AN ACOUSTIC, HISTORICAL, AND DEVELOPMENTAL

ANALYSIS OF SARIKOL TAJIK DIPHTHONGS

by

PAMELA S. ARLUND

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

December 2006

ACKNOWLEDGEMENTS

ABSTRACT

AN ACOUSTIC, HISTORICAL, AND DEVELOPMENTAL

ANALYSIS OF SARIKOL TAJIK DIPHTHONGS

Publication No. _____

Pamela S. Arlund, PhD.

The University of Texas at Arlington, 2006

Supervising Professor: Jerold A. Edmondson

Sarikol Tajik has been reported to be an unusual language, containing up to twelve diphthongs and being the only language in the Gorno-Badakhshan family to not contain a short and long vowel distinction among monophthongs. However, the basis of such claims is not clear and could be accounted for by any of several factors. For example, different researchers have utilized different definitions of diphthongs, some researchers have (perhaps unknowingly) studied different dialects at different times, while others simply failed to account for the variation that is often found in languages in change. This dissertation reevaluates these claims and the potential reasons behind such

claims by examining field recordings from Sarikoli speakers in three locations, utilizing prototype theory, time analysis, and the nature of diphthongs.

Spectrographic analysis reveals that, contrary to previous reports, Sarikoli contains both long and short vowels and three diphthongs. After examining the definition of a diphthong, this dissertation takes the approach that diphthongs can be more or less prototypical instantiations, showing that some Sarikoli long monophthongs are very diphthong like, particularly in their release pattern. Spectrographic and statistical analysis also revealed dialectal differences in the instantiations of the diphthongs across dialects, with the eastern most dialect (Burungsali) containing lower, more central vowels than the eastern most dialect (Tashkorgani).

After examining the data, a prediction is made about the developmental pattern among the Pamiri mountain languages and the three dialects of Sarikoli. Implicational scales show that the Burungsal dialect of Sarikoli is the most advanced in a process of diphthongization of long monophthongs. This change is traced historically in relation to Avestan and across four current Pamiri mountain languages. All things being equal, it is predicted that all of these languages will continue in a pattern of developing diphthongs from long monophthongs.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

xi

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 DIPHTHONGS

The exact nature of diphthongs is a ripe field of controversy in linguistics. It seems that for as long as linguistics has existed, linguists have long disagreed over, debated about, and deliberated upon what a diphthong is (or is not). This disagreement has arisen in part because different theoretical assumptions have lead to different field methods. For example, one linguist might hear a particular segment and transcribe it as a complex vowel with movement, while another might hear the same segment and transcribe it as a vowel followed by a semi-vowel, while yet another might hear two vowels in two different syllables. These varied theoretical assumptions employing different transcription conventions have in turn made it difficult to obtain consistent phonetic information about diphthongs (or even vowels in general) across languages. This lack of information is even more pronounced when it comes to understudied languages.

This controversy about diphthongs has been reflected in the linguistics literature, in which there have been two primary competing viewpoints of what a diphthong is. These two viewpoints can be seen as two endpoints on a continuum with various positions lying in between. On the one end is the duality (sequence) view that diphthongs are two discrete simpler vowels with a transition (called a glide) in between

(e.g. Bond 1978, Gottfried 1993, Jha 1985, Lindau 1990a). That is, those who advocate the sequence view stipulate that the two vowels that comprise a diphthong have simple vowel counterparts. These simple vowels are linked by a gliding segment, i.e. a segment that generally displays rapid formant change from the onset steady state to the offset steady state. If this were the end of the definition, it would already be complex enough, but even among linguists advocating this definition, there is agreement to disagree about minor aspects of diphthongs. That is, even though linguists from this point of view agree in their basic understanding of a diphthong, they still argue among each other about specific diphthongs in specific languages. They finally agreed to disagree by stating that some aspects of diphthongs must be language specific and cannot be stated categorically across languages (Lindau 1990a).

On the other end of the diphthong continuum is the unity (single vowel) view, which states that diphthongs are simply a vowel that has continuously changing formant qualities (Kent 2002, Ladefoged 1996). Proponents of this definition argue that the two endpoints of a diphthong need not coincide with any of the simpler vowels in the language. Furthermore, there need not necessarily be two steady state points with a glide in-between in all diphthongs. All that is necessary to qualify as a diphthong under this definition is movement. It is perhaps worth noting that linguists who seem to hold this view are those with more of an acoustic phonetics perspective. Also, while these linguists have extensive field experience and knowledge, the specific works in which they propose these definitions are more theoretical and do not hold any specific

2

language in mind. Thus, it seems that purely theoretical works are more willing to accept more variation in their definition of a diphthong than those linguists who have approached the problem of diphthongs from the perspective of one or a small group of specific languages. That is to say, those with a more theoretical perspective seem to hold a less rigid view of what a diphthong is.

