

3. RESEARCH DESIGN

Below I present the research questions which guided my study; I also present information regarding my consultants, procedures for data collection, and methodology.

3.1 *Research Questions*

My central question is

- (1) What linguistic principles define code switching boundaries within sentences?

This question leads to some subquestions, shown in (2), which I shall also address using the Spanish-Nahuatl data I collected in Southeast Puebla. I will use the data pertaining to these subquestions to address question (1).

- (2a) What are the descriptive facts of Spanish-Nahuatl code switching in Southeast Puebla?
- (2b) How do the Spanish-Nahuatl data differ from those of other code switching corpora? What might explain conflicting findings?
- (2c) Does the Functional Head Constraint of Belazi, Rubin and Toribio (1994) account for the Spanish-Nahuatl code switching data?
- (2d) Does the “Null Theory” of code switching proposed in Mahootian (1993) account for the Spanish-Nahuatl code switching data?
- (2e) Does the MLF Model of Myers-Scotton (1993b) account for the Spanish-Nahuatl data?
- (2f) Can minimalist grammars explain the Spanish-Nahuatl data? If so, what are their advantages?
- (2g) Can a minimalist account be extended to data from other code switching corpora?

3.2 *Consultants*

Consultants used in a study on code switching ought to be selected according to particular criteria. Below these criteria and the extent to which my population conforms to them will be discussed.

3.2.1 Selection Criteria for Target Language Population

As discussed in section 2.1, consultants used in any study of intrasentential code switching should be native bilinguals, relatively evenly dominant in both languages, have actively used both languages since infancy, have had continued, sustained exposure to both languages, and appear to have generally high verbal fluency in both languages. This selection criterion is a methodological precaution to guard against critical period effects in the data.

Furthermore, Valdés (1981) and Lipksi (1978) have emphasized that code switching data may only reliably be gathered from communities in which code switching is a socially valued speech style. Otherwise, there is a risk that speakers may be reluctant to code switch, or may simply not engage in code switching at all, despite their bilingual ability.

3.2.2 Description of Consultants

During my fieldwork in Mexico I worked primarily with four consultants. Their names (pseudonyms) and the number of hours spent with each were

Jesús	17.0
Mario	10.3
Alberto	35.0
Jorge	38.5

Jesús, now 36 years old, grew up in the town of San Juan Tetelcingo in the Alto Balsas region of the state of Guerrero. He spoke Nahuatl exclusively until the age of seven; he then began to learn Spanish in elementary school. He graduated from a business administration college in Mexico City and now teaches a Nahuatl class on Saturday afternoons while attending graduate school.

I quickly learned from working with Jesús that he had very negative attitudes toward code switching. When I asked him to express judgments on particular code-switched sentences he reacted with great discomfort. He believed that his language was losing ground among its people and the mixing of Spanish and Nahuatl was a great political disservice to the Aztec community. Although Jesús appeared to be highly proficient in both Spanish and Nahuatl, his strongly negative attitudes toward code switching made him an inappropriate consultant for this study. Therefore, none of the code switching data collected from Jesús is presented in this dissertation.

San Sebastián Zinacatepec, the town in which I conducted nearly all of my fieldwork, is located near Tehuacán in southeast Puebla and has a population of about 7,000, according to the 1990 Mexican census. About 65% of the population is bilingual in Spanish and Nahuatl, about 31.5% speak only Spanish, and 271 inhabitants (about 3.9%) reported knowledge only of the indigenous language; 68% of the population can read and write in Spanish. Zinacatepec has one secondary school and four primary

schools. (See relevant tables in INEGI (1994b).) In the last local election, under tight police supervision, the Partido Ecológica ousted the ruling PRI.

Mario, another consultant, is a 26-year-old male from Zinacatepec. His mother and father are both monolingual speakers of Nahuatl, and his siblings do not know Spanish as well as he does, according to his assessment. Like Jesús, Mario began learning Spanish in school at age seven. He speaks both Spanish and Nahuatl on a daily basis, and he has been training to become a bilingual teacher in a school located in the mountains surrounding the Tehuacán Valley. The school is about a day's journey from San Sebastián, including four hours of hiking. Mario finished *preparatoria* (high school) and attends the technical university on Saturdays as part of his bilingual teacher preparation program.

Mario, with whom I worked about ten hours in total, was comfortable with producing code switching judgments but is not accustomed to speaking in this way. Because his parents are monolingual in Nahuatl and his siblings do not know Spanish well, Mario says that he is not often in a conversational situation with bilinguals which is informal enough to allow code switching. Nonetheless, Mario's judgments were consistent with those of Alberto and Jorge, both bilingual since infancy and very comfortable with code switching.

Alberto is 26 years old and has three sisters and two brothers. He grew up in San Sebastián speaking Nahuatl with his grandparents, who were essentially monolingual and lived with his family as he was growing up; he mainly spoke Spanish with his siblings and parents. His parents, who are both fluent bilinguals, speak both languages between

themselves. He is married and speaks Nahuatl with his wife, a monolingual. He has two children, four and two years old, who are learning to speak Nahuatl at home. Alberto works in the country where he grows crops on a small parcel of land he inherited from his father. Out in the fields Spanish and Nahuatl are both used daily.

Alberto is very comfortable with code switching. He explained that in the company of close friends and peers it is the typical way of talking, but it is looked down upon by monolingual Spanish-speakers in the big city and sometimes by older Nahuatl-speakers in San Sebastián. I usually worked with Alberto and Jorge together at the same time.

Jorge and Alberto are very comfortable together because they have known each other all their lives and are now family (Alberto is married to Jorge's sister). Jorge is 24 years old and is the youngest of six children. He and all his siblings are native in both Spanish and Nahuatl, and have spoken both languages among themselves since infancy. However, his grandparents, who also lived with them, spoke Nahuatl and very little Spanish. His parents, on the other hand, speak both languages but are dominant in Nahuatl. Jorge is single and lives at home; he speaks only Nahuatl with his mother but both languages with his siblings. Like Alberto, he works in the fields where he grows produce to sell at market. His parcel of land is now very small, but he can still make a living with it.

Jorge, like Alberto, reports that he often code switches with close friends and family members. Language mixture came very naturally to him, and he often offered his own theories about why some switches were not allowed. For the experimental data, I

relied most heavily upon Jorge and Alberto because of their strong bilingual ability and their high level of comfort with code switching. Although Mario was also consulted, most of the data reported in section 4.1 are due to Jorge and Alberto.

The naturalistic data were gathered at an elementary school in San Sebastián Zinacatepec. Although there are bilingual schools in the Tehuacán Valley (in San José Hiahuatlan and Altepexi), the school in San Sebastián did not offer first-language support for Nahuatl-speaking children. The fifth grade teacher, herself a Spanish-Nahuatl bilingual, invited children who self-identified as bilingual in Spanish and Nahuatl to participate in the project.

A proficient Spanish-Nahuatl bilingual adult (Mario, described above) came along as an assistant in the project and interviewed five children for about twenty-five minutes. The children spoke to one another and to the adult bilingual regarding a variety of topics, freely mixing Spanish and Nahuatl in conversation. These interactions were videotaped and later transcribed with the help of the adult bilingual. All but one of the children appeared to be conversationally fluent in Nahuatl; the other child understood much of what was said to him but could not answer in Nahuatl. The code switching facts observed in this session, among both the children and the adult bilingual assistant, are reported in section 4.2.

3.3 Data Collection Procedures

As already mentioned, I used two procedures for data collection, naturalistic observation and sentence judgment (SJ) tasks. Although some researchers in code switching have a strong preference for one of these data types over the other, both used in

combination may ultimately be a useful way of establishing the best description of code switching for a particular language pair.

3.3.1 Naturalistic Observations

Labov (1971, 1972) favored naturalistic data over elicited data because, according to Newmeyer (1983), he did not accept the distinction between competence and performance. While naturalistic data may be useful for obtaining initial findings in a natural setting, it does not tell us what *cannot* occur; thus, it is of somewhat limited use in constructing a theory of the sort pursued here. However, the naturalistic findings may provide evidence which falsifies some of the proposals discussed in 2.2.2 (and revisited in section 5.2.1) and corroborate the experimental findings.

Naturalistic data were transcribed with the assistance of a native bilingual and later coded for morphological, syntactic, and lexical information relevant to code switching. Roughly following MacWhinney's (1991, 1995) CHAT format, and influenced by Curtiss, Schaeffer, Sano, MacSwan, and Masilon (1996), each code switch was marked for its syntactic category on a separate code switching tier (designated %cdg:),. as illustrated in (3).

(3)	*PED:	¿Tlamón tikchiwa in recreo?
	%trn:	¿Qué haces en la hora del recreo?
	%cdg:	N recreo

Here Pedro (whose speaker tier is marked *PED:, following CHAT conventions) asks a Nahuatl question during conversation. A Spanish rendition is presented in a translation

tier (%trn:),⁵⁶ followed by observations regarding the utterance in the code switching tier (%cdg:). In (3), the observation is that a code switch occurs at the Spanish noun *recreo*.

Coding the transcript in this way allowed for the document to be automatically searched using a simple frequency counter on the %cdg: tier. These occurrences may then be manually inspected to determine what other properties may be relevant to the code switch (major phrase boundaries, the presence of functional categories, and so on).

In addition, borrowed elements are coded on the speaker tier using the suffix @B, as in (4).

- (4) *HEC: ¿Pero@B tlan okichiwato?
 %trn: ¿Pero qué fue a hacer?
 %cdg:

Because *pero* has been borrowed into Nahuatl (see section 2.3.1), its presence is not marked on the %cdg: tier. However, it may be of some peripheral interest to keep track of the frequency of borrowed words, so these are coded with the suffix @B, as shown in (4), and are reported in section 4.2.13.

3.3.2 Sentence Judgment Tasks

Before asking consultants for judgments on code-switched sentences, I discussed the general and widespread practice of code switching, presenting the case of Spanish-English code switching in Los Angeles as an example, an especially useful illustration

⁵⁶It was much more convenient to use Spanish translations, provided by my Spanish-Nahuatl bilingual assistant, rather than English. In (3), the question *¿Qué haces en la hora del recreo?* means “What do you do at recess time?”

given that Spanish is the language of prestige in this bilingual community, unlike the situation in Los Angeles.

Then, in an effort to make very clear what sort of information I was looking for, I presented code-switched constructions which I suspected would be extremely bad (based on comparisons with English-Spanish mixtures; see (6c) in Table 1, page 68). These involved the mixture of a Spanish accusative clitic with a Nahuatl verb, already marked for object agreement, as shown in (5).

- (5) *Yo la *niktlasojtla*
 Yo la ni-k-tlasojtla
 I ACC.CLITIC 1S-3Ss-love
 ‘I love her’

This construction, which my consultants judged to be extremely ill-formed, served as a comparison for relatively better mixtures. Consultants were always asked to provide gradient rankings for sentences; (5) was continually referenced as a prototypical worst case. In addition, each sentence was written, read aloud, and sometimes discussed at some length by consultants.

The constructions I presented were for the most part modeled after those summarized in Table 1 (page 68). In addition, my choice of constructions was sometimes influenced by Baker’s (1996) Mohawk corpus.

3.3.3 Conventions and Abbreviations Used for Presentation of Data

Finally, I will present data in this dissertation using the following format:

- (10) Le dije *ke kitlasojtla in Juan sikpanoah*
 le dije ke 0-ki-tlasojtla that in Juan sikpanoah
 DAT.CLITIC PAST/1Ss/say 3S-3Os-love IN Juan a.lot
 ‘I told him that she loves Juan a lot’

The first line of (10) is the datum; the second line is a morphological “parse” of the datum; the third line is a gloss, morpheme by morpheme, followed in line four by an approximate translation into English (in single quotes).

Utterances prefixed with a star or asterisk (*) are those regarded as ill-formed by consultants; expressions regarded as degraded but possibly acceptable are prefixed with (multiple) question marks (?, ??, ???), the more the worse. Material in parentheses may be omitted without altering the judgments denoted by *, ?, ?? or ???. Sentences with no *- or ?-prefix are well-formed.

A code switch is indicated with *italic text*, as is standard in the literature, but here the language italicized will not be presumed to have a special status (that is, it will not, for my purposes, play the role of the “embedded” language as opposed to the “matrix” language, a distinction important in some models).

In the gloss line, bound morphemes are separated by a hyphen (-), free morphemes by a space. Portmanteau morphemes are glossed using slashes (/), as illustrated in (10) for *dije*, an irregular Spanish verb, glossed as *PAST/ISs/say* (order irrelevant). The meanings of many functional/inflectional morphemes are glossed in SMALL UPPERCASE, as is conventional in linguistics discussions. When a single morpheme requires two or more separate words for its gloss, a period (.) is used to separate these words; this is illustrated in (10) for *le*, glossed as *DAT.CLITIC* (dative clitic, where *DAT* abbreviates *dative*), and for *sikpanoah*, glossed as *a.lot*. A zero (0) introduced in the second line, as shown in (10) where *kitlasojtla* is parsed as *0-ki-tlasojtla*, indicates the presence of a phonetically null element to be glossed in line three.

The particular terms used for functional elements were selected in the interest of expository convenience. Nowhere should these choices be taken to represent a theoretical commitment; for instance, by labeling Nahuatl *ki-* as 3O (third person object agreement), I do not mean to preclude an analysis of this item as an accusative clitic (an ACC.CLITIC) in chapter 5, or even to claim that the distinction is an important one.

Abbreviations used in the glosses are given in Table 6.

Table 6: Abbreviations of Terms Used in Glosses (Alphabetical)

<i>Abbreviation</i>	<i>Morpheme</i>
1O	first person object agreement (unspecified for number)
1Op	first person plural object agreement
1Os	first person singular object agreement
1S	first person subject agreement (unspecified for number)
1Sp	first person plural subject agreement
1Ss	first person singular subject agreement
2O	second person object agreement (unspecified for number)
2Op	second person plural object agreement
2Os	second person singular object agreement
2S	second person subject agreement (unspecified for number)
2Sp	second person plural subject agreement
2Ss	second person singular subject agreement
3O	third person object agreement (unspecified for number)
3Op	third person plural object agreement
3Os	third person singular object agreement
3S	third person subject agreement (unspecified for number)
3Sp	third person plural subject agreement

Table 6 (continued): Abbreviations of Morphemes Used in Glosses

<i>Abbreviation</i>	<i>Morpheme</i>
3Ss	third person singular subject agreement
ACC.CLITIC	accusative clitic
COND	conditional
DAT.CLITIC	dative clitic
DUR	durative morpheme (like Spanish <i>-ando</i> or Nahuatl <i>-toc</i>)
FUT	future tense
IMP	imperative
IN	Nahuatl determiner <i>in</i> , similar to English <i>the</i> or <i>a</i>
INDEF	indefinite morpheme (Nahuatl <i>tla-</i>)
INF	infinitive marker
LOC	locative suffix
NOM.CLITIC	nominative clitic
NSF	noun suffix (sometimes called absolutive)
PAST	past tense
PL	plural marking (on nouns or verbs)
PRES	present tense
PRT	particle
REF	reflexive clitic or pronoun
SING	singular
SUBJ	subjunctive mood
VSF	verb suffix
1SPOS	first person singular possessive
2SPOS	second person singular possessive
3SPOS	third person singular possessive
1PPOS	first person plural possessive
2PPOS	second person plural possessive
3PPOS	third person plural possessive

Finally, I should say a word about the orthographic system used in this dissertation. Some Nahuatl communities, influenced by the Spanish writing system, use *j* for /h/, while others used *h*. In some contexts, the /h/ is deleted. My consultants, for instance, spelled the verbal suffix *-kej* variously as *-kej*, *-keh* and *ke*. In addition, the Nahuatl determiner element *in* is sometimes cliticized to an element which follows it,

pronounced as a vocalic alveolar nasal; in this case it might be written simply as *n*.⁵⁷ The consultants for my experimental data consistently used the more traditional *in* spelling, while the consultants who helped with the naturalistic data preferred *n*. As a result, the orthographic system used in this text is a mixture of different Nahuatl orthographic systems. Since there is no established conventional writing system for this language, and since the “inconsistencies in spelling” originate with the Nahuatl speaking community itself, these differences will not be of concern to me in the following chapters. I have essentially used whatever spelling my consultants wished. (See 2.5.3.3 for a brief word on Nahuatl orthographic traditions.)

3.4 How the Research Questions Will be Addressed

Returning briefly to the research questions set out in section 3.1 of this chapter, some comments should be made regarding how the particular questions will be addressed in the findings and analysis chapters of this dissertation.

Question (1) will be informed by answers to its subquestions in (2). The descriptive question, (2a), will be answered as the facts are presented in chapter 4. The descriptive characteristic of question (2b) will also be answered in chapter 4, and an explanation for apparent conflicts in findings will be offered in section 5.3.

Questions (2c) through (2e), each of which pertains to a prior account of code switching, will be specifically addressed in section 5.2.1. Finally, in sections 5.2 and 5.3,

⁵⁷R. Joe Campbell (personal communication) suggests that the cliticization takes place only before a vowel for purposes of syllabification.

minimalist grammars will be used to explain both the Spanish-Nahuatl facts and some of those reported in other corpora, as addressed in questions (2f) and (2g).