Chapter 2

Methodology

This chapter explains the methodology used for data collection and coding. A detailed explanation of stimuli is included along with descriptions of the project participants and the procedure. Twenty-four verbal stimuli in sentential contexts were randomly distributed among distractors and presented orally to sixty-nine bilingual participants for a translation task. A brief questionnaire asking each participant to describe aspects of his or her bilingualism was applied following the elicitation process. Based on age and contact scores derived from the questionnaire, participants were assigned to groups for the analysis. A description of the piloting process can be found in the second half of the chapter.

2.1 Data collection methodology

2.1.1 Participants

Sixty-nine bilingual speakers of Veneto and Spanish participated in the study: 35 older speakers ranging in age from 58 to 85 and 34 younger speakers ranging in age from nine to 20. All participants were approached by the investigator by taking advantage of the dense social networks in the bilingual community (see Milroy, 1980, for details on the use of social networks for data elicitation). Almost every participant who was interviewed provided the researcher with possibilities for future interviews.

Age of potential interviewees was the major factor in selecting participants, as the researcher was interested in finding either older or younger speakers of Veneto.

2.1.2 Advisors

Two bilingual advisors participated in the design of the current study and the elaboration of stimuli. The first advisor, an undergraduate student of Language and Literature in his final year, is a member of the Chipileño community trained in linguistic theory but is a non-native speaker of Veneto. The second advisor lacks training as a linguist but is both a member of the community and a native speaker of Veneto. Both have carried out ethnographic observations of the use of Veneto in the community as part of a continuing personal project. These observations, which served as an initial foundation for this study, include observations of lexical changes in Veneto which range from the borrowing of cultural and core Spanish terms (Myers-Scotton, 1993,) as well as changes in syntactic frame of verbs.

2.1.3 Materials

2.1.3.1 Questionnaire

A short questionnaire was adapted from Hall and Smith (unpublished) and was translated to Spanish by the researcher. This questionnaire was further adapted for the younger and older interviewees. The purpose of this questionnaire was to collect sociolinguistic information about each informant which would allow distribution into groups. Questions uncovered two types of information about each speaker: 1) the level

of bilingualism and 2) the degree of contact with Spanish. A copy of both questionnaires can be found in Appendix 1.

2.1.3.2 Stimuli

A list of verbs that differ from their Spanish translation equivalents in syntactic frame and were suspected to be undergoing syntactic frame changes was constructed with the help of the bilingual advisors. This was done primarily by using the Veneto-Spanish dictionary elaborated by MacKay (2002) which contains partial information about the syntactic frame of Veneto words (for nouns: gender; for verbs: reflexivity; use of prepositional complements for phrasal verbs only). Information obtained from this source was supplemented and/or tentatively confirmed by ethnographic observations made by advisors.

The dictionary was estimated to contain approximately 1700 verbs, based on the average number of verbs found on several randomly selected pages multiplied by number of pages in the dictionary. However, not all 1700 verbs would be appropriate for study. The first step of the process of selecting specific verbs consisted of an evaluation of culturally-bound verbs. Chipilo began as an agricultural town, and although it remains primarily so today, antiquated and terminologically specific verbs that would most likely be known only by the older speakers and not by the younger ones were not included. Although the loss of these words may be interesting from anthropological or ethnographic standpoints, this is beyond the reach of this project.

The Spanish equivalents of the verbs that remained were studied closely in order to determine the syntactic frame. The focus of this step was to find Spanish verbs that

either differed in prepositional complement or differed in reflexivity from the Veneto word. These two frame aspects were chosen because it was easier to work with the non-linguist advisor if the syntactic frame was more easily intuited and transparent, especially since the input from the native speaker was relied on more during this stage of stimuli selection than that of the Spanish-dominant linguistically-trained advisor.[1]

The process of narrowing down verbs yielded a list of approximately 400 verbs in Spanish. Working with the primary advisor, the meanings and syntactic frames of all 400 words were carefully studied. Nuances of the Veneto and Spanish translation equivalents were studied for two reasons. Firstly, a bilingual dictionary gives a definition of a word via its translation equivalent, but often it is only through use that the underlying concept of each word is understood. This is especially true for abstract words, which verbs tend to be. Secondly, it was important to ensure that the use of even a contextualized Spanish verb could not elicit a large number of corresponding Veneto verbs, since the goal was to elicit and study specific Veneto verbs. Many more possible verbs were eliminated based on these criteria. Syntactic frames of the remaining Veneto verbs were obtained and contrasted with the syntactic frames of the Spanish equivalent. One final elimination was made based wholly on the intuition of one advisor (see 4.5 for a discussion) that phrasal verbs such as *caier do* (Sp. 'caerse', Eng. 'fall down'), though they contrasted with the Spanish equivalent's syntactic frame, would not yield strong CLI results[2]. By identifying verbs whose syntactic frames differed from Spanish, a list of 24 verbs that could potentially be affected by inter-lexical

[1] The reasons that the Veneto-dominant advisor was used almost exclusively in this process were to receive a) native-speaker judgements on grammar, b) native-speaker judgements on lexical meaning, and c) input on actual use of verbs in the community.

[2] Cross-lexical influence at the frame level

influence was elaborated. Table 1 contains the list of verbs that were studied. The list contains the Spanish translation equivalent form and frame, followed by the Veneto form and traditional frame. Based on the contrast between the Spanish and Veneto frames, a hypothesized Veneto frame was hypothesized. This hypothesized frame was coded for the change that could logically take place. Accordingly, there are six types of changes (See table 1).

AP	Added Preposition
DP	Different Preposition
NR	No Reflexivity
AR	Added Reflexivity
DN	Different Preposition, No Reflexivity
DA	Different Preposition, Added Reflexivity

Table 1 Coding key

	Spanish stimuli		Veneto			
	form	frame	form	Traditional frame	Hypothesized innovated frame	code
1	Aprender	V, [__ a]	inparar	V, [__]	V, [__ a]	AP
2	Empezar	V, [__ a]	scominzar	V, [__]	V, [__ a]	AP
3	Hablar	V, [__ en]	parlar	V, [__]	V, [__ in/inte/ente]	AP
4	Insistir	V, [__ en]	insister	V, [__]	V, [__ in/inte/ente]	AP
5	Fiar	V, [__ de]	infidar	V, [__ in]	V, [__ de/da]	DP
6	Jugar	V, [__ a]	dugar	V, [__]	V, [__ a]	DP
7	Oler	V, [__ a]	nasar	V, [__ da]	V, [__ a]	DP
8	Preguntar	V, [__ por]	domandar	V, [__ de]	V, [__ par]	DP
9	Saber	V, [__ a]	saber	V, [__ da]	V, [__ a]	DP
10	Tropezar	V, [__ con]	ingambarar	V, [__ par]	V, [__ co]	DP
11	Recargar	V, [__ en]; V, [+refl.]	puyar	V, [__ su par]; V, [-refl.]	V, [__ (su) in/inte/ente]	DA
12	Atrever	V, [__ a]; V, [+refl.]	osar	V, [__]; V, [-refl.]	V, [__ a]; V, [+refl.]	DA
13	Cansar	V, [__ de]; V, [+refl.]	stracar	V, [__]; V, [-refl.]	V, [__ de/da]; V, [+refl.]	DA
14	Encontrar	V, [__ con]; V, [+refl.]	catar	V, [__ para]; V, [-refl.]	V, [__ co]; V, [+refl.]	DA
15	Fijar	V, [__ en]; V, [+refl.]	infisar	V, [__ de]; V [-refl.]	V, [__ in/inte/ente]; V, [+refl.]	DA
16	Pelear	V, [__ con]; V, [+refl.]	brancar	V, [__ par]	V, [__ co]; V, [+refl.]	DA
17	Soñar	V, [__ con]; V, [-refl.]	insuniar	V, [__ de]; V, [+refl.]	V, [__ co]; V, [-refl.]	DN
18	Caber	V, [-refl.]	logar	V, [+refl.]	V, [-refl.]	NR
19	Descansar	V, [-refl.]	destracar	V, [+refl.]	V, [-refl.]	NR
20	Necesitar	V, [__]; V, [-refl.]	ocorer	V, [__]; V, [+refl.]	V, [__]; V, [-refl.]	NR
21	Irse	V, [+refl.]	ndar	V, [__]	V, [+refl.]	AR
22	Levantar	V, [+refl.]	levar	V, [__ su]; V, [-refl.]	V, [__]; V, [+refl.]	AR
23	Parecer	V, [+refl.]	someiar	V, [-refl.]	V, [+refl.]	AR
24	Reir	V, [__ de]; V, [+refl.]	rider	V, [__ de]; V, [-refl.]	V, [__ de/da]; V, [+refl.]	AR

Table 2 Spanish-Veneto stimuli list with frames

For each verb in the list, a sentence with the Spanish translation equivalent was designed. These sentences included not only the verb in question, but also its preposition(s) or a complement, according to the syntactic frame in question. These sentences were written collaboratively by the researcher and the bilingual advisors. Suggestions for possible sentences as well as modifications were based on 1) providing sufficient context to the sentences to minimize ambiguity in meaning, concentrating especially on providing appropriate context for the verb, 2) considering appropriate social use of the language and, 3) the articulation of the verb and its complement. Sentence validity was confirmed through back-translation and some sentences underwent further modification after pilot results. These sentences had an average of 5.2 words. In addition, an equivalent number of distractor sentences was elaborated. These sentences contained an average of 6.3 words and often included verbs that were of interest to the advisors but was not in the scope of this present study. The purpose of including these distractors was to ensure that informants would be unaware of the specific objective and focus of the study. Stimuli were then randomized along with the distractors. Minor changes in ordering were made to avoid over-grouping of verbs with similar frames and stimuli sentences. A copy of this list of sentences can be found in Appendix 2.

Finally, taking into consideration the mental fatigue involved in doing oral translation tasks, especially for those who have not received formal training, it was necessary to ensure that any one verb would have an equal opportunity to be found at the beginning, middle, or end of the task. Therefore, the 48 randomized sentences

were divided into four groups. The order of each group of 12 sentences changed as the data collection proceeded.

2.2 Procedure

Access to informants was gained principally through being introduced into the community as a "friend of a friend" (Milroy, 1980). As Chipilo is a tight-knit community, sessions with each informant provided opportunities to meet new potential informants. All sessions were therefore conducted in a fairly impromptu fashion, since pre-selection of participants was not deemed necessary.

The researcher insisted that all sessions be carried out in a one-on-one fashion. Since these were done in people's homes or places of work, sessions were often briefly interrupted by family members or visitors. During interruptions, the session was momentarily suspended.

The first part of each session involved filling out the language history questionnaire. This was usually done by the researcher, especially during sessions with older interviewees or the very young, who were unaccustomed to filling out questionnaires.

The second part of each session involved administering the oral stimuli. Informants were instructed that they would hear sentences in Spanish which they had to translate to Veneto. As this part of the interview was being audio recorded, informants were also instructed to speak into the microphone. The translation activity typically lasted between five and ten minutes. The time needed to fill out the questionnaire and do the translation activity lasted no longer that half an hour.

2.2.1 Data coding procedure

Recordings were studied as data collection proceeded. Using the stimuli list in Table 1 as support, a spreadsheet was made noting for each participant the form of the verb used, the frame of the verb, and observations about the form and frame. These observations consisted of categorizing each elicited verb translation as borrowed from Spanish or as having a classical Veneto or innovated frame

2.2.2 Questionnaire coding procedure

As mentioned earlier, two different questionnaires were administered according to the age of the participant. The purpose of the questionnaires was to have an objective way of determining the degree of contact each participant has with Spanish. Since the questionnaire included questions about the individual's language acquisition and linguistic abilities in addition to domains, not all answers to the questionnaire were used to calculate the contact score. Answers to individual questions about domains were scored as +1 when there was a tendency to use more Veneto and +2 when the tendency was to use more Spanish. Although participants were asked to make clear choices between these options, some insisted that they used as much Spanish as Veneto for some domains; in these instances a score of +1.5 was given. Answers to questions about frequency of language use were scored on a scale of 1 to 4, with 1 indicating a higher frequency use of Veneto and 4 indicating a higher frequency use of Spanish. Raw scores were divided by the highest possible score and converted to a contact score.

Questions about language use at work were reserved for older speakers, although some of the older adolescents were employed part time. Adults fell into three basic categories regarding work: 1) housewives, having never worked outside of the home or stable; 2) retired workers, having worked previously in Chipilo, Cholula, or elsewhere; and 3) currently employed. Raw scores were again divided by the highest possible score and converted to a contact score, but due to work status, the adults' scoring procedure was slightly different from that of the adolescents. Housewives had no additional points added to their score. Retirees or people who were still employed had additional points added to their raw score based on whether they were employed in Chipilo, whether Veneto was the language used at work, and the amount of time spent working outside the home. These raw score were also converted to contact scores. For an example of how scores were calculated, see appendix 3.

2.2.3 Participant assignment procedure

Participants were assigned twice to groups; the first distribution based solely on age, allowing the researcher to compare all 69 speakers. A second distribution of participants was based on the contact score from the language history questionnaire, providing the researcher with a more objective standard for determining degree of contact with Spanish. Originally, it was hoped that not only would there be two clear groups according to age, but also that the people in both age groups could be easily assigned to the subgroups of +/- contact. However, it was found that for both older and younger speakers there was a continuum between less contact and more contact, meaning that there was no clear cut-off point in the middle to divide either group.

Instead, the polar extremes for both groups were used for comparison. The cut-off points were arbitrary in a sense, but effort was made to consider clusters of scores so that individuals with almost identical scores were grouped together instead of being separated.

In the case of the younger speakers, the contact scores ranged between .407 to .944 (out of a possible range of .370 to 1.000), whereas among older speakers the scores ranged from .384 to .725 (out of a possible range of .357 to 1.000). Eight younger speakers with scores from .407 to .463 were compared to eight younger speakers with scores from .610 to .944; from the older group, seven speakers with scores ranging from .384 to .393 were compared to nine speakers whose scores ranged from.496 to .725. The distribution of these informants is given in Table 2.

	Less contact with Spanish	More contact with Spanish
Younger than 20	8	8
Older than 54	7	9

Table 2 Distribution of bilingual informants according to contact

2.3 Pilot study

The pilot was carried out to resolve questions about the methodology that would be used. Initially, two tasks were proposed: an oral translation task and an oral elicitation task, similar to an oral elicitation task (OET). Later, a written task was also considered. The purpose of piloting these three instruments was to ascertain which

would yield the best elicitation rate of the target Veneto verb and its complement, and to make any necessary modifications to the procedure.

2.3.1 Pilot study participants

Five participants, distributed across three generations of the same family, were chosen for the pilot. The participants represented different age groups as well as different educational levels and contact with Spanish. The general breakdown of the five is as follows:

Age	Less contact with Spanish	More contact with Spanish
20's	1	1
40's	1	0
60'2	2	0

Table 3 Distribution of bilingual informants for pilot

2.3.2 Pilot study materials

Three data collection techniques were elaborated: translations, oral elicitation tasks, and a written task. All instruments used as a guide the list of Veneto verbs that are hypothesized to be undergoing innovations.

First, a list of 31 sentences was designed with the Spanish translation equivalents of verbs that are hypothesized to be undergoing a shift to the Spanish syntactic frame. These sentences included not only the verb in question, but also its preposition(s) or a complement, according to the syntactic frame in question. Suggestions for possible sentences were made through collaboration between the researcher and one of the bilingual advisors. For the pilot, these sentences were randomly ordered, yet no distractor sentences were incorporated. (See appendix 4).

Second, an Oral Elicitation Task was elaborated. These tasks consisted of a list of 24 multiple-component situations with a linear logic between the beginning and middle, and an open end to be provided by the informant. These were written to reflect social situations endemic to Chipilo, translated to Veneto, and audiotaped with the voice of the second informant, a native speaker of Veneto. (See appendix 5).

The third instrument used written language instead of oral language. Words in the same thirty-one sentences used in the oral translation activity were randomized and the verb's prepositions and reflexive pronouns were removed. These prepositions and reflexive pronouns were included in a list at the head of the activity to be used as needed to complete the sentences while the participants wrote out these sentences. (See appendix 6).

2.3.3 Pilot study procedure

Piloting the above materials took place during two different sessions, one for the translation and oral elicitation tasks, and a second for the scrambling task, a month later. Not all informants participated in all tasks. Each session took place at the home of the participant and the only people present during the sessions were the participant and the investigator. However, since they were carried out in participants' homes, sessions were often briefly interrupted. During each interruption, the session was suspended momentarily.

The translation exercises lasted approximately five minutes and the oral elicitation task exercises lasted approximately 15 minutes. Both activities were audio taped. For the translation task, participants were told that they would hear a sentence

in Spanish and they should immediately give the translation equivalent in Veneto. For the OET part of the session, participants were told that they would hear a truncated situation in Veneto and that they would have to supply an appropriate ending.

The written task was carried out at a later date. Verb particles in 31 sentences were removed and the remaining words were scrambled by the researcher. These scrambled sentences were then presented to the participants. Participants were asked to unscramble the words, inserting words presented in a word list as necessary. Furthermore, as Veneto has no established written system, a hispanicized orthography was used in order to facilitate reading comprehension.

In order to gather data on language history, the researcher asked questions regarding parents, age, education level, and general contact with Spanish. The notes taken during this session served as the basis for the distribution of the participants as seen in Table 1.

2.3.4 Pilot study results and discussion

The translation task elicited responses for all of the sentences. As can be seen in table 4, the translation task elicited the target verb with a precision varying between 70% and 77% of the responses. Furthermore, for those responses that did not elicit the target verb, between 6 and 10% of the responses involved a borrowed Spanish verb. While the borrowing of Spanish verbs is not the central focus of this study, this still yields interesting data regarding contact between two languages.

	Participant 1	Participant 2	Participant 3
Target verb use	77%	71%	70%
Lexicalized Spanish word	10%	6%	7%

Table 4 Translation task results

The OETs also elicited responses for all situations. As can be seen in table 5, the elicitation rate varied between 39% and 57%, with the two younger participants using Veneto verb in question at higher rates. However, when the OET responses are analyzed for the frequency in which the verb along with its syntactic frame are elicited, the rate drops to between 9% and 17%.

Two observations can be made from this data. Firstly, the OET were not as successful as the translation task at eliciting the target verb. And secondly, the tasks was 18% less successful for the older speaker in comparison to the younger ones.

	Participant 1	Participant 2	Participant 3
Target verb use	57%	57%	39%
Target verb with frame	17%	19%	17%

Table 5 OET results for elicitation of target verb with frame

Finally, the written exercise showed the largest gap between older and younger speakers. While the 20- and 40-year-old speakers were able to do this task, the older speakers were not. The difference in abilities may be due to differences in exposure to

written Veneto as well as level of schooling, including literacy skills, leading therefore to differences in abilities to do school-like activities.

Based on the results of the pilot, the translation task was used for the interviews with only minor changes made, principally to ensure that all the prepositions would be articulated better. The OETs, on the other hand, were eliminated due to questions of feasibility. Regarding the written data collection method, as the older participants were unable to perform the task, this method was also eliminated.