

Chapter 3

Sociolinguistic survey design

3.1 Goals of survey and research questions

This chapter describes the design of the sociolinguistic survey fieldwork carried out in December 2009 and February 2011. The first survey was designed under the supervision of Noel Mann and the survey instruments for the second trip were designed under the directions of Nathan and Carey Statezni. The two field work trips were carried out by the author, a few friends from the Payap MA linguistics program and with help from a few local interpreters.

This survey was designed by first identifying three broad goals for the survey and then within each goal, specific research questions were formulated. The fourth goal was added after the first trip to make use of knowledge gained in that trip. The survey instruments were then designed to collect sufficient information to provide at least a partial answer to each research question. In considering the details of this particular survey it is good to bear in mind the context, which is that the Meung Yum community had formed a committee to work on language development and needed data to make good decisions for the whole people group.

The goals and associated research questions are described in the following sections. The goals and related research questions are based on templates for survey designs given in the RAID tool by Nahhas et al (n.d.).

3.1.1 Goal 1: Assess the need for vernacular language development

The first goal of this survey was to determine the need for vernacular language development for Meung Yum speakers, especially by investigating the potential for them to use existing written materials in the LWCs. This broad goal is made specific by the following research questions:

- Research Question 1.1: Do Meung Yum speakers understand Shan adequately³?
- Research Question 1.2: Do Meung Yum speakers have negative attitudes toward the existing written Shan that would keep them from using these materials?
- Research Question 1.3: Do Meung Yum speakers understand Lachid adequately?
- Research Question 1.4: Do Meung Yum speakers have negative attitudes toward the existing written Lachid that would keep them from using these materials?
- Research Question 1.5: Do Meung Yum speakers understand Burmese adequately?
- Research Question 1.6: Do Meung Yum speakers have negative attitudes toward the existing written Burmese that would keep them from using these materials?
- Research Question 1.7: Do Meung Yum speakers understand Chinese adequately?
- Research Question 1.8: Do Meung Yum speakers have negative attitudes toward the existing written Chinese that would keep them from using these materials?
- Research Question 1.9: Do Meung Yum speakers understand Standard Wa adequately?
- Research Question 1.10: Do Meung Yum speakers have negative attitudes toward the existing written Standard Wa that would keep them from using these materials?

3.1.2 Goal 2: Investigate the readiness

If existing materials were found to not be adequate for them, the survey sought to investigate their readiness for vernacular language development, especially by investigating language vitality, population size, interest and unity. The following specific research questions investigate this broad goal.

³ The meaning of 'adequately' in this context is determined by the broader goal to which the research question contributes. A working definition is 'adequately' means do Meung Yum speakers understand [LWC] enough to be able to benefit from a literacy program in [LWC] or would lack of comprehension greatly inhibit the learning of literacy in that language.

Research Question 2.1: Does it appear likely Meung Yum will continue to be spoken by future generations?

Research Question 2.2: What is the approximate population of the Meung Yum and Savaïq?

Research Question 2.3: How interested are Meung Yum people in language development in their own language?

Research Question 2.4: How many Meung Yum speakers have sufficient education to help with language development?

Research Question 2.5: How unified is the Meung Yum community about orthography and language development?

3.1.3 Goal 3: Determine the most suitable variety for Meung Yum language program

If vernacular language development was found to be needed, the third goal was to determine how many and which varieties should be selected for development. The following specific research questions investigate this broad goal.

Research Question 3.1: What Meung Yum varieties are understandable to speakers of other Meung Yum varieties?

Research Question 3.2: What varieties are prestigious and important to the Meung Yum?

Research Question 3.3: What are the types, natures, and extents of interaction between and within the Meung Yum varieties?

Research Question 3.4: What are the linguistic relationships between Meung Yum and other Palaungic varieties?

3.1.4 Goal 4: Could Savaïq be incorporated in a Meung Yum language program?

This goal was added as a result of the survey team hearing on their first fieldwork trip that Savaïq was very closely related to Meung Yum. The final goal was to determine whether Savaïq language development could be achieved by a joint program with Meung Yum. The following specific research questions investigate the broad goal.

Research Question 4.1: Can Savaïq speakers understand Meung Yum?

Research Question 4.2: Do Savaïq speakers have negative attitudes to Meung Yum?

Research Question 4.3: What are the types, natures and extent of interaction between the Meung Yum and Savaïq?

The following sections describe how the site selection was made, which instruments were used, how the subjects were selected and how each instrument was used to answer the research questions.

3.2 Site selection and summary of data collected

Meung Yum/ Savaïq sites were preferred if they were the largest and had a high proportion of Meung Yum/ Savaïq inhabitants. Also there was a desire to visit as many distinct speech varieties as possible, with at least 2-3 sites per variety.

Factors in selecting sites included population size, areas between which there is little and frequent contact, locations reported to speak the 'purest/best' variety, locations reported to speak 'differently', and locations with strong language vitality. Historical centers, trade centers chosen from different village groups⁴ and locations that are remote are also included in selecting survey sites. Consideration was also given to ease of access for the surveyors to visit the sites. The sites the team visited are shown in Table 8. The villages marked with (*) were visited by the researchers for data collection in December 2009 and the rest of the villages were visited in February 2011.

All the villages are located in Kunlong Township. The reason for selecting each village as a survey site as well as the research instruments used in each site is also listed. The first four listed villages were fieldtrips conducted in December 2009. Of those four villages, two of the villages are Buddhist and the other two are Christian. These villages were also selected because they were easier to access and had good relationships with the local survey guides. Another four villages were chosen to survey on January 2011.

⁴ In Myanmar, the village group is an administrative level below the township and above the village.

The two Savaiq villages were chosen in order to study a variety reported to be closely related to Meung Yum. They were chosen from different village groups. One village was big and the other was small. . The villages marked with (*) were visited by the researchers for data collection.

Table 8: Site selection

No.	Village	Village Group	Variety Group	Why selected	Instruments Used
1	Kaung Sar*	Wasoke	Meung Yum	Buddhist village	Wordlist, Knowledgeable Insider (1), Individual SLQ (12), informal interviews
2	Pan Tan*	Wasoke	Meung Yum	Buddhist village	
3	Man Kyu*	Wasoke	Meung Yum	Meung Yum Christian villages	
4	Man Phan*	Wasoke	Meung Yum		
5	Namt Yoke	Taptu	Meung Yum	to pilot test the RTT story; pure Meung Yum village; 40 households	Wordlist, Meung Yum RTT (12), Knowledgeable Insider (1), Individual SLQ (12), Religious Leader Interview (1), Dialect Mapping Tool
6	Kaung Sang	Nawng Mo	Meung Yum	biggest Meung Yum village; pure village; 100+ households	
7	Man Kan	Nam Kyin San	Meung Yum	pure Meung Yum village; 30 households	
8	Man Pein	Taptu	Meung Yum	big village; pure Meung Yum village; 100 households	
9	Man Gyat	Namt Kyin San	Savaiq	big village; 100 households: 90 Savaiq households and 10 Lhaovo households	
10	Thein Tan	Pang Hai	Savaiq	about 20 Savaiq households	

3.3 Survey instruments

The following section describes in detail the instruments used to collect data.

3.3.1 Wordlist collection and procedures

One wordlist was collected in each village visited during the survey trip. When collecting a wordlist for a particular speech variety in a particular village, the

following three screening questions were used to determine whether a person can represent the particular variety.

Three criteria were used to choose suitable subjects in each site. (1) The subject is “from a Meung Yum/ Savaiq⁵ village.” This is defined as growing up in a Meung Yum/ Savaiq village, living in a Meung Yum/ Savaiq village at present, and, if they have lived outside the area, their time elsewhere is not over five years. (2) The subject speaks the elicited variety first and best. (3) Both of the subject’s parents are mother-tongue speakers of the variety and both parents spoke the variety with him/her when he/she was a child. If he/she is not representative of that speech variety as spoken in that village, thus was not asked to participate in the wordlist collection. In each village at least two speakers participated in the elicitation of the wordlist. This had the benefit that they could discuss any words where there was uncertainty.

After eliciting wordlists, one of the participants was asked to pronounce transcription of each word and for recording. The speaker chosen for this task must be free of obvious speech impediments such as missing teeth or a lisp. This person should be the best available representative of the native variety spoken in the village. Wordlists were collected and transcribed by the researchers using the International Phonetic Alphabet (IPA). The wordlists were then recorded using a mini-disc recorder. Wordlist elicitation took place in Burmese.

The items in the wordlist are grouped in broad categories: natures, plants, foods, animals, body, people, home, verbs, numbers, dimensions, appearance, taste, and feelings. Appendix A gives the 454-item wordlist that was used.

3.3.2 Knowledgeable insider sociolinguistic questionnaire

The purpose of the Knowledgeable Insider Questionnaire is to collect information about the community as a whole. This questionnaire was only administered once in each location. The village leader is the person chosen to answer the questions on this questionnaire. The questions are grouped in sections which are described in Table 9. The full questionnaires are given in Appendix B for the questionnaires that were used in 2011 and 2009. The questionnaire used in 2011 is revised from the 2009

⁵ Meung Yum subjects must be from Meung Yum villages; Savaiq subjects must be from Savaiq villages.

questionnaire by adding a few more questions such as interaction and contact among the people from different parts, and about the prestige dialect.

Table 9: Description of sections on Knowledgeable Insider Sociolinguistic Questionnaire

Section Title	Summary of Information Elicited
Subject Demographics	Information about the subject and his/her family
Tribal group information	Information about the history of the Meung Yum/ Savaiq people as a whole or migration history of the village
Village name and population	Information about how outsiders refer to the language used/ inhabitants of the village and the ethnic profile of the village
Group name and village information	Information location of the village, its official and other alternative names
Schools	Information of the language mix in the village school (if any) and whether children go elsewhere for some or all of their schooling
Language maintenance	Information about the fluency on their mother tongue and on the other languages
Language of wider communication	Information about the use of each of the LWCs used in the village
Contact, festivals	Information about ways of interactions among the people group and the outsiders
Intermarriage	Information about extent and convention of intermarriage between Meung Yum/ Savaiq and other ethnic groups
Prestige dialect	Information about dialect variations, central dialect and important location among the people
Orthography and language development	Information about desire, reasons, and preferred script for language development

3.3.3 Religious leader interview questionnaire

The purpose of this questionnaire is to collect information about the religious language used among the community as a whole. This questionnaire was only administered once in each location in the 2011 fieldtrip. A monk in each location

was chosen to answer the questions on this questionnaire. There are no Meung Yum/ Savaiq Christian village in the selected sites in 2011. The questions are grouped in sections which are described in Table 10. The full questionnaire is given in Appendix B.

Table 10: Description of sections on Religious Leader Interview Questionnaire

Section Title	Summary of Information Elicited
Subject Demographics	Information about the subject and his/her family
Village monastery information	Information about the history, number of monks, novices in the monastery
Festivals	Information about kinds of religious festivals, ethnic groups and language use in the those festivals
Language use	Information about language use in religious activities
interest in written language	Information about language use in religious writings
Orthography and Language Development	Information about desire, reasons, and preferred script for language development

3.3.4 Individual sociolinguistic questionnaire

The purpose of this questionnaire is to gather information from individual relating to the research questions under investigation in the survey. This questionnaire was administered to twelve subjects in each location. The subjects are chosen according to the criteria given in Section 3.4. The questions are grouped in sections which are described in Table 11. The full questionnaires are given in Appendix B.

Table 11: Descriptions of sections on Individual Sociolinguistics Questionnaire

Section Title	Summary of Information Elicited
Subject Demographics	Information about the subject and his/her family
Contact	Information about ways of interactions among the people group and the outsiders
Attitudes toward written LWCs	Information about proficiency and attitudes on the written LWCs
Interest in Literacy	Information about interest in reading in the LWCs
Bilingual Proficiency	Information about the subjects' fluency in speaking LWCs
Children language use and language maintenance	Information about children language use in the village and attitudes on their children acquisition of other languages
Domain of Language use	Information of language(s) use in various domains
Ethno-linguistic Identity	Information about which ethnic group the people think of themselves

3.3.5 Recorded text test

The team used a modified form of Recorded Text Test (RTT) to discover how well the people in the various villages visited understand the Namt Yoke variety of Meung Yum. To construct the RTT, a story was elicited in the Namt Yoke variety. This is called the "Test Story". This RTT was then pilot tested with Meung Yum subjects in Namt Yoke village, using the Test Story, just as it was used in other villages, except that this Pilot Test contained 25-30 questions on the Test Story. A short Practice Story was created and played first to every subject to help them become familiar with the requirement of an RTT, listening and then answering questions about the text. The questions were asked orally in Shan or Burmese, depending on what was more convenient for the subject. The full RTT story with questions used are given in Appendix D. The questions that the Namt Yoke villagers could not answer correctly or any other problematic questions were eliminated, leaving 12 questions for RTT testing in other areas.

During the survey trip, the team tested 12 subjects and administered the Individual SLQ to them in each village, using the Practice Story and Test Story just as was done during the pilot testing. However, it was not convenient for old men to answer the

questions directly; instead, the team asked them to retell the phrase they heard by translating it into Shan. If the subject did not mention the portion asked about in the question, the team then asked them directly. Thus, the retelling method was mainly used and only a few questions were asked to the old male subjects. The procedure to develop this RTT is described in Table 12.

Table 12: RTT preparation steps

Day	RTT steps to develop the Meung Yum story
Day 1-2	Elicit and record some personal experience stories in the Namt Yoke variety of Meung Yum and choose one
	Break the story into phrases (chunks). This gives each phrase as a separate recorded track.
	Transcribe and translate the story into Burmese
	Make up questions for each phrase (a total of 20-25); translate them into Burmese and Shan
Day 3	Make a track table – this is a numbered list of the recorded phrases
	Construct the pilot test disc with the Meung Yum warm-up story
	Prepare pilot test questionnaires
	Prepare pilot test answer sheets
Day 4-5	Pilot test with 12 subjects
	Write down all the answers and decide what answers to consider correct for each question
	Score the subjects, i.e., allocate scores to each subject based on their answers to each question.
	Choose the 12 best questions for the RTT. This omits questions that subjects found unclear or often got wrong.
	Construct the final RTT disc. This has the whole story as the first track followed by tracks for each individual phrase.
Day 6	Update the track table to show which tracks have questions associated with them and which do not.
	Prepare RTT questionnaires – these use the reduced set of 12 questions
	Prepare RTT answer sheets

Then, in each village, 12 subjects were tested on this Meung Yum RTT along with the Individual SLQ. The Meung Yum RTT was elicited in Yangon before this survey trip and was pilot-tested in Namt Yoke village during the survey trip. Twelve subjects were used in pilot-testing the Meung Yum RTT in Namt Yoke village.

1050032433

3.3.6 Dialect mapping tool

The mapping tool is to assist a group of speakers of a specific variety in discussing what they know about the varieties of their language. Another purpose is to encourage them to think about which varieties of their language could share a single set of written or oral materials. The tool is participatory in that it uses simple materials to create a visual display of speech varieties and groups these varieties according to various criteria.

Firstly a group of village members is gathered who are knowledgeable about other villages and their speech varieties. The group typically numbered around 12-14 people. The group is asked to name all the speech varieties⁶ that are related to their own speech variety. The survey team then writes these names down on pieces of paper and lays them out on the floor so that the entire group can see. The group is then asked to group the names, firstly according to how much they understand of each variety; then according to how much contact they have with each named variety; then how they interact with speakers of the other varieties. For example, can they use their own speech variety to communicate, and if so how much do they understand. For a more detailed description of the methodology for this tool, see Appendices Appendix B for the English and Burmese versions of the Dialect Mapping Tool steps.

3.3.7 Observation notes

The survey team frequently noted any observations that were relevant to the research questions, such as what languages they heard in use in the villages when they were not conducting interviews.

3.4 Subject selection

There are two aspects to selecting subjects for the sociolinguistic questionnaires: screening subjects to ensure they are members of the community in question who speak the specified variety and correctly sampling individuals in each location. These are described in the following subsections.

⁶ In this context, the word rok [rɔk] 'Palaungic people group' was used as it was easier for the people to understand.

3.4.1 Screening criteria

When administering individual SLQs, the target population for each variety in a village consists of people from that village who are mother-tongue speakers of that variety. This is formalized by using the following criteria for subjects. If a subject did not meet all three criteria for any one variety, then he/she is not part of the target population for that variety and, thus, was not tested.

1. The subject is "from a Meung Yum⁷ village." This is defined as growing up in a Meung Yum village, living in a Meung Yum village at present, and, if they have lived outside the Meung Yum area, their time elsewhere is not over five years.
2. The subject speaks the mother tongue either first or best.
3. The subject has at least one parent who is a mother-tongue speaker of the variety and that parent spoke the variety with him/her when he/she was a child.

3.4.2 Sampling

Quota sampling was used, with age and gender as the quota stratification variables. In all cases, the three age categories: age 15-30, age 31-45, and above 46. This results in six strata with the sample sizes desired in each stratum shown in Table 13. The stratification variables were chosen because it was thought that the answers to the research questions might differ by gender and age and so that no segment of the population would be excluded. The sample size was limited to only 12 people per village due to the time constraints of the research team members. The planned sample is shown in Table 13.

Table 13: Planned sample size by age and gender

Gender	Age			Total
	15-30	31-45	46+	
Female	2	2	2	6
Male	2	2	2	6
Total	4	4	4	12

The team also tried to make sure that the subjects they interviewed included a wide range of educational backgrounds (none, basic, and high). Also, an effort was made to sample in all parts of the village and to sample at a time of day when the villagers

⁷ Savaiq speakers must come from a Savaiq village

are normally at home (rather than out in their fields, for example). These ways were used to avoid making the results biased by the sample selection.

3.5 Methods of analysis

In this section, methods of analysis are described.

3.5.1 Phonetic transcription

IPA transcriptions of the 454-item wordlists were entered into Excel and double-checked for accuracy using the recordings. Detailed phonological analysis was not done. Charts of phonetic segments found in each variety were compiled so that they could be compared.

3.5.2 Lexicostatistics

This section describes the lexicostatistic approach used for computing the lexical similarity percentages. This lexicostatistic approach is an approximation of the percentage of cognates shared by two or more speech varieties. In making lexical comparisons, only core vocabularies are used. Mann (2004) identified a list of 118 core items used in Southeast Asia. There are 105 of those items on the 454-wordlists so the lexical similarity percentages in this thesis are based on these 105 items.

Waic languages are from the eastern Palaungic language family. Palaungic languages have monosyllabic and polysyllabic roots. Polysyllabic forms may include non-root syllables. These non-root syllables are supplemental information. For a lexical comparison attempting to approximate cognate percentages between speech varieties, it is often misleading to include non-root syllables in the comparison. Thus, in this study, only the root syllables are considered; non-root syllables are ignored. For example, consider the data in Table 14 [non-root syllables are shown in brackets]

Table 14: Data with root and non-root syllables

Variety	Village	'ten'	'leaf'	'branch'
Meung Yum	Man Pein	kau	ɿaʔ [k ^h auʔ]	kak [k ^h auʔ]
Savaiq	Man Gyat	kau	ɿaʔ [k ^h auʔ]	kak [k ^h auʔ]
Muak Sa'aak	Wan Fai	kul	la: ² [su ²]	tan ¹ [su ²]

In Table 14, the word 'ten' is a monosyllable; thus no further analysis is required and these forms can be directly compared. Looking at the words 'leaf' and 'branch', the varieties has the same morpheme [k^hau?]/ [su²] meaning 'tree'. These morphemes provide supplemental semantic information which is not necessary to the core meaning of the root syllable and are not analyzed. Applying these basic steps, the data can be clarified by eliminating minor and supplemental syllables, as shown in Table 15.

Table 15: Data with root forms only

Variety	Village	'ten'	'leaf'	'branch'
Meung Yum	Man Pein	kau	ɭaʔ	kak
Savaiq	Man Gyat	kau	ɭaʔ	kak
Muak Sa'aak	Wan Fai	kul	la	tan

Phonetically-similar consonants and vowels (adapted from Burquest 2001) are shown in Figure 4, Figure 5, Figure 6, and Figure 7. Initial and medial consonants which are phonetically similar in at least two features are shown by connected lines in Figure 4. For example, [p] and [ʔ] share two phonetic features, plosive and voiceless, shown by a connecting line in the figure.

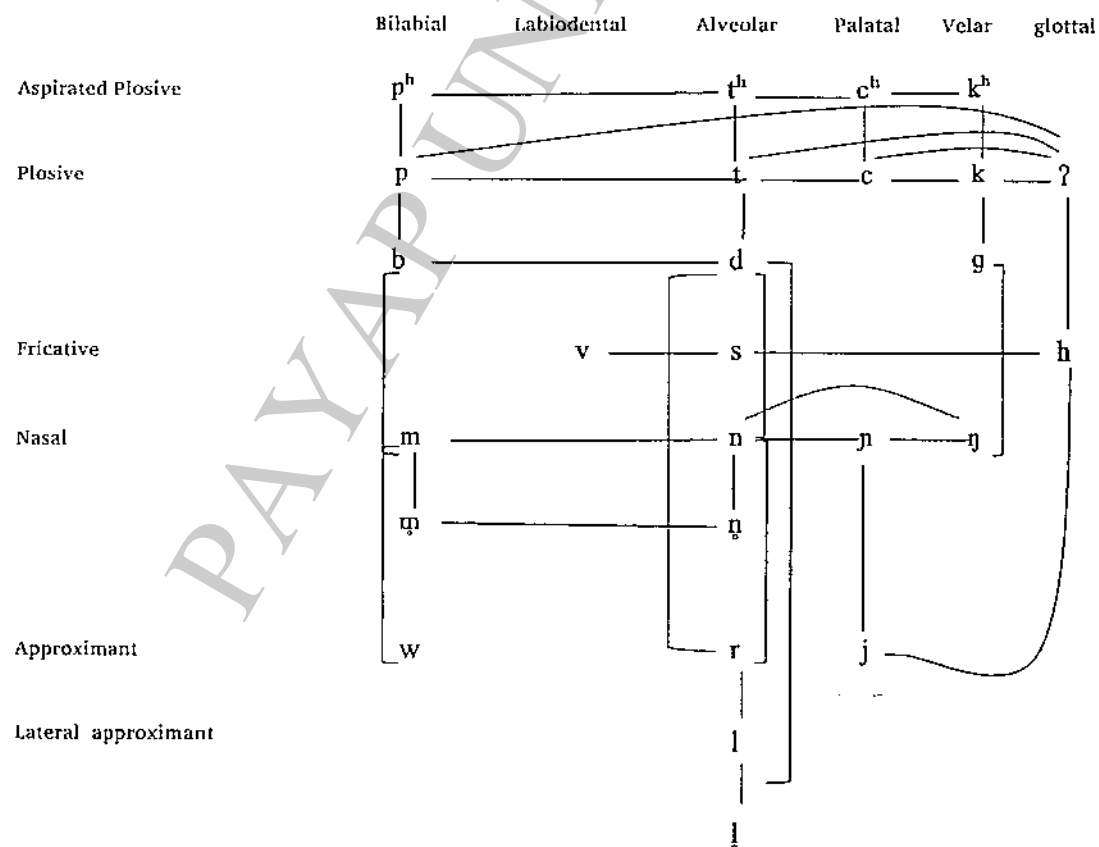


Figure 4: Initial and medial consonants which share at least two features

Final consonants which are phonetically similar in at least two features are shown in Figure 5.

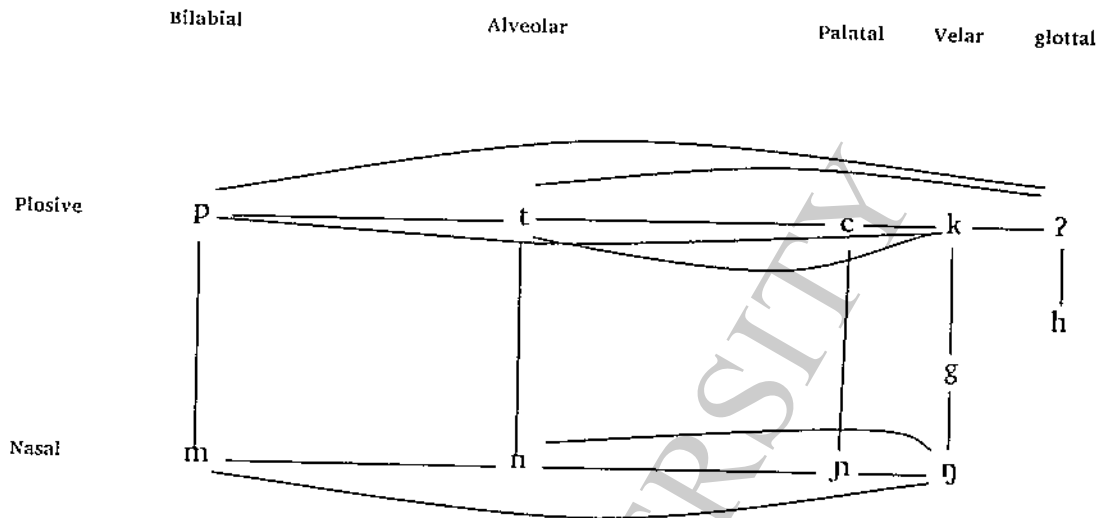


Figure 5: Final consonants which share at least two features

Phonetically-similar vowels are shown by connecting lines in Figure 6. Connected vowels differ by no more than two features. For example [i] and [e] differ in only highness whereas [i] and [u] differ in backness and roundness.

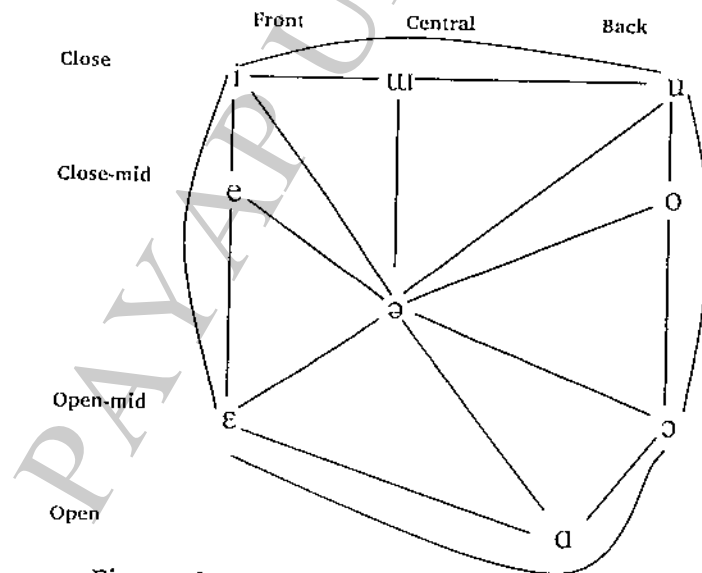


Figure 6: Vowels differing by one feature

Simple vowels with phonetically-similar diphthongs are displayed in Figure 7.

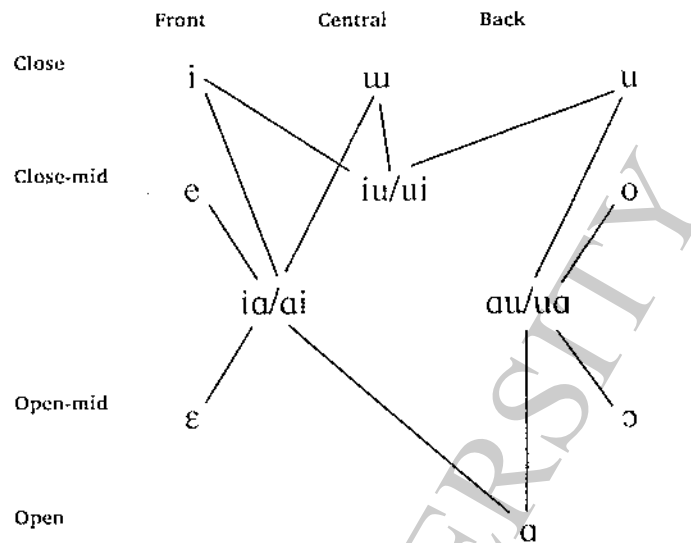


Figure 7: Simple vowels with phonetically-similar diphthongs

The criteria in Table 16 are used to judge whether phonological segments of a word from two varieties are phonetically similar or not.

Table 16: Criteria for segment comparison (Blair 1990: 31-32)

- Category A:
- (a) Identical consonants
 - (b) Identical vowels or phonetically-similar (connected) vowels
 - (c) Phonetically-similar (connected) consonants that appear in a total of 3 or more word pairs (over the whole wordlist)
- Category B:
- (a) Phonetically-similar (connected) consonants in fewer than 3 word pairs
 - (b) Vowels that are not connected
 - (c) [r]/[l]/[x] and nothing after another consonant
- Category C:
- (a) Non-phonetically-similar (not connected) consonants
 - (b) A correspondence with nothing in fewer than 3 word pairs (over the whole wordlist)
- Ignore:
- (a) The vowel [ə] between consonants
 - (b) A correspondence of a consonant or a vowel with nothing in 3 or more word pairs (over the whole wordlist)
 - (c) A correspondence between [ʔ]/[h] and nothing for final consonants
 - (d) Suprasegmentals such as tones, breathiness

When comparing diphthongs and vowels, for example [ia] and [ie] Figure 7 was consulted first. For diphthongs not included in that figure, the diphthong should be connected to all the sounds in between the two sounds in the diphthong. Thus, [uɔ] is connected to [u], [o] and [ɔ]. Segments were considered phonetically similar if either (1) they are shown connected on the charts of phonetically-similar segments or (2) they appeared in three or more word pairs within the 454-item wordlist.

In this example, for the word for “ten”, Meung Yum and Savaiq have identical initial consonants. Thus, the correspondence between the first segments of the word pair are assigned to Category A; sub-point (a). They also have two identical vowels, so these vowels are assigned to Category A; sub-point (b) and Category A; sub-point (b) respectively. From here on the word sub-point will be omitted.

For Muak Sa-ak and Savaiq, the correspondence between the segments [k] - [k] or [a]-[a] are identical with Muak Sa-ak, thus they are also assigned to Category A (a) and A (b) respectively. But the last segment [l] is a correspondence with nothing in 3 or more word pairs (over the whole wordlist). Thus, it is assigned to be ignored (x).

The results of applying the criteria in Table 16 to the data in Table 15 are shown in Table 17.

Table 17: Application of similarity for lexical item ‘ten’

Comparison of ‘ten’	Meung Yum	Savaiq	Muak Sa-ak	Categorization
Meung Yum- Savaiq	kau	kau		A(a)-A(b)-A(b)
Meung Yum-Muak Sa-ak	kau		kul	A(a)-A(b)-(x)
Savaiq-Muak Sa-ak		kau	kul	A(a)-A(b)-(x)

It should be noted that by ignoring phonations, tones, and registers the lexical similarity percentages might be raised. However, Phung (p.c) reported that only around ten percent of Meung Yum words have breathy vowels.

Once the categories have been assigned for all of the phones, Table 18 is used to determine whether the words thus compared are lexically similar or not. The determination is based on the number of phones and certain conditions the word forms must meet in order to be considered lexically similar. The comparisons must match one of the specifications listed in the matrix in order to be considered lexically similar.

Table 18: Acceptable category combinations for lexical similarity

No. of Phones	=	Category		
		A	B	C
1	=	1	0	0
2	=	2	0	0
2	=	1	1	0
3	=	3	0	0
3	=	2	1	0
4	=	4	0	0
4	=	3	1	0
4	=	3	0	1
4	=	2	2	0
4	=	2	1	1
5	=	5	0	0
5	=	4	1	0
5	=	4	0	1
5	=	3	2	0
5	=	3	1	1

No. of Phones	=	Category		
		A	B	C
6	=	6	0	0
6	=	5	1	0
6	=	5	0	1
6	=	4	2	0
6	=	4	1	1
6	=	3	3	0
6	=	3	2	1
7	=	7	0	0
7	=	6	1	0
7	=	6	0	1
7	=	5	2	0
7	=	5	1	1
7	=	4	3	0
7	=	4	2	1

To use Table 18, first, the number of phones in each word is counted. If one word is longer than the other, the number of phones in the longer word is used. Then, the number of phones in each category (Category A, Category B, Category C) is counted. For example, for the word 'ten' and 'leaf' in the comparison between Meung Yum and Savaiq in Table 17, they have three phones, all the three phones are Category A (Phones 3, Category A=3, Category B=0, Category C=0). Since this case is listed in Table 18, the words compared are lexically similar.

Another example is the word 'branch' in the comparison between Savaiq and Muak Sa-ak in Table 15. Of the three phones, two are in Category C (a) and the other one is in Category A (b). This case is not listed in Table 18 (Phones=3, Category A=1, Category B=0, Category C=2), and therefore these two words are not lexically similar. Table 19 gives the results for each of the cases in Table 15.

Table 19: Lexical similarity analysis

Comparison	'ten'	'leaf'	'branch'
Meung Yum- Savaiq	similar	similar	similar
Meung Yum-Muak Sa-ak	similar	similar	not similar
Savaiq-Muak Sa-ak	similar	similar	not similar

From these comparisons, the percentage of lexical similarity can be computed and a matrix can be generated that depicts the lexical similarity relationship between speech varieties. The percentage presented in Section 6.1.2 was based on comparisons of approximately 100 words (see Appendix A), not just the words shown in this example.

3.5.3 Recorded text test

The procedure for RTT followed Nahhas (2007: 68). After administering the test, the answer keys were scored, 0 or 1. Then the RTT scores for each subject are computed for the average and standard deviation. Do this for all the subjects combined. Also do it by the stratification used in the sampling (i.e. by age and gender). Then, the scores are lined up from the subjects with higher to lower scores, to evaluate the effect of contact on comprehension as shown in Appendix D. The average and standard deviation is calculated by computer software; Microsoft Excel.