severe implementation problems in the bilingual programs, involving typical political corruption; children in remote areas often do not receive supplies, and teachers from more urbanized towns have been threatened and intimidated by local political bosses. In one case, a bilingual teacher was said to have been murdered for criticizing local party officials involved in embezzlement in the town to which he had been sent to teach.[55] Such obstacles further weaken any hope for language revitalization in schools.

Thus, while Figure 7 may reflect nothing more than the natural consequences of population growth, Figure 6 suggests a transition to Spanish-Nahuatl bilingualism,

[54]Regarding the Mexican bilingual education programs, see SEP (1994); for an assessment of the promise of the new bilingual programs to create additive bilingualism, see MacSwan and MacSwan (1997).

[55]The difficult political situation and the severe problems of implementation were reported to me in interviews with anonymous officials in the Mexican Secretaría de Educación Pública (SEP); the case of a murdered teacher was reported by a student who lives and works in the region. See MacSwan and MacSwan (1997) for discussion.

perhaps from there to Spanish monolingualism; this transition, it might be conjectured, is largely due to social and economic pressures imposed (or arranged) by Spanish-speaking corporate and industrial elites in urban centers.

Nonetheless, at the moment both languages are alive and well in towns I visited in the Tehuacán Valley. In particular, San Sebastián Zinacatepec, the town in which data in chapter 4 was collected, is about 65% bilingual in Spanish and Nahuatl and about 31.5% monolingual in Spanish, with 68% of the population able to read and write in Spanish. Only 271 residents (about 3.9%) are monolingual in Nahuatl. (See INEGI (1994b).) More specific detail regarding the town and my consultants will be provided in chapter 3.

Finally, a note on the term *Nahuatl* may be appropriate here. As mentioned, there are three words commonly used to refer to the language--Nahuatl, Mexicano, and Aztec. According to Hill and Hill (1986), the term *Nahuatl* did not begin to appear widely in scholarly work until the end of the nineteenth century. Among the Aztec people of Mexico, *Mexicano* is much more widely used, although *Nahuatl* is also known in some communities. In some indigenous communities, the preference for the term *Mexicano* is often reported to be related to a desire to sustain the memory that Mexico belonged to the speakers of Mexicano, or that they are the aboriginal Mexicans. *Nahuatlaca* was the name given to the tribes which came from Aztlan and settled in Central Mexico; it means "that which sounds good" (Siméon, 1995 [1885]).

Hill and Hill (1986) prefer the term *Mexicano* because their consultants did not know what Nahuatl was. My consultants, in contrast, referred to their language variously as Mexicano and Nahuatl. I have adopted the term *Nahuatl* in the present work because

it is one of the words used by my language consultants, and also happens to be the one more widely used in the scholarly literature.