A Dialectal and Phonological Analysis of Penghu Taiwanese

by

Alexander T. Ratte

Professor Nathan Sanders, Advisor

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Bachelor of Arts with Honors
in Linguistics

Williams College

Willliamstown, Massachusetts

May 2009

Table of Contents

Abstract

This study investigates voiced oral stops as they are expressed in the dialect of the Southern Min language spoken in Penghu County of Taiwan. Penghu Taiwanese exhibits various changes to the underlying voiced stops /b/ and /g/, and research was carried out in two ways to investigate these sound changes. First, the prevalence and distribution of these shifts were tracked across Penghu County in a dialectal study, and second, acoustic data was collected from Penghu residents in a phonetic study to measure the voice onset times and burst intensities of stops. In the dialectal analysis, I discovered that Magong City had a higher concentration of changes to /g/, and that these sound changes all appear to be part of a conspiracy against voiced stops in Taiwanese. In the phonetic analysis, acoustic data collected from speakers in Magong City showed not only common deletion of voiced stops, but also wholesale replacement of underlying /g/ with [k] and even forms of voiced consonants that had undergone mild lenition but not total deletion. These results not only confirm the sound changes tracked in the dialectal study, but also introduce new ways in which a phonological conspiracy may be occurring against voiced stops in Penghu. Some of the results of the acoustic data were unanticipated. When /g/ undergoes devoicing to [k], it becomes similar to [k] with respect to voice onset time but remains similar to [g] with respect to burst intensity. Thus, /g/ devoicing in Magong Taiwanese does not create a full phonetic merger with /k/.

Acknowledgements

Chapter 1
Background on Taiwanese

1.1 Introduction

The Taiwanese language is the native language of the majority of Taiwanese people, and
has a unique history of both suppression and acceptance in recent history. Deemed a sub-
standard or even politically dangerous language, Taiwanese has been actively suppressed
throughout most of the 20th century by both the Imperial Japanese and the Nationalist
Chinese. However, through regime changes, social upheavals and technological advances,
Taiwan has changed enormously in the 20th century, and with it so too have the languages
of its peoples. Now, in the 21st century, Taiwanese is spoken and heard alongside Mandarin
in Taiwan, and enjoys a status of full acceptance as a major language of Taiwan, though
education is done in Mandarin. Taiwan has undergone massive industrialization and quan-
tum leaps in its infrastructure and economy within the past 100 years, and this has forever
altered the way that most Taiwanese live. The Penghu Islands are one such place. Origi-
nally an archipelago with many fishing towns, Penghu County has since turned to tourism
after overfishing depleted natural resources (Taipei Times, 2002). Additionally, a lack of
economic opportunities in Penghu have caused younger residents to leave the islands for
work, causing the population to shrink.

Along with their economy and their lifestyle, the language of Penghu residents has
also changed in the past 50 years; older generations continue to speak Taiwanese almost
exclusively, but younger generations speak Mandarin as well, and Taiwanese will almost
certainly see more demographic change in the following 50 years as well. Previous research
(Kubler, 1978) has shown that Penghu Taiwanese has changed greatly in the past 30 years;
sound changes, such as the diminishing of unique dialectal features, have changed Penghu
Taiwanese, and there are many explanations for why this may be. One explanation may
be that linguistic universals are driving sound change for natural reasons; another could be

that Mandarin has become the language of status and education since the 1950s and come to pervade every aspect of Taiwanese life, and influence from Mandarin may be affecting people's speech. To potentially answer these questions, this honors thesis is both a dialectal and acoustic investigation of Taiwanese voiced stops. The following study hopes to capture these changes in progress and apply a linguistic analysis to them, so that we can not only better describe the unique linguistic situation of Penghu Taiwanese, but also frame these questions in linguistic terms to better understand how and why this situation has come to be.

1.2   The Taiwanese Language

The language referred to as "Taiwanese" in this paper has many different names. Because this language came to Taiwan along with settlers from the modern-day Fujian province of mainland China, it is often called the same names as its sister languages/dialects across the strait. These are often collectively called "Southern Min" (Minnanhua or Minnanyu), and the word "Min" is the traditional name for the area called Fujian from which these languages are thought to originate. They are also called Amoy. Likewise, the language can also be called Fujianese, Fukkienese or Hokkien, based again off of the provinces name. Some call it a "dialect" of a greater Chinese language, a somewhat misleading classification as many of these so-called "dialects" of China are not mutually intelligible and are linguistically separate languages.

These different names also carry some political baggage; Taiwanese people are more likely than mainland Chinese people to call the language "Taiwanese" (Taiwanhua or Taiyu, lit. "Taiwanese language"), and some Chinese people feel that the use of the word "Taiwanese" contains overtones of Taiwanese independence from China by stressing the languages uniqueness to Taiwan. Although "Taiwanese Min" is a more politically neutral term for this language, I have chosen to use "Taiwanese" in this analysis, both for simplicity as well as because my informants and participants called it that. The reader should simply understand that these many different names can be used to refer to the language studied in this paper, and that the choice of name is often a politically tinged decision.

1.3   A Brief Introduction to Taiwanese History

The original inhabitants of Taiwan are Austronesian peoples speaking languages in the Austronesian language family. The Taiwanese language came to Taiwan and to the Penghu

2

Islands in the 17th and 18th centuries, when large waves of Han Chinese people emigrated from Chinas modern-day Fujian province and settled in the islands.

During the turn of the century, the Japanese fought and won the Sino-Japanese war against Qing dynasty China. As part of the concessions of Chinas defeat, China ceded Taiwan to Japan, and Taiwan became a colony of Japan. The Japanese were responsible for the subsequent modernization and industrialization of Taiwan, but proved to be ruthless occupiers who inculcated the idea that the Taiwanese people were now second-class citizens of a Japanese empire. For the next 50 years, Japanese became the language of education and status, and all Taiwanese schooling was done in Japanese. Japanese became a lingua franca of sorts for the Taiwanese; people of disparate ethnicities who could not speak each others language or dialect used Japanese as a method of communication.

After the end of World War II, Taiwan reverted to de facto Chinese control and Japanese ceased to be taught in schools. Following the defeat of the Nationalists to the Communists during the Chinese Civil War in 1949, Chiang Kai-shek and the Nationalist government fled to Taiwan. At this time, a large population of several million Mandarin-speaking mainland Chinese moved to Taiwan, many of them soldiers, flooding the area with Mandarin speakers. Afterwards, the government in Taiwan established martial law, and speaking the Taiwanese language was forbidden in an effort to suppress the Taiwanese identity of Taiwan. For a Nationalist government whose legitimacy was based on its claim to all of China, this Taiwanese identity was found to be dangerous, and Mandarin became the language of education and prestige. Following the democratization of Taiwan and a liberal upheaval of society in the 1980s and 1990s, the Taiwanese language underwent a mainstream revival and gained new prominence as a unique cultural birthmark. Now, Taiwanese can be heard everywhere in Taiwan, both in the streets and on television. Some schools now teach Taiwanese as a second language, including Wenguang Elementary School where I conducted research.

Currently, approximately 70% of the population of Taiwan speaks Taiwanese as a native language, 15% speak Hakka, about 12-13% speak Mandarrin and the small remainder speak aboriginal Austronesian languages. However, these figures are complicated by the fact that Mandarin is the language used in the classroom, meaning virtually all Taiwanese who attended school since the end of WWII possess a strong facility, if not native fluency, in Mandarin. Due to both this and Mandarins higher prestige, more and more Taiwanese people are beginning to speak only Mandarin, These figures are only meant to give an estimate of the percentage of the population who speaks a language as their first and primary language, and there are clearly many people who grow up speaking Taiwanese and learn Mandarin and vice versa.

1.4   The Penghu Dialect of Taiwanese

The Penghu islands, also known as the Pescadores Islands, are a group of small islands to the west of Taiwan owned by the Republic of China (Taiwan). Penghu is split into roughly 6 different regions: Magong, Huxi, Baisha, Xiyu, Wangan and Qimei. The largest city in Penghu County is Magong, and the total population of the islands numbers around 90,000.

The Penghu islands lie between China and Taiwan, and almost all residents of Penghu are ethnic Han Chinese and speak Taiwanese, a dialect of the Southern Min language (also called Minnan, Hokkien, or Fujianese) in the Sino-Tibetan language family. Min is hypothesized to have split from Ancient Chinese and the other Chinese languages at around 0 CE, compared to languages such as Mandarin, Cantonese (Yue) and Shanghainese (Wu), which split from each other at around 600 CE. In this way, Min and Taiwanese are less similar to the other Chinese languages.