This difference between less and more theoretical viewpoints is worth acknowledging because this dissertation examines diphthongs from the viewpoint of a specific language: Sarikol Tajik, an Indo-Iranian language of China. Previous descriptions of the language have been in widespread disagreement on how many diphthongs Sarikoli might have or, indeed, if it has them at all. To address this issue, the present dissertation analyzes the vowels of Sarikol Tajik from a phonetic perspective, using spectrograms, statistics, and historical information to determine the exact nature and structure of each of the vowels. Ultimately, this dissertation shows how Sarikol Tajik is much more like the other members of its immediate language family than previously supposed and examines some of the variation currently exhibited among individual speakers of the language.

1.2 SARIKOL TAJIK

Since this dissertation will be concerned with the study of one little known language, it is worth taking time now to provide some context about the language under investigation.

Sarikol Tajik is an Indo-Iranian language of western China. It is the only Indo-European language spoken exclusively in modern China. (Russian, the other Indo-European language spoken in China, is also an official minority language of China, but it is spoken in many other countries as well.) Sarikol Tajik is located at the extreme western edge of the Xinjiang Uighur Autonomous Region, where China borders Tajikistan, Afghanistan, and Pakistan/India (Kashmir). This can be seen in Figure 1.1, where the Chinese Tajik area has been outlined in black. China is estimated to have 41,000 Tajiks (Xinjiang Statistics Bureau 2004). However, this number, like most Chinese minority statistics, can be misleading. The Chinese term 'Tajik' includes three distinct sub-groups: Sarikolis, Wakhans, and Tor Tajiks. Sarikolis are the predominant group, numbering perhaps around 25,000. Their language is used as the language of wider communication among Tajiks of China. The Wakhan Tajik live in China, Afghanistan, Pakistan, and Tajikistan. They have their own language, which is grammatically nearly identical to Sarikoli but lexically divergent. The Tor Tajiks have Tajik customs but speak a variety of a Turkic language that is intermediate between Uighur and Uzbek.

As confusing as the Chinese term 'Tajik' is, the English term 'Tajik' can be equally confusing to those not familiar with the area. 'Tajik' is used traditionally to denote a Persian people group of Central Asia - as opposed to a Turkic one. This means that there are many different kinds of 'Tajik' people. When the distinction between them is not important, they usually will refer to themselves simply as 'Tajik'. Only

when the distinction is important will Tajiks usually refer to themselves by saying what specific kind of Tajik they are. Some might recall that this was also the case with Turkic peoples until the solidification of Soviet control over Central Asia. For example, Turkic peoples were previously referred to by terms such as 'a Turki of Kazak origin' or a 'Kazak Turki,' etc. Tajiks have retained this older custom of naming themselves. Throughout this dissertation, the Sarikol Tajiks are referred to simply as 'Tajiks'. If any other kind of Tajik is referred to, their more specific title is used.

Figure 1.1 Map of China showing location of Tajiks in black
(Chinatour.com: 2003)

1.3 STUDIES OF SARIKOLI

In fact, the political situation in Central Asia has affected more than just the naming standards of people groups in the area. In particular, studies of Sarikoli have been strongly influenced by the political situation in Central Asia in general and China

5

in particular. Certainly the question of political ownership of the Tajik areas was not totally settled until the first half of the 1900s. Russia, England, and China each vied for control of Central Asia in what has come to be known as The Great Game. (For a detailed account of the history of this time period see Peter Hopkirk's *The Great Game*.) As the representatives of the countries participating in The Great Game traveled through Central Asia, they quickly documented languages they encountered and then moved on to the next assignment. In fact, this is roughly how the first English account of Sarikoli, a study made by the English agent and linguist R.B. Shaw in 1876, came into being. When the dust settled and the Russians had won Tajikistan, the British had won India, and China had won the eastern side of the Pamirs, the mountain Tajik languages (also known as Pamiri languages) were divided into three different spheres of influence. Among these, Sarikoli became the most isolated and most understudied of the languages because it was confined to a remote border area of China.

Over time, Sarikoli became viewed by Indo-Iranianists as being a language that was interesting but, practically speaking, difficult to study. For many, the language was simply too remote geographically from their Central Asian perspective. To this day, one wishing to study the language has to first arrive in Beijing, take two more plane flights from there, and then ride in a car over a poor quality road for another 6 - 10 hours. Researchers who were not daunted by the journey to the area were often put off by the thought of communicating in Chinese. Typically, linguists attracted to Indo-Iranian languages as their primary interest normally do not require any Chinese. Others simply

could not cope with the amount of red tape introduced by the Chinese government. At present, anyone wishing to do any amount of detailed study on the language is still met with suspicion and will undoubtedly have to endure many police visits.

As a consequence of these geographic and political barriers, research on Pamiri mountain languages became fragmented. Fragmentation has, in turn, produced widely conflicting reports about languages of the area. For example, Hattaway (2000) claims that Sarikoli is mutually intelligible with Shughni while Grimes (1996) explicitly states that it is not. In contrast, Edelman (1980) and Sims-Williams (1996) claim that all the Shughni group languages are mutually intelligible and have only been divided into separate languages for socio-linguistic reasons. These reports are further muddled by the fact that none of these researchers explained how they reached their conclusions. In fact, none of these researchers ever studied Sarikoli themselves; they simply relied on publications written by other field linguists.

The main field study upon which most researchers have relied is the work of T.N. Pakhalina (1966, 1971). Her work is by far the most extensive and reliable work to date on Sarikol Tajik. She published both a short dictionary and a comprehensive grammar of Sarikoli. Although her research was published in the 1960s and 1970s, most of the actual field work that led to the writing of the books was done in the 1950s. Recall that the Soviet Union and China had a parting of ways in the 1960s, which led to no other Soviet researcher ever being given the kind of freedom she had been allotted

while doing her field research. Her work was used as the cornerstone for all following Soviet writings that included Sarikol Tajik.

The main Chinese researcher of Sarikol Tajik has been Gao Erqiang (1963, 1985). He contributed the volume *Tajikeyu Jianzhi* to the Chinese series of books that explores each of the minority languages of China. In that volume, he gives a basic overview of the phonology, morphology, and grammar of Sarikoli and also includes a comparison of Sarikol and Wakhan Tajik. This inclusion of Wakhan Tajik is in keeping with the Chinese point of view that Wakhan and Sarikoli are merely dialects of the same language. However, it is well known that Chinese scholars and government officials count dialects and languages differently than most of the world. In truth Wakhan and Sarikoli are mutually unintelligible languages belonging to people who are ethnically similar but retain their individual group identity. As such, most non-Chinese scholars would consider them to be different languages. The book contains some other problems as well. When asked, most Tajik scholars consider the book to be adequate but also point out that it contains a number of errors in it and sometimes contains words that no one seems to have heard of. As a result, this work must be approached cautiously and can only be used as a guideline for further research.

There is only one Tajik scholar that has published anything at all on Sarikol Tajik. Shirin Corban (1994) is a Tajik professor at Xinjiang University, the province's largest academic institution. In his book on the Tajik people, he includes a small section on the language, but an analysis of the language is not his area of interest. Indeed, much

of his book seems to have been copied word for word from Gao. He is much more interested in the customs and practices of the people rather than in linguistic description. Therefore, the section on language is both short and provides little new information.

The only non-Chinese to study Tajik in recent history was a team of two Americans, Eric and Ellen Peters, who taught English in Tashkorgan for a short period of time in the late 1990s. After leaving the area, they continued to study Tajik at Xinjiang University and compiled a dictionary as they studied. They both have some minimal training in linguistics and are skilled language learners, but never wrote a grammar of the language. Nevertheless, their lexicon seems to be of excellent quality, though it has never been published. In addition, their phonetic understanding of the language appears to be good. Not being linguists, however, they never applied any instrumental analysis to their work.

In summary, these four authors, a Russian, a Han Chinese, a Sarikol Tajik, and two Americans are the only ones in the past hundred years to have based their studies on their own field data. While Sarikoli is often mentioned or included in other works (particularly articles and publications by Soviet scholars), those other authors used one of these four studies as their primary sources. As can be seen by this brief survey, field data has been both scarce and has often been flawed or merely incomplete. Therefore, if the data exists and is complete, it is still scarce and often quite old now. In fact, even the most reliable information on the language, Pakhalina's work, is based on information that is already half a century old.

1.4 THE PAMIRI MOUNTAIN LANGUAGE FAMILY

Given this lack of information, it is not surprising to find more contradictions when trying to establish Sarikoli's relationship to other languages in its immediate vicinity. However, to understand the present study of Sarikoli diphthongs, it is helpful to clearly grasp the relationship between it and the other Pamiri mountain languages. To that end, the most widely accepted division of the so-called Gorno-Badakhshan group is given in Figure 1.2 (Bashiri 1997). (The Sarikol language is spelled here as Sarykol.) Not all sources, however, agree with this representation either. Sims-Williams says, '…it does not seem possible to regard the Eastern Iranian group as a whole - even excluding Parachi and Ormuri - as a genetic grouping.…It is therefore more plausible to conceive of Eastern Iranian as a 'Sprachbund' or areal grouping of languages. (1996:659)'

Indo-European

Indo-Iranian

Iranian

Western Iranian   Eastern Iranian

Yaghnobi   Tajiki   Gorno-Badakhshan

northwestern   central   southern   southeastern

Ishkashemi   Shughno-Rushan   Wakhi   Yazgulemi

Rushan   Shughni

Khufi   Bartang   Oroshor   Sarykol   Bajui   Shahdar
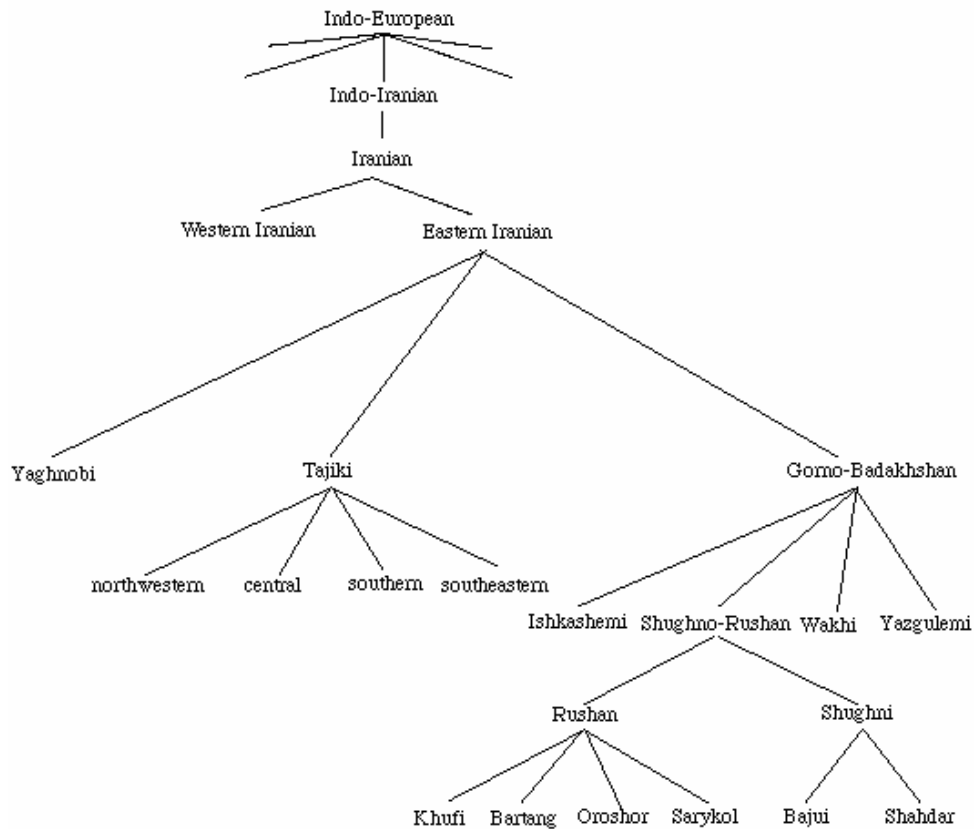
Figure 1.2 A representation of the relationships among Eastern Iranian languages

Perhaps the most reliable source to turn to for languages of the area would be the Soviets. The most comprehensive work done on this group of languages was carried out under the leadership of I.I. Zarubin of the Soviet Academy of the Social Sciences in the 1950s and 1960s. As previously mentioned, the main field researcher for Sarikol Tajik was T.N. Pakhalina. Pakhalina's work and the work of her colleagues in other Pamiri languages were then pooled together and analyzed by V.S. Sokolova (1967).

Sokolova then conducted a study on the genetic relationships of the Shugnan languages and Yazghulami (the language in which she had done her own field work). The Russian data is by far the most comprehensive work ever pooled on Tajik languages. Nevertheless, the data for Sarikoli is often incomplete, simply left blank. In addition, the work is now dated, indicating that more study is needed of Sarikoli specifically and all Pamiri mountain languages in general. The researchers also had little access to the equipment and theories that have appeared in the last 50 years. For example, they had little access to computers (though they did do some acoustic work on languages other than Sarikoli). Likewise, significant advances in historical studies and time-based linguistic analyses have been made in the intervening years, theories which would have undoubtedly have aided their analyses greatly had they been available.

The only other major work to study the relationships among the Shughnan group languages is Georg Morgenstierne's (1974) etymological dictionary. He used the works of the Russian researchers as his data, redacting them into one dictionary and making historical hypotheses about each entry. He does not make any claims about the languages' relationship to each other but confines his attention to each word. He never looked beyond the scope of each word to draw any conclusions about the overall relationships between the languages. He is hampered by the same lack of data with which Sokolova was and has many alternate entries or blanks for Sarikoli.

1.5 SARIKOLI VOWELS

Given the state of research into Sarikoli Tajik, it should not be surprising that reports on the vowel system of Sarikol Tajik are particularly contradictory. There is general (though not absolute) agreement on the number of simple vowels but widespread disagreement on the number of diphthongs. Schwarz (1984) quoting Gao (1963) claims that Sarikoli has nine diphthongs. Later, Gao (1985) modified his earlier claim, positing that Sarikoli has eight diphthongs. Meanwhile, Shirin Corban (1994) claims that Sarikoli only has two diphthongs. Similarly, Pakhalina (1971) and Peters (1996) both transcribed two diphthongs. All other articles that have appeared in the west have been based on Pakhalina's work, and have therefore followed her transcription conventions.

Of the four studies that provide an overview of Sarikoli based on their own field research, each has made different claims about Sarikoli vowels. These claims are summarized in Table 1.1. Some of the differences in transcription probably result merely from different interpretations of the same sound. Other differences may be based on substance. The two western sources, Pakhalina (1971) and Peters and Peters (1996) have both transcribed more vowels than the Chinese scholars. The question arises: What is the exact nature of these sequences?

Table 1.1 Cross-study comparison of Sarikoli vowels

	Gao (1963)	Gao (1985)	Pakhalina (1971)	Ku'erban (1994)	Peters & Peters (1996)
Vowels					
i	✓	✓	✓	✓	✓
e	✓	✓	✓	✓	✓
ɛ			✓		✓
æ			✓		✓
a	✓	✓	✓	✓	✓
u	✓	✓	✓	✓	✓
ɯ	✓	✓	✓	✓	✓
ʊ			✓		
o	✓	✓	✓	✓	✓
ɔ					✓
ə	✓		✓	✓	
Diphthongs					
oi	✓	✓	✗	✗	✗
ɛi	✓		✗		✗
ai	✓		✗	✗	✗
ui	✓	✓	✗		✗
ɯi	✓	✓			✗
iu	✓	✓			✓
eu	✓	✓	✓	✓	✗
əu	✓				
ɯu	✓	✓			
ei		✓		✓	✗
ou		✓		✗	
ao			✓		
ɛu			✗		✗

Note: A <✓> indicates that a segment was transcribed by the researcher as presented in the table. A <✗> indicates that the sequence of sounds was present in the data but had been interpreted as a Vowel + Consonant/Glide sequence with no indication that the researcher classified them as a diphthong.

14

To make an obscure situation even more obscure, Pakhalina (1966, 1971) and field informants contend that there are three different dialects of Sarikoli, which differ from one another primarily in the realization of diphthongs (Table 1.2).

Table 1.2 Comparison of diphthongs in three dialects of Sarikoli

Central (Tashkorgan) 塔什库尔干	Near Eastern (Vacha) 瓦恰乡	Far East (Burungsal) 布龙夏村
ɛj	æj	aj
ɛw	æw	aw

Given the already widespread disagreement on the number of Tajik vowels, this difference in local vernaculars seems suspicious indeed. Could it be that different scholars at different times heard different dialects and therefore reached inconclusive results? Or could it be that the language is in a state of development and this development can be seen (at least partially) in the dynamic change among speakers? In other words, is the widespread variation in reports about diphthongs because there really are so many diphthongs or do the reports conflict because different dialects have led some researchers to think there are more diphthongs?

1.6 OVERVIEW OF THE DISSERTATION

To try to answer some of these questions, this dissertation concerns itself primarily with the analysis of vowels in Sarikol Tajik. Specifically, the current research

seeks to determine how many diphthongs might be in the language. To do so, however, it must enter into a discussion of what a diphthong is, whether or not different diphthongs are present in different dialects of Sarikoli, and how these diphthongs might be developing over time. After analyzing some of these questions a final proposal will be made on the structure of the vowels in Sarikoli, especially in light of other languages in the language family. Ultimately, it will be shown that Sarikoli Tajik contains three diphthongs along with a short and long vowel distinction. These long vowels are in the process of developing into diphthongs, a process which is most advanced in the more eastern areas. The process already begun in the easternmost lects seems likely to continue across all Pamiri mountain lects until long vowels have developed into diphthongs in all lects.

To reach that final conclusion, Chapter 2 begins with an examination of the controversies surrounding diphthongs and some of the major previous studies that have been conducted into diphthongs in various languages of the world. Chapter 3 analyzes Sarikoli vowels by looking at spectrographic evidence. Since it is has been proposed that some dialects might have different diphthongs than others, Chapter 4 analyzes each of three dialects to determine if they truly have different diphthongs. This will lead to Chapter 5, which examines how these segments have developed in the language and across dialects of the language. To determine how they have developed and what they might be developing into, Sarikoli will be compared with both Avestan and other Pamiri mountain languages. Finally, a synchronic and diachronic/developmental proposal will

be made that solves the question of the number of diphthongs in the language, brings Sarikoli into more conformity with other languages of the Gorno-Badakhshan group, and maps the history and possible future of the language. Chapter 6 places all of the previous chapters in a theoretical perspective, showing that the present study has implications for theoretical understandings of diphthongs and perhaps for theoretical approaches to linguistic problems in general. Chapter 6 ends with some proposals for further research.

2.1 INTRODUCTION

Diphthongs are far from being understudied in linguistics (Miret 1998). In fact, one linguist even went so far as to say 'diphthongs are a positive migraine' (Lass 1984:95). Another ended his study of diphthongs by asking and answering this question: 'What do the present results tell us about the perception of 'genuine' diphthongs? The answer to this is: very little' (Schouten 2000:38). Despite many studies of diphthongs and many sincere attempts to define them, much about diphthongs remains a mystery.

For many years under the U.S. structuralist school of linguistics led by Bloomfield, diphthongs were excluded from discussions of vowel systems altogether, left to be interpreted as a sequence of a vowel + glide. This interpretation meant that diphthongs were not even included in an analysis of vowel systems but were relegated to some other portion of the grammar. Even linguists outside of the structuralist school have excluded diphthongs from their discussion of vowels on the basis of their intuition or on either some sort of operational or analytical grounds (Lass 1984).

However, if diphthongs are relegated to some other part of the phonology, it seems that certain aspects of the 'oneness' of diphthongs and other simple or long vowels are lost. For example, In English and German neither long vowels nor

diphthongs can occur before /ŋ/. Also in the English Great Vowel Shift long vowels ([iː]

and [uː]) changed into diphthongs ([ai] and [au]. Likewise diphthongs can change into

long vowels. For example, early in Germanic /ei/ became /iː/ (i.e. a diphthong became a

long monophthong) while in Lithuanian [oː] became [uo] (i.e. the opposite happened,

where a monophthong became a diphthong) (cf. Stampe 1972, Szemerényi 1999). If

diphthongs are not integrated methodologically into accounts of vowel systems of

languages, then surely highly relevant similarities between diphthongs and more

traditional simple vowels or long vowels is being lost. On the other hand, if diphthongs

are integrated into phonological representations of vowel systems, then what is the best

way to do so? After examining several pages of possible ways to integrate diphthongs,

long vowels, and short vowels into one coherent system, Lass finally says, 'This leaves

us with no really good solution, but a multitude of half-baked approaches' (1984:101).

Unfortunately, this seems to be the current state of affairs regarding diphthongs: many

approaches and ideas, but none truly able to account for their complexity in any truly

systemic way.

It is possible that many linguists have found it easier to simply exclude

diphthongs from overall vowel systems because diphthongs are phonetically and

temporally complex. While other vowels seem to be steady and unchanging in their

character, diphthongs are always changