CHAPTER 3

METHODOLOGY

3.1 Data Sources

There were two different data sources available for this study: digital recordings collected by this researcher in Laos and by Dr. Jerold Edmondson in Vietnam, and various phonological sketches of Kim Mun varieties spoken in China as presented by other researchers.[3]

The phonological analysis of this thesis is based on a 503-item (Laos) and 441-item (Vietnam) wordlist[4]. The data for the Kim Mun in Laos was collected and recorded in Luang Namtha by this researcher in 2006. Since the only data available for the Vietnam variety was already 80 years old[5], Dr. Jerold Edmondson kindly provided his digitized recordings and transcription notes of this variety taken in Lao Cai, Vietnam, in 1999. In order to provide maximum consistency, only the author's transcriptions of Edmondson's data were used in this study.

---

[3] The following sources also contain wordlists: Mao (2004) (China, Yunnan and Guangxi Province); Liu, et. al. (1998) (China, Yunnan Province); Shintani (1990) (China, Hainan Island)

[4] In Laos time for data collection was highly limited due to rice-harvesting season. The data from Vietnam was limited to what this researcher received from Dr. Edmondson.

[5] Purnell (1970) based on a dictionary by Savina (1927).

The data from the mainland of China, including Yunnan Province and Guangxi Province, came from data published in 2004 (Mao 2004). A further source was discovered through a Chinese resource on Minority Language and Alphabets (Liu, et. al. 1998). Liu, et. al. (1998) does not provide an analysis of Kim Mun, though he does provide a few tables that include initials, finals, and tones. Liu, et. al.'s largest contribution to Kim Mun studies is a 1500-item wordlist from a Kim Mun village named Jianlichuan in Wen Shan District, Funing County, Huajia Township in Yunnan Province. The data collected by Shintani (1990) on Hainan Island was gathered during 1987-88.

Since there are no texts available for this study, grammatically motivated differences could not be taken into consideration. Also, apart from some general background information about the people, this thesis will not include any socio-linguistic aspects because the data sources used for the phonological comparison do not provide this type of information.

3.2 Data Gathering

The 503-item wordlist used for data gathering in Laos was divided into different sections based on semantic domains (cf. Appendix A). The mediating languages used were Lao and Thai. Three male Kim Mun

speakers, approximate ages 23, 25, and 55 years old, first discussed which word in Kim Mun best fit the Lao word.[6] All three language consultants gave feedback for the transcriptions of the wordlist.[7] Because the two older Kim Mun speakers considered the articulation of the 23-year old speaker to be the best, only his articulation of the wordlist was recorded and used for further editing of the raw data.

The data from Dr. Edmondson came in the form of digitized recordings organized by semantic domains and Dr. Edmondson's transcription notes. Data from Dr. Edmondson's work generated a 441-item wordlist (cf. Appendix A).[8]

3.3 Data Analysis

The data were processed with the computer programs Speech Analyzer[9] and Praat[10]. After inputting the data, the recordings were broken down into individual words, tagged, transcribed, and glossed. Waveforms,

---

[6] It was not possible to find more language consultants because it was rice-harvesting season.

[7] One limitation of this study is that the analysis of Lao Kim Mun is based on only one wordlist.

[8] One limitation of this study is that the analysis of Vietnam Kim Mun is based on only one wordlist.

[9] Speech Analyzer is a speech analysis tool that provides pitch listings, sound waves, and spectrograms and has the ability to organize a wordlist for quick and easy reference. All the data are saved directly in the audio file, which makes the audio easy to access when referencing a word in the wordlist.

[10] Praat (version 4.4.31) is a phonetics analysis program created by Paul Boersma and David Weenick. Praat was originally created in 1992 and has undergone many version updates since. Praat provides detail pitch listings for tone analysis among many other features. Pitch listing from Praat were used in conjunction with Microsoft Excel to create tone charts for tonal analysis.

spectrograms, and pitch listings were consulted to verify the transcriptions. The program Cool Edit Pro 2000[11] was used to cut individual words from the word list for analysis. The wordlists obtained through Speech Analyzer were then exported into Phonology Assistant[12] for further analysis. General phonology following the functional approach (Burquest 2006) was employed for the analysis, description, and comparison of the data.

3.4 Data Transcription

All data from the Lao and Vietnam varieties were transcribed using the International Phonetic Alphabet (IPA), except for alveolo-palatal segments (cf. Section 4.3.2.2 for a discussion of the alveolo-palatal place of articulation and its application to the IPA system). Mao (2004) transcribed the data from Yunnan and Guangxi using the alveolo-palatal segments. The Lao and Vietnam data sets were also transcribed with the alveolo-palatal segments, following Mao.

---

[11] Cool Edit Pro 2000 is audio editing program for editing, converting, and fixing audio files, among other things.

[12] Phonology Assistant is a phonological analysis tool created by SIL International in 2006. It is an upgrade from Speech Manager and was in its beta stage at the time of this analysis.