1.5   A Brief Linguistic Description of the Taiwanese Language

Taiwanese, like other Chinese languages, is a tonal language, and every syllable has an underlying tone. Taiwanese is approximately an analytic language with some limited inflections, such as an ending "-n" that marks a plural pronoun. Taiwanese does form a large number of words using monosyllabic morphemes that combine to form disyllabic words. However, Taiwanese words do not generally exhibit different forms and do not mark for case, gender, plurals, tense, etc.

Table 1.1: Taiwanese consonants

voiceless Stop	p	t	k	ʔ
voiced Stop	b		g	
nasals	m	n	ŋ	
affricates	dz	ts	tsʰ	
fricative		s		h
lateral		l		

Taiwanese has voiceless stops /p/ /t/ /k/, voiced stops /b/ /g/, and voiceless aspirated stops /pʰ/, /tʰ/, /kʰ/. Taiwanese also has nasals /m/, /n/, /ŋ/ and affricates /ts/, /dʒ/, /tʃ/ and /tʃʰ/. Before the high front vowel [i], affricates and fricatives can become alveo-palatal. In the coda, only the nasals and the voiceless stops /p/, /t/, /k/ and /ʔ/are allowed.

4

Taiwanese is noteworthy in that while most of the Chinese languages have a two-way phonemic distinction between voiced and voiceless consonants /ɡ k/ or voiceless and voiceless-aspirated consonants /k kʰ/, Taiwanese has a 3 way phonemic distinction: /ɡ k kʰ/.

Table 1.2: Taiwanese vowels

i		u
e	ɔ	o
	a	

Taiwanese has a 6 vowel system of /i/, /e/, /a/, /o/, /ɔ/, and /u/. In addition, syllabic [m] and [ŋ] are also allowed. Taiwanese syllable structure is strict, allowing at most an onset consonant, vowel (monophthong or diphthong), and a coda consonant. All consonants may appear in the onset except the glottal stop. Phonemically speaking, only the nasal consonants, the glottal stop, and the plain (unaspirated, devoiced) stops are allowed in codas; all others are barred. However, phonetically speaking, assimilation of a coda to the following liquid [l] allows [l] to appear in the coda as part of a geminate.

Traditionally, Chinese scholars have counted up to 8 tones in the Southern Min languages. However, syllables with stop codas have traditionally been counted as separate "tones" for their short duration, despite the fact that in modern Taiwanese the tones of [p] [t] [k] coda syllables can be described as shorter allophones of other tones. Additionally, the tones classically numbered 2 and 6 are merged, meaning that linguistically speaking, Taiwanese has 5 phonemic tones.

Taiwanese is also known for having complex tone sandhi rules. Syllables have an underlying tone that is expressed in isolation, but changes to another tone by specific rules when in an utterance. All syllables undergo a tone change when either at the beginning of an utterance or anywhere in the middle; only the last syllable in an utterance is unaffected by tone sandhi rules.

Taiwanese has multiple strata of words in its lexicon, representing words that entered the Southern Min language at various points in history. Having split off at an early point from the other Chinese languages, Taiwanese has a stratum of lexical items that remain from Old Chinese, such as:

- [twa] 'big' (cf. Mandarin [ta], Cantonese [taai])

- [gwa] 'me, I' (cf. Mandarin [wo], Cantonese [ŋo], Sino-Korean [a], Sino-Japanese [ga])

- [tjɔŋ] 'middle' (cf. Cantonese [tʃʊŋ], Sino-Korean [tʃuŋ]

5

Additionally, Taiwanese has a stratum of literary morphemes that entered the language at some time during the first millenium, well after Southern Min had diverged from the other Chinese languages. These literary morphemes are used in words like:

- [taihak] 'university' (c.f. Cantonese [taaihɔk], Sino-Korean [tɕhak])

- [sam], literary variant of '3' (cf. Mandarin [san], Cantonese [sam], Sino-Korean [sam], compare with Taiwanese non literary variant [sã])

This stratum can be a "false friend" in the study of Southern Min historical linguistics, and the study of this stratum can give the appearance of a much closer relationship than we know to be possible between Min and Mandarin. Taiwanese also has many words which are not cognates with Mandarin, such as the word [lang] meaning 'person', a word unrelated to Mandarin [ren] but instead derive from a word meaning 'peasant' ([nong] in Mandarin).

Figure 1.1: General Map of Penghu

Chapter 2
Dialectal Analysis

2.1   Previous Research: Prof. Cornelius Kubler

The first portion of this analysis is directly based on fieldwork collected by Prof. Cornelius Kubler of Williams College. In 2003, Kubler collected a large amount of linguistic data from a number of schools across Penghu County as part of a longitudinal study of changes in the Penghu dialect. As a comparison to similar research of the Penghu dialect performed 25 years earlier (Kubler, 1978), Kubler collected data to determine to what extent certain dialectal features unique to areas of Penghu County were still present, and to track dialectal change across the county.

   In his study, Kubler interviewed 178 Penghu residents between the ages of 12 and 16 using a word list containing 50 lexical items. All items on the list were written in Chinese characters, and participants in the study were asked to read these items out loud. Kubler then took notes on what word they said and the way they pronounced it on a separate, identical word list with the participants identifying information on it. This identifying information for participants included their name, where they were born and raised, where they attend school, and the extent to which they speak Taiwanese at home. Participants interviewed came from over 15 different school, and on average there were approximately 2 to 3 participants representing each small town of Penghu.

2.2   Dialectal Analysis

In this data, many Penghu respondents deleted an underlying /g/, and moreover there was a great deal of variation in the data as to when /g/ was retained or deleted. Further examination showed that /b/ was also undergoing sporadic deletion, leading to the conclusion that Penghu Taiwanese may be exhibiting a general trend to delete voiced stops. To determine

if this was the case, Kubler's data was examined on a large scale, and several different types of changes to voiced stops were discovered:

- /g/ remains [g]

- /g/ nasalizes to [ŋ]

- /g/] becomes [ji]

- /g/ deletes

- /b/ remains [b]

- /b/ nasalizes to [m]

Furthermore, these sound changes vary in their degree of ubiquitousness; /g/ deletion is relatively common across Penghu, whereas a change from /gi/ to [ji] is quite rare. Thus, the focus of this dialectal analysis is to get a broad sense of the extent to which these different sound changes occur all across Penghu, and on examining whether specific areas may be characterized by the way they pattern to these sound changes. In order to classify the various areas of Penghu as belonging to one dialect group or another, 4 broad sounds were examined in the data:

- Deletion of the glottal stop /ʔ/

- Deletion of /w/

- Nasalization of /b/ to [m]

- Shifts in /g/ to [ŋ], [j] or deletion

Although the first two sound shifts (deletion of /ʔ/ and deletion of /w/) are not the primary focus of this research, they are nonetheless still valuable because their alternation is binary and their changes are relatively wide-spread, allowing us to see which areas are patterning with each other.

To measure the degree to which shifts were occurring, each word for which the majority of participants responded with either the target sound or one of its allophones was examined. The total number of attestations for each were then counted and summed for each town, and an overall distribution of phonological behavior by town was catalogued. Next, the distribution of each town was compared to its neighboring towns and its geographic area, and lines marking dialect regions of similar behavior were drawn. After

grouping towns into dialect regions, data from each town within the region were summed and the frequency of the underlying form or the allophone was calculated as a percent of the total. Due to variation in the data, a few towns have been included in dialectal areas with which they do not pattern ideally; however, the following divisions attempt to portray dialectal divisions that are true to how both the town and the greater area patterns. Of note is the fact that because acoustic data was not recorded during Kublers interviews, it was impossible to determine whether this sound change indicated a full deletion or whether there was some kind of gradation to this deletion, such as milder kinds of lenition which might go unnoticed in a purely impressionistic analysis.

Figure 2.1: Deletion of Glottal Stops; The dark grey regions are where /ʔ/ is deleted 50% of the time or more, while the light grey regions are where it is deleted 25–50% of the time. White regions show areas where /ʔ/ is deleted less than 25% of the time.

Figure 2.2: Deletion of /w/; Again, the dark grey regions are where /w/ is deleted 50% of the time or more, while the light grey regions are where it is deleted 25–50% of the time. White regions show areas where /w/ is deleted less than 25% of the time.

Based on this dialectal mapping, we can tentatively identify 5 regions which appear to pattern together: Magong+western Huxi, eastern Huxi, Pengnan, Baisha, and Xiyu[1]. Pengnan, an area historically categorized as belonging to the Magong sub-county, appears to be patterning separately from Magong, suggesting that Pengnan may be its own dialect region. Interestingly, the data also indicates that the western parts of Huxi sub-county pattern similarly to the city of Magong with respect to these 4 sound changes. It is possible to tentatively hypothesize that the dialect of Taiwanese spoken in Magong may be spreading to neighboring regions of Penghu County. An airport serving the Magong area is located

---

[1] Much thanks to Nathan Sanders for his great help in constructing the maps.
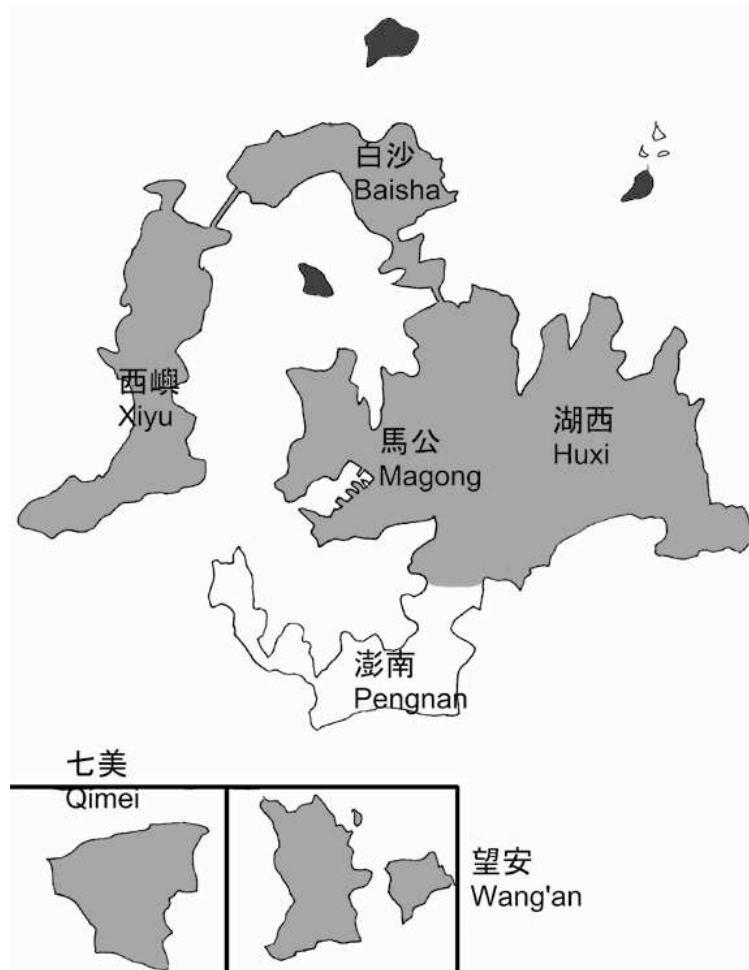
Figure 2.3: Nasalization of /b/ to [m]; The dark grey regions are where /b/ is nasalized 50% of the time or more, light grey regions 25–50% of the time, and white regions show areas where /b/ is nasalized less than 25%.

in the town of Chenggong in the western part of Huxi, and major roads connect Magong to this airport. The influx of people going to and from Magong through Chenggong and the ease of transportation between these two places may help to explain the similarities seen between Magong and western Huxi.
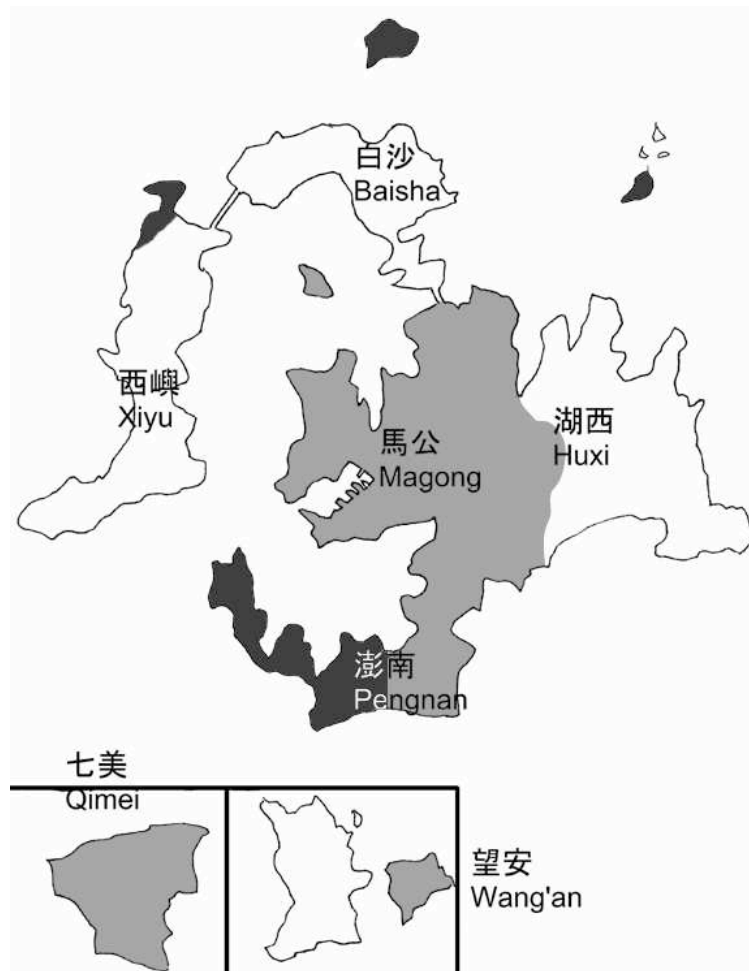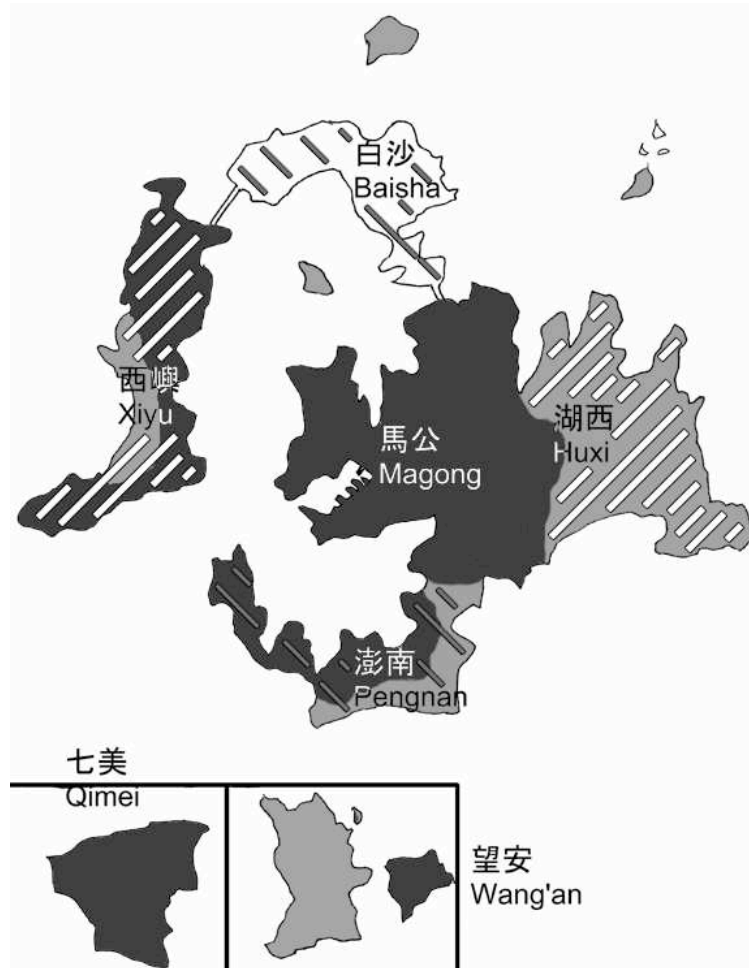
Figure 2.4: Deletion and shifts in /g/; The dark grey regions are where /g/ is deleted 50% of the time or more, while the light grey regions are where it is deleted 25–50% of the time. White regions show areas where /g/ is del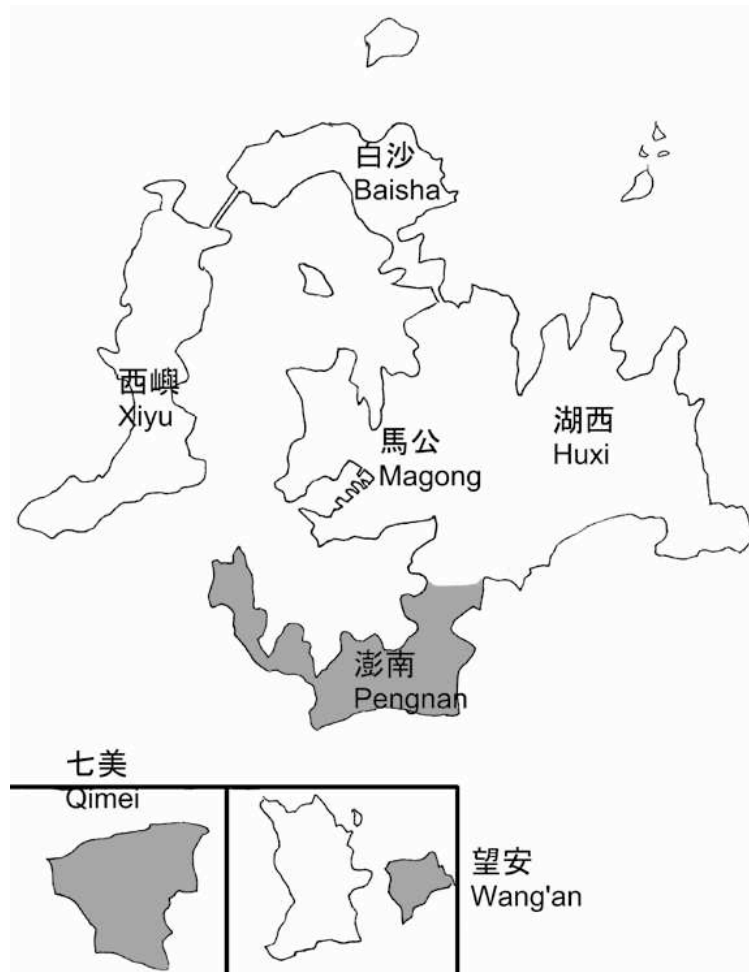eted less than 25% of the time. In addition, regions with dark falling stripes indicate areas where /g/ is being nasalized at least 10% of the time, while regions without dark falling stripes show that nasalization of /g/ is occurring less than 10% of the time. Similarly, areas with rising light stripes indicate areas where /gi/ is becoming [ji], and areas without rising light stripes show areas where this occurs less than 10% of the time.

2.3 Analysis of Voiced Stops

As we see in Figures 2.3 and 2.4, the voiced stops /b/ and /g/ are undergoing shifts in various areas of Penghu. The following phonological analysis focuses on how voiced stops /g/ and /b/ are realized throughout the Penghu Islands. We find the voiced stops /b/ and /g/ undergoing several changes, such as nasalization, /g/ shifting to [j], and even complete deletion:

- Retention of /g/: 'moon; month' pronounced as [gwaʔ44]

- Nasalization of /g/ to [ŋ]: for example, for 'goose', [gɔ24] alternates with [ŋɔ24].

- Deletion of /g/: for example 'moon; month' alternates between [gwaʔ44] vs. [waʔ44] and [gwe44] vs. [we44].

- Alternation of /g/ and [j] before [i]: as in the word for 'silver', [gin24] becomes [jin24].

- Nasalization of /b/ to [m]: for example, in 'younger sister', [bwe22] alternates with [mwe22].

What is occurring here is a conspiracy against /g/, an overarching tendency to avoid [g] by different speakers in different ways, but all for the same underlying reason. Also, if we look at [b], we see a similar conspiracy on a smaller scale, leading us to conclude that the same universal drive to avoid voiced stops may be fueling both sets of changes.

2.3.1 Why delete or change voiced stops?

Universally speaking, voiced stops are more difficult to articulate than voiceless stops or other voiced sounds. However, specifically in Taiwanese, there is a strong reason for a tendency to exist to change voiced stops. Taiwanese, as mentioned before, makes phonemic distinctions between /g/, /k/, and /kʰ/, something relatively rare in the Chinese languages. Because these three consonants are articulated in very similar ways with only minor differences, they are difficult to acoustically differentiate, especially [g] and [k]. Thus, a natural tendency arises to increase the distinctiveness of these consonants, and an easy and natural way is to increase the sonorant nature of [g] and turn it to similar (and more acoustically differentiable) sounds like [ŋ]. In addition, another important reason is that [g] is more difficult to produce than either [k] or [kʰ], making it natural for [g] to be the one changing and not [k] or [kʰ]. Similarly, since Taiwanese also differentiates between /p pʰ/, it is natural for the same reasons to see /b/ nasalizing to [m] and not to see /p/ or /pʰ/ undergoing shifts.

2.4 Conclusion

The data suggests that the Penghu Islands can be categorized into 5 different dialect regions by the different degrees to which they undergo different sound changes. Deletion of /g/ is clearly most prevalent at Magong and in the surrounding area; whether this is due to the actual spread of this change to neighboring areas or coincidence remains to be ascertained, but is likely given the relative prestige of Magong compared to its neighboring towns. We have also identified the linguistic universals inciting a conspiracy against Taiwanese voiced stops /g/ and /b/ and identified different ways in which this conspiracy may be acting to change voiced stops.

Since it appears that the Penghu /g/ is in the process of moving away from [k] to either delete or become [ŋ], it is possible that words with an underlying /k/ are being affected by this change as well. This would be due to the open space that [g] leaves after migrating, opening up the possibility for the voice onset time of [k] to become shorter, a conclusion which would also demonstrate a general conspiracy against /g/ in Penghu Taiwanese. As previously mentioned, given Kubler's non-acoustic data, it is impossible to determine whether more minor lenitions of voiced consonants may be taking place in addition to full deletion and nasalization. Acoustic data may potentially corroborate the phonological hypotheses presented here in.

Chapter 3
Research Methodology

3.1   Interviews

In order to test the hypothesis of a general trend of "conspiracy" acting against voiced stops in Taiwanese, acoustic data was also collected from school-age children in Penghu County using an interview format. I spent approximately one week in the city of Magong in Penghu County, and collected phonetic data at Wenguang Elementary School and Wenguang Middle School. The interview consisted of three parts: asking participants to describe their average day, asking participants to read a word list out loud, and asking participants to read several sentences out loud. During this time, I not only collected the raw data necessary for my analysis, but also was able to speak with Penghu residents about local opinions on language and the future of Taiwanese. Acoustic data was collected primarily to answer two questions concerning Taiwanese stops: whether acoustic properties of [k] were becoming more similar to those of [g], and whether /g/ deletion is indeed taking place and to what phonetic extent.

In part 1, instructions were given to each participant in standard Mandarin to describe their average day from start to finish. During the instructions, emphasis was placed on using colloquial, natural Taiwanese that participants would speak at home or with friends. In addition, instructions were given in a colloquial and chatty manner meant to calm whatever anxieties participants might have felt during the interview. The purpose of part 1 was two-foldto put participants at ease and allow them to reach a state of speaking natural, colloquial Taiwanese, but also to gather data for potential use in future studies of the Magong Penghu dialect. Participants generally took about 1 or 2 minutes to recount their daily activities; some spoke for longer periods and some only volunteered a few sentences. When it appeared that participants had nothing more to say about their daily routines, the interview proceeded to its next part.

In part 2, each participant was then asked to read a word list out loud in Taiwanese that they would speak at home or among friends. Although a Romanization for Taiwanese exists, Taiwanese people themselves write Taiwanese using Chinese characters (as mentioned previously), prompting the choice of a word list in Chinese characters and not a Romanization. In addition, the choice of a character word list was somewhat necessary; Taiwanese children use the Zhuyin Fuhao system and not Pinyin to phonetically write Chinese in school, which meant that familiarity and comfort with the Roman alphabet could not be guaranteed. For these reasons, the word list was written entirely in Chinese characters. This did not mean, however, that word list items were all written in Mandarin; many common Taiwanese words are etymologically different from the equivalent Mandarin words and are written using Taiwanese-specific characters (cf. "today" Mandarin [ɕintjen] vs. Taiwanese [kinadʒit]). The Taiwanese way of writing a Taiwanese word was used whenever possible in constructing the word list. This helped ensure that participants read a true Taiwanese word as opposed to attempting some kind of morphological substitution between Mandarin and Taiwanese. After finishing the word list, participants

In part 3 of the interview, participants were asked to read several scripted sentences in Taiwanese, the vocabulary and grammar of which were extremely basic in the hope that they would have no difficulty reading. Most of the words in the sentences were repeated from the word list to determine if participants would read them differently in isolation (excluding the tone, which necessary would be different due to tone sandhi). Part 3 served two main purposes; first, to act as a second data source for a comparison with the word list in part 2, and second, to provide possible future data for an acoustic analysis of tone in Penghu Taiwanese. The items on the word list and the sentences were again written using Chinese characters in a semi-standard Taiwanese orthography, using Taiwanese characters whenever possible to produce natural reading and to reduce the possibility of translation from Mandarin to Taiwanese.

3.2   Issues in the Collection of Taiwanese Data

Because Taiwanese is predominantly a spoken-only language that lacks a standardized written language, some unique issues arise in the collection of any data. There are different ways of writing Taiwanese, the two most common being a jumble of Chinese characters used for meaning and phonetic value, and a romanization. Although romanization is used to teach Southern Min in the United States, the Taiwanese themselves use a non-standardized system of Chinese characters to write Taiwanese. This system sometimes employs characters for their phonetic value in approximating Taiwanese words, sometimes

employs characters for their meanings in Mandarin, and sometimes borrows characters to represent cognate morphemes. In many instances, characters that do not exist in Mandarin have been created for Taiwanese words, and often these "dialect characters" have a long and entrenched tradition in society. Moreover, variant characters exist for words as basic as "no, not"—for those without strong fluency in both Taiwanese and Mandarin, this system can be confusing and unintuitive.

Issues arise when we consider the difficulty of phonetic solicitation using written forms in a non-alphabetic, non-standardized language like Taiwanese. Romanizations are likely unknown, and the use of characters cannot guarantee that the target word will even be articulated. Some respondents in Kubler's study seemed to be reading the character prompts using Taiwanese equivalent morphemes, suggesting that they were simply reading by character and failing to translate into Taiwanese; others translated the character prompts out of Mandarin but gave a multitude of Taiwanese words; some students could not identify certain words at all. Furthermore, character prompts may be priming Mandarin-like forms since Mandarin is the language of education.

These kinds of confusion among participants expose the problems faced when conducting phonetic and phonological studies of languages like Taiwanese with (somewhat) non-phonetic and non-standardized writing systems. Before beginning any of the following analyses, the author of this study took care to discount attestations that likely stemmed from student confusion.

3.3   Issues Pertaining to the Collection of Acoustic Data

Several issues arose concerning the methodology specifically relating to the collection of this acoustic data. Instructions for all three parts of the interview were given in standard Mandarin to the children, a factor that could potentially be inducing a bias among participants to produce prestigious or Mandarin-like Taiwanese forms during the interview. This was due in large part to my limited facility in Taiwanese and the necessity for clarity in relaying instructions. Additionally, previous discussions have already detailed the difficulty of providing accurate written cues in the Chinese morpho-syllabic writing system, especially in a school setting where Mandarin is cued. Finally, Penghu County is an area of Taiwan in which Caucasians are rarely seen, and the presence of a Caucasian male performing an acoustic study may have had some unpredictable effect on participants, in particular on participants feelings of anxiety. Some participants (mostly older ones) were clearly nervous during the interview, a factor which may also have unpredictable effects on their responses.

However, I predict that any additional bias produced by the factors mentioned above will likely be minimal for several reasons. All participants were students in elementary and middle schools where the primary language of instruction and communication is Mandarin, making the use of Mandarin in giving instructions unlikely to have any strong impact on participant responses. Also, as previously mentioned, orthography used by Taiwanese people themselves was used as much as possible in the word list and sentences, hopefully minimizing the effect of Mandarin influence and lessening participant uncertainty. Whatever consequences may have arisen out of the experimenter being a Caucasian male I have deemed unavoidable; with a limited time frame of several days within which to collect acoustic data, training Taiwanese assistants was impossible. There also seems no reason for me to assume that data collection by a Taiwanese assistant would have produced more authentic or different results; I believe that when participants felt nervous, it was due to the act of being interviewed and not due to my race, and these participants most likely would have also felt nervous in an interview by a Taiwanese assistant. In short, steps were taken to minimize unwanted effects on participant responses, and I am relatively confident that the results of collecting data in this fashion represent the way participants actually speak Taiwanese.

3.4   Demographics

For this study, 10 elementary school children were interviewed at Magong City of Penghu County. At my request, the teachers of Wenguang Elementary School chose 7 students from their classes who spoke Taiwanese at home and allowed me to record their speech in a small room adjacent to the headmaster's office. Of these 7 students, 3 were boys and 4 were girls, and all said that they lived in Magong City. On the following day, I was driven to Wenguang Middle School, and the teachers there chose 3 students for me to interview, all girls. In all, 7 girls and 3 boys were interviewed, and all said that they spoke Taiwanese at home.

3.5   Recording and Analytical Instruments Used

Two extremely important issues to consider when collecting acoustic data are the quality of the recorded sound and the impact that undesired factors will have on data analysis. In collecting this data I was of course unable to access the comforts of a phonetic laboratory,

such as a sound-proof room or advanced recording equipment, and was instead forced to rely on recording devices that I personally brought to Magong from the United States.

To minimize the effects of unwanted sounds, I chose to interview the students only when classes were in session, thus guaranteeing at least in part that the environment outside of the room in which I recorded data would be less contaminated with students' footsteps and conversations. Gigaware Digital USB Stereo Headphones with a built-in microphone and the software program Audacity were used to directly record the students' responses into a computer, and the phonetics software Praat was then used to analyze the recorded sounds.

3.6   Measurements

To measure shifts in underlying /k/ and to verify lenition or deletion of underlying /g/, consonant burst intensity and voice onset time (VOT) were measured for each consonant. Voiced stops have less intense bursts than voiceless stops, and differences in VOT are widely considered to be the phonetic bases for distinguishing voiced from voiceless stops (Lisker and Abramson, 1964). VOT and burst intensity were both measured for every underlying voiced, voiceless and voiceless aspirated velar and bilabial consonants. In the event that a consonant clearly had no burst, the burst intensity was marked as 0. Furthermore, burst duration was measured to determine whether it would correlate with voicing in oral stops. Because different speakers spoke at varying degrees of loudness, the maximum burst Intensity of the syllable was also recorded as a means to normalize the consonant burst intensity and eliminate discrepancy due to soft or loud speech. Similarly, the duration of the vowel and coda was also measured as a potential way to normalize burst duration.

Chapter 4
Analysis and Results

4.1   Results of Interviews with Penghu Residents

I stayed with a family who ran a bed and breakfast next door to the elementary school, and this family took me around the area and introduced me to a number of people across Penghu County. During this brief stay, I learned much from both my host family as well as other acquaintances about the attitude of the local Taiwanese toward Taiwanese and Mandarin, as well as general information about Penghu itself. The following is information I learned by word of mouth while staying in Magong, and thus should not necessarily be construed as the truth but simply as the results of obtaining information from a number of Penghu Taiwanese people. According to my hosts and other residents with whom I spoke, Penghu is a quiet area with a large amount of linguistic diversity. Some expressed concern that young people were moving away from Magong to find more job opportunities elsewhere. Although it is not an especially wealthy area, Penghu is considered a vacation spot during the summer months, and its economy is partly based on the production of local goods and tourism.

Non-governmental residents of Penghu County generally had a long history in the area and were not recent immigrants. My hosts said that many families (including their own) came to Penghu during the Qing dynasty, and these immigrants were from Fujian across the strait and not from across China. Also, they said that after the Kuomintang (Guomindang or KMT) nationalists fled to Taiwan and established their government there, a number of soldiers were stationed in Penghu but immigration had otherwise been minimal. For these reasons, virtually all residents of Penghu speak Taiwanese, and many have been living there for centuries. According to my hosts and the different people I came across in Magong, fewer families speak only Taiwanese at home than before. This was said to be because many parents speak to their children in Mandarin and avoid speaking Taiwanese,

preferring to emphasize the language of education and employment to boost their childrens success in school.

Of the many people I met, I noticed that elderly people (generally 50+) in Penghu County tended to speak only in Taiwanese amongst themselves (whether they were familiar with Mandarin or not), and preferred not to use Mandarin. It is an interesting fact that many of the oldest Taiwanese people were children when the Japanese still controlled Taiwan, attending school in Japanese and becoming bilingual in Japanese and their native Taiwanese. These elderly people had left school by the time the KMT nationalists took over Taiwan and imposed Mandarin in schools, and consequently never learned Mandarin.

The middle-aged generation (around 35–50) seemed most versatile in their use of Mandarin and Taiwanese. These people, among them my hosts and some teachers, were equally comfortable in either Taiwanese or Mandarin, and seemed to use them in different social contexts. Among non-immediate family members such as other middle-aged siblings or cousins, only Taiwanese was spoken; with small children, including their own, they spoke mostly in Mandarin with some Taiwanese; in more formal non-familial situations, such as in school or at the workplace, conversations were in Mandarin; and in informal non-familial social situations, Taiwanese seemed to be preferred. Thus, one can see that there is a general socio-linguistic trend to use Taiwanese as the language of familiarity and Mandarin as the language of education and employment, a conclusion that supports a previous assertion that Mandarin is a more prestigious language. However, this socio-linguistic description was by no means binding; there were varying degrees of variation in the way people used these two languages. This age-group also had the most code-switching and code-mixing, and many conversations would begin in one language and switch back and forth. When asked about code-switching and code-mixing, one adult replied that she finds Taiwanese to be more precise in conveying certain nuances and prefers Taiwanese in cases where clarity of emotional expression is important.

Young adults showed as much variation in their speech as middle-aged adults, but seemed more likely to speak Mandarin during regular conversations. Most of the teachers at Wenguang Elementary were young adults ranging from ages 20 to 30. These young adults had no problems using either Mandarin or Taiwanese, but chose Mandarin as their primary language of communication amongst themselves. Code-switching and code-mixing were also extremely common among young adults. Generally, young adults seemed just like middle-aged adults in their sociolinguistic use of Taiwanese and Mandarin; the only difference was that young adults seemed to default to Mandarin.

Children showed mixed use of Taiwanese and Mandarin, but the children I encountered at Magong showed a preference for using Mandarin as a default amongst themselves.

The children of the family with whom I stayed mostly spoke Mandarin, though they under-stood spoken Taiwanese and were taking a Taiwanese course at school. This family owned a small hotel and was probably economically advantaged compared to other residents, but most children I encountered used Mandarin almost exclusively.

4.2   Analysis of Acoustic Data

As shown in Chapter 2, deletion and shifts of /g/ were most common in Magong, leading me to hypothesize that the data collected from Wenguang Elementary School and Wenguang Middle School would show a large amount of /g/ deletion. This hypothesis turned out to be correct. Every participant that I interviewed had at least some total deletion of /g/, though there were varying degrees of this deletion. Most also had some partial changes of /b/, where /b/ in standard Taiwanese was articulated as something similar to [v] or [m].

Some dialectal differences existed even among students of the same elementary school who had grown up in the same area. For the morpheme meaning 'language', some speakers said [gu] instead of the standard [gi], a switch that my Taiwanese correspondents said was typical of certain areas of Penghu such as Xiyu subcounty. In addition, all respondents used [gun] instead of [gwan] for 'we (exclusive)', suggesting that the use of this form may be characteristic of all Magong speech.

The acoustic data collected not only showed deletion and retention of [b] and [g], it also showed some intermediary forms that were not and could not have been part of the data used in the dialectal analysis. In the place of [b] and [g], many of the participants interviewed articulated voiced consonants of varying stop-ness, consonants in between an approximant and a stop or in between a nasal sonorant and a stop. Overall, participants were more likely to delete or change /g/ and were more likely to express /b/ as a true voiced stop.

Very few underlying /g/s in the data were actually expressed on the surface as voiced stops; many were deleted, such as the word 'language' (/gigjen/ or /gugjen/ expressed as [uen]), some were partially deleted to a sound somewhere in between [g] and [ŋ]. Sur-prisingly, many underlying /g/s were expressed on the surface as [k]. Every respondent deleted [g] in the word I /gwa/, an observation in line with our hypothesis that common words would undergo this deletion to a greater extent. Generally, underlying /g/ was mostly deleted or changed to a different sound.

On the other hand, most underlying /b/s did not undergo any deletion or change; the words for to sell, to buy and to not have were almost never deleted or changed. One notable exception to this trend is the word for America /bikok/, a word where most people articu-

lated a partially-deleted sound in between [b], [m], and [v]. This phenomenon is difficult to explain; we would expect higher-frequency words to undergo this sound change more than lower-frequency words, but words like not have, buy and sell are clearly more common in regular speech than the word for America. In fact, many students at Wenguang Elementary school did not understand that America was a country, and instead thought that America meant some foreign place, leading me to believe that they had had very little exposure to American culture and knew little about the world outside of Taiwan. This effect may be due to some unknown influence of the high-front vowel /i/ that follows /b/.

Finally and most surprisingly, [k] for most speakers was unchanged; rather, for many speakers, [g] was conversely becoming more like [k]. Approximately 1/4 of all [g] attestations had bursts preceding the onset of voicing, and certain participants were more likely to do this than others. A possible explanation for this may be the relative articulatory ease of [k] versus [g]. Another explanation may be Mandarin influence; since Mandarin lacks voiced oral stops, speakers applying Mandarin phonology to Taiwanese would be likely to underlying shift /g/ to [k].

4.3   Statistical Analysis of Acoustic Data

As detailed in Chapter 3, burst intensity, burst duration, the syllable's maximum intensity, vowel+coda duration, and voice onset time were measured for every syllable beginning with /k^h/, /k/, /g/, /p/, and /b/ in the word list data. An analysis of mean VOT for bilabial and velar oral stops is listed in the table below:

Table 4.1: Mean VOT Results

k^h	.0698
k	.0206
g	-0.0102
p	0.0035
b	-0.0218

The majority of tokens collected were of /g/ and /k/; fewer /b/ were included in the word list, only 2 instances of /p/ existed in the word lists, and only 1 dependable example of /k^h/ was able to be used for each speaker. It is important to note that while acoustic analyses of [k^h] and [p] may be insightful in understanding voicing in Taiwanese, they both rely on less than 10 total data points and any statistical comparisons made with them will most likely be statistically insignificant.

4.3.1   Analysis: Highest and Lowest Quartile [g] VOT

The voiceless stop [k] showed comparatively little variation in its surface forms, and thus will not be discussed in as much detail. However, voice onset times for underlying /g/ showed much more variation than VOTs for underlying /k/. A little over half of all underlying /g/'s had undergone total deletion in the data. Conveniently, all [g] articulations with negative voice onset times fell entirely within the bottom quartile of voice onset times for all [g] data, and similarly all [g] articulations with positive voice onset times fell entirely within the top quartile. Thus, calculating the mean VOT for the highest quartile of [g] VOTs was functionally equivalent to calculating the mean VOT of all "[k]-like" [g] attestations; similarly, calculating the mean VOT of the lowest quartile was functionally equivalent to calculating the mean VOT of all "[g]-like" [g]. The two remaining mid-high and mid-low quartiles consisted solely of deleted [g] with a VOT of 0. Mean VOTs are shown below:

- The mean VOT for the lowest quartile of [g] (all negative VOT values) was -0.0639.

- The mean VOT for the mid-high quartile of [g] was 0.

- The mean VOT for the mid-low quartile of [g] was also 0.

- The mean VOT for the highest quartile of [g] (all positive VOT values) was 0.0231.

A t-test of the significance of the difference between these two means yields a p-value of less than .0001, indicating a difference that can be considered extremely statistically significant. These positive [g] VOTs that fall within the highest quartile show that what is being pronounced on the surface in these cases is not phonetically [g] but [k]. The mean VOT calculated for this highest quartile of [g] VOT values is strikingly similar to the mean VOT for all [k] values. This can be taken as evidence corroborating an analysis where many speakers replace underlying voiced consonants with voiceless ones.

In the same vein, the extreme difference between the mean VOT for the lowest quartile of [g] VOTs (-0.0639) and the mean VOT for all attestations of [g] (-0.0102) merits attention; the lowest quartile contained all instances of [g] with a negative VOT, suggesting that these may be surface forms with true voicing and thus be categorically distinct from other surface forms of underlying /g/. This conclusion is logical; the mid-high and mid-low quartiles of [g] VOT contain only VOTs of 0, most of which are /g/ which have undergone deletion or lenition and are no longer oral stops, and since we hypothesize that the highest quartile of [g] VOT in reality may be wholesale replacement by [k], it makes sense to posit that [g] attestations with fully negative VOTs may be the only truly voiced surface forms. A t-test assessing the significance of this difference again yields a p-value of less than 0.0001,

indicating a highly statistically significant difference. Thus, the difference between mean VOT for the lowest 25% of [g] VOTs is statistically different from the mean VOT of all [g] attestations, corroborating an analysis of negative VOT [g] attestations as surface forms with true voicing and categorically distinct from the other 3 quartiles.

4.3.2   Analysis: Comparison of non-zero [g] and [k] Burst Intensities

Voiced stops generally have lower intensity bursts than voiceless ones do, and thus we can compare the mean burst intensities of [g] and [k] attestations to verify the hypothesis that fully voiced [g] and devoiced [g] are categorically distinct. Below is a table listing the results of measuring the average burst intensities of [k], HQ [g] and LQ [g] (HQ stands for "Highest Quartile VOT" and LQ stands for "Lowest Quartile VOT").

Table 4.2: Mean Burst Intensity Results of [k] and [g]

k	48.5761
HQ g	43.0067
p-value	0.0001
k	48.5761
LQ g	46.8000
p-value	0.0215
HQ g	43.0067
LQ g	46.8000
p-value	0.0662

The average burst intensity of [k] is clearly higher than that of [g], and the mean burst intensity of LQ [g] seems higher than that of the HQ. A t-test of the difference between [k] and HQ [g] yields a *p*-value of 0.0001, showing high significance; a t-test of the difference between [k] and LQ [g] also yields a significant value of 0.0215. However, the same test between HQ [g] and LQ [g] yields a value of 0.0662, making the difference questionably statistically significant (by convention, a *p*-value of 0.05 can be considered statistically significant for most comparisons). This is an unexpected result; we would expect the burst intensity of HQ [g] to pattern similarly to [k], and vice versa, we would expect that of LQ [g] to be more similar to non-zero [g]. Instead, we see [g] with positive VOT having a lower mean burst intensity than [g] with negative VOT, although the difference is not quite statistically significant.

Finally, the normalized burst intensity averages show a large difference between [k] and [g], but almost no difference between HQ [g] and LQ [g]. A t-test comparing the difference between the normalized means of LQ [g] and all [k] yields a *p*-value of 0.0395,

Table 4.3: Mean Normalized Burst Intensity Results of [k] and [g]

k	0.8231
HQ g	0.7801
p-value	0.0452
k	0.8231
LQ g	0.7742
p-value	0.0395

making this difference statistically significant; the same test for HQ [g] and all [k] also shows significance with a *p*-value of 0.0452. The same test yields no significant difference between HQ [g] and LQ [g]. These calculations confirm what is universally attested–that voiced stops have lower burst intensities than voiceless ones, but also that the burst intensity of [k] is still significantly higher than the average burst intensity of [g] that appears on the surface as [k]; this shows that devoiced [g]'s are not patterning like underlying /k/.

This makes for an interesting and puzzling result; although the average burst intensities of HQ [g] and LQ [g] differ, a normalized average shows no difference between them. The root of the original difference lies in an unanticipated factor: that the maximum intensity of syllables containing LQ [g] is significantly higher than the maximum intensity of syllables containing HQ [g]. This is most likely due to coincidence, and a future statistical test may show this to be the case.

Thus, we conclude that both the both absolute average burst intensity and normalized average burst intensity of [k] is significantly higher than that of [g]. We also conclude that a somewhat significant difference in absolute average burst intensity between HQ [g] and LQ [g] exists, a difference that disappears when these values are normalized against maximum syllable intensity. The most surprising conclusion of this burst intensity analysis is that there is a statistically significant difference in syllable intensity between /g/ that is expressed as a voiced stop and /g/ that is expressed as a voiceless stop.

Chapter 5
Optimality Theoretic Analysis

5.1   Optimality Theory Analysis of Penghu Voiced Stops

The manifestation of a conspiracy against certain sounds is prime support for an Optimality Theory (OT) analysis (Prince and Smolensky, 1993/2004), an analysis that allows us to the bypass language-specific nature of rule-ordering and instead focus on universal linguistics principles. OT can be used to rank universal hierarchies in the different dialects of Penghu Taiwanese. We can construct a constraint ranking in OT for the 5 most commonly attested patterns regarding /g/ and /b/:

- /g/ remains [g]

- /g/ nasalizes to [ŋ]

- /g/ deletes

- /b/ remains [b]

- /b/ nasalizes to [m]

5.1.1   OT Constraints

5 constraints can be used to form an OT analysis:

- Ident-[nas], barring changes in the nasal feature of a sound

- ONSET, requiring syllables to have an onset

- MAX, barring deletion from the input to the output

- *g, barring [g] from appearing in the output

- *b, barring [b] from appearing in the output

5.1.2 OT Rankings

(1) /g/ remains [g]

	gɔ	IDENT-[NAS]	MAX	*g
a.	ŋɔ	*!		
b.	ɔ		*!	
c. ☞	gɔ			*

(2) /g/ nasalizes to [ŋ]

	gɔ	*g	MAX	IDENT-[nas]
a.	gɔ	*!		
b.	ɔ		*!	
c. ☞	ŋɔ			*

(3) /g/ deletes

	gɔ	IDENT-[nas]	*g	MAX
a.	ŋɔ	*!		
b.	gɔ		*!	
c. ☞	ɔ			*

(4) /b/ remains [b]

	bin	IDENT-[nas]	MAX	*b
a.	min	*!		
b.	in		*!	
c. ☞	bin			*

(5)   /b/ nasalizes to [m]

	bin	*b	MAX	IDENT-[nas]
a. bin		*!		
b. in			*!	
c. ☞ min				*

Also, the use of ONSET as a constraint in this hierarchy would account for the occasional appearance of [ji] from underlying /gi/, where a ranking could be created to explain the epenthesis of a glide [j] after [i] to prevent syllables from being onsetless.

5.2   Variation in OT

These 5 attested constraint rankings do not represent individual speakers or dialect areas, but rather are hierarchies that describe rankings specific to attested sound changes. Because this analysis examines a large number of participants across all of Penghu with many thousands of attestations, a large amount of variation can be expected; however, the presence of dialectal variation creates problems for the analysis of a phonological phenomenon. Specifically, what is problematic for a purely phonological analysis is that in many instances, individual speakers show constraint rankings that violate universally attested hierarchies. For example, some speakers show [mwe] as well as [gwe], leading to a constraint ranking where *b outranks *g, a conclusion that has been cross-linguistically attested as false. These contradictions necessitate the construction of multiple, separate constraint rankings to account for social variation.

5.3   Acoustic Analysis in OT

In addition to Kubler's 2003 data, the acoustic data detailed in chapter 3 and analyzed in chapter 4 can also be used in Optimality Theory to give an account of /g/ devoicing.

(6) /g/ devoices to [k]

	gwa	*g	MAX	IDENT-[voi]
a.	gwa	*!		
b.	wa		*!	
c. ☞	kwa			*

(7) /g/ remains [g]

	gwa	IDENT-[voi]	MAX	*g
a.	kwa	*!		
b.	wa		*!	
c. ☞	gwa			*

In 18.52% of the relevant data collected in January of 2009, /g/ changed to [k], meaning *g, MAX outrank IDENT-[voi], and in 19.75% /g/ remained [g], meaning IDENT-[voi], MAX outrank *g. In the remaining 61.73% of the time, /g/ was either deleted or had no burst, meaning it was no longer an oral stop. Clearly, a large amount of variation is present in the data, but this is to be expected; variation on the speaker level can ultimately be random, and we cannot predict which speaker will use which form. However, using the framework of OT, a phonological analysis can be created that accounts for social variation by following different rankings for different attested patterns. Coincidentally, if one considers all 24 possible permutations of rankings of *g, ONSET, MAX, and IDENT-[voi], 10 out of 24 result in a surface form [k], and 10 out of 24 result in a surface form [g], a curious fact when one considers the fact that the percentages of surface [g] and surface [k] attested from underlying /g/ in the acoustic data are almost identical. This symmetry appearing in the acoustic data may reflect the convergence of different possibilities predicted by OT.

Chapter 6
Conclusions

6.1   Conclusions of This Study

The conclusions reached from this study are intriguing and at many times quite surprising. Data gathered by Kubler from a large number of locations in 2003 suggests that the Penghu Islands can currently be categorized into 5 different regions by the degrees to which they undergo the 4 sound changes tracked in this study, and specifically by the degree to which areas delete or change voiced stops. In some areas, such as Baisha, voiced stops rarely underwent a sound change; in others, such as Magong, voiced stops were regularly changed via nasalization or total deletion.

Data gathered specifically for this study at Magong City in Penghu County in 2009 confirms the existence of these sound changes and furthermore shows that different kinds of changes are occurring than previously anticipated. Focusing primarily on [g] and [k], acoustic data not only reveals that deletion is occurring, it also showed that /g/ frequently devoiced to [k], another manifestation of a general conspiracy against voiced stops.

Deletion of /g/ is clearly most prevalent at Magong and in the surrounding area; whether this is due to the actual spread of this change to neighboring areas or to coincidence remains to be ascertained, but is a likely hypothesis given the prestige of Magong compared to its neighbors. This study has also identified the linguistic universals inciting a conspiracy against Taiwanese voiced stops /g/ and /b/ and applied an Optimality Theoretic analysis to this conspiracy, keeping in mind that variation at the individual level is random and probabilistic by nature.

Voice onset times measured in this research showed surprising results. The differences in VOT between devoiced [g] and voiced [g] were statistically significant, and furthermore there was no significant difference between the devoiced /g/ and underlying /k/ on the surface. This result suggests that for some people and in some Taiwanese words, [g] and

[k] are undergoing a merger. However, measuring burst intensity showed an opposite effect; all underlying /g/s that retained bursts on the surface (meaning those that had not undergone deletion) had similar burst intensities, and this burst intensity was statistically different from that of /k/. Thus, stop devoicing in Penghu Taiwanese causes /g/ to become similar to [k] with respect to VOT but not with respect to burst intensity.

Examining the bigger picture, Penghu Taiwanese may be undergoing these sound changes for a variety of reasons. Voiced stops are universally more difficult to articulate than voiceless ones, and the many manifestations of deletion and lenition affecting voiced stops likely represent a phonological conspiracy against them. However, the discovery that Mandarin is widely spoken in Penghu County also raises the possibility that influence from Mandarin may also be another factor; Mandarin lacks voiced stops but has voiceless ones, and the less-prestigious Taiwanese may be changing to become more similar to Mandarin by deleting /g/ or by expressing /g/ as its acoustically similar voiceless counterpart /k/.

The true answer may be any one of these reasons, but is likely a complex combination of both; language change is motivated by many different factors, both societal as well as phonological, and their interaction can bring rise to a multitude of different sound changes. As the OT analysis of chapter 5 points out, language variation on the speaker level is ultimately random and unpredictable, and it is ultimately impossible to pinpoint exactly why sound change may be occurring. This analysis has explained how sound changes are affecting voiced stops (and specifically /g/) in Penghu Taiwanese, and in the future may be able to better understand why.

6.2   Future Avenues of Study

This analysis has opened up many possibilities for future research of Taiwanese. In the spirit of Kubler's 2003 dialectal analysis, another widespread examination of Penghu Taiwanese will certainly be necessary in the future to determine the extent to which a conspiracy against voiced stops has spread. If possible, it would be especially illuminating to gather data from other cities of Taiwan that may have contact with Magong, such as Taipei, Gaoxiong, or Tainan. These cities may also be getting rid of voiced stops, allowing this analysis to have broader implications for the Taiwanese language.

Also, during the course of measuring syllable length, an unpredicted effect on codas was discovered: syllables beginning with the liquid [l] caused assimilation of preceding codas to [l], thus creating geminate liquids. This unexpected phenomenon will likely be a topic of future study in Taiwanese. Additionally, a future study of burst and syllable

intensity in Taiwanese may shed more light on how and why certain syllables seem to be significantly less intense than others.

Finally, it will also be important in the future to conduct new research to determine whether the sound changes affecting /b/ and /g/ will have created any mergers. For example, when [b] nasalizes to [m], will this sound change create a phonemic merger of /b/ and /m/ in Taiwanese, or will they remain underlyingly separate phonemes with subtle phonetic cues to distinguish them? This analysis shows that /g/'s that have devoiced to [k] by increasing their VOT are in fact retaining burst intensity cues of their underlying form [g], suggesting that a full merger is not occurring in spite of a statistically insignificant difference in VOT. A new study of devoiced stops in Penghu Taiwanese may verify or refute this claim.

6.3   Improvements on a Future Study

There are several improvements that could be made in a replication or continuation of this study. Firstly, access to a phonetics laboratory would drastically improve the quality of measured sounds and eliminate some of the uncertainty that arose during the acoustic analysis. Without outside disturbance, phonetic data will be more accurate and participants may feel less distracted.

Secondly, if acoustic data is to be gathered again for a similar study, it may prove more prudent to avoid using a word list to elicit tokens. While a word list is an efficient and easy way of eliciting specific sounds from speakers, the use of a word list may be priming pro-Mandarin biases in responses. Many of the children interviewed misspoke and first spoke a Mandarin word which they then corrected, illustrating potential for the intrusion of Mandarin phonology into Taiwanese words. A more precise way of eliciting Taiwanese words would be to use pictures as visual cues; this method bypasses the problems inherent in using Chinese characters on a word list.

Thirdly, the use of tokens taken from natural, fluent speech would likely provide data that more closely resembles the way that participants actually use Taiwanese in everyday speech. The inability to use tokens taken from natural conversation was primarily due to two factors: the anxiety of participants and the author's lack of fluency in spoken Taiwanese. In future interviews, the exclusive use of Taiwanese would be preferable to instructions given in Mandarin and may help draw out speech that is more natural from participants. This semester, sentences in part 1 of the interview (free conversation) were written down with the help of a native speaker of Taiwanese, Ms. Ying-hua Hsiao. In future analyses, this portion of the data will be used to a much greater extent in the hope that it may provide acoustic data that is more representative of actual Penghu Taiwanese speech.

Appendix

This is the word list used during the January 2009 interviews at Wenguang Elementary School and Wenguang Middle School, Magong. Part 2 of the interview consisted of a word list, and part 3 consisted of sentences.

word list used		*translations*	
你好	學校	hello	school
公園	美國	park (n.)	America
加拿大人	中國人	Canadian (n.)	Chinese person
外國人	台灣	foreigner	Taiwan
語言	日月潭	language	Sun Moon Lake
五	國語	five	Mandarin
國家	我	country	I/me
家己	麻煩	self	bother
明年	今年	next year	this year
去年	九月	last year	September
明天	今天	tomorrow	today

word list used (continued)   *translations*

昨天	問題	yesterday	problem
關係	自己	relation	self
關於	買	about	buy
賣	問	sell	ask
無	五月	not have	May
決定	名	decide	name
兩個人	禮拜	two people	week
台北	便當	Taipei	bento
一萬元	文具店	10,000 yuan	stationary store
原子筆	牛肉麵	ballpoint pen	beef noodles
阮	吳小姐	we (excl.)	Miss Go
黃小姐	誤會	Miss Ng	misunderstand

sentences used   *translations*

－老師，gau早！
Good morning, teacher!

－伊是什麼人？
Who is he?

－我想伊是日本人，不過伊的英語真好。
我不確定。
I think he is Japanese, but his English is very good. I'm not sure.

阮只會曉講台語一點仔，因為阮是加拿大人。
We can only speak a little Taiwanese because we are Canadian.

－你來美國偌久啊？
How long have you been in America?

－我甲我的外國朋友來差不多九個月啊。
阮是舊年五月來的。你呢？
My foreign friend and I have been here for about nine months. We came last year in May. How about you?

－你不知影無？我是台北人，已經住五十八年啊！
You don't know? I am from Taipei; I've already been here for 58 years!

－你愛牛肉麵無？
Do you like beef noodles?

－真愛！你有在賣無？
I do! Do you have them for sale?

－有，在外面！
Yes, outside!

－多謝!
Thank you!

References

Bodman, N. C. 1954. *A linguistic study of the Shih Ming: Initials and consonant clusters*. Cambridge, MA: Harvard University Press.

Boersma, P., and D. Weenink. 1993/2008. Praat: Doing phonetics by computer. v. 5.0.40. http://www.praat.org/

Branner, D. P. 1999. *Problems in comparative Chinese dialectology: The classification of Miin and Hakka*. Berlin: Walter de Gruyter

Kubler, C. 1978. *A dialect study of the Penghu Islands*. Masters thesis, National Taiwan University.

Lee, V. 2002. Penghu residents, experts question value of gambling. *Taipei Times*. 29 July.

Lisker, L. and A.S. Abramson. 1964. A cross-language study of voicing in initial stops: Acoustical measurements. *Word* 20. 384–422.

Norman, J. 1988. *Chinese*. Cambridge University Press.

Pan, H. 1994. Phonetic variants of Taiwanese "voiced" stops. *Journal of the Acoustical Society of America*. 95 (5). 2875.

Prince, A, and P. Smolensky. 1993/2004. *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell Publishers.

Ramsey, S.R. 1989. *The languages of China*. Princeton University Press.

Zheng, L.W., N.Q. Fang, and S.W. Zhao. 1990. *Shenghuo Taiyu beginning Taiwanese*. Zili Baoxi Chubanbu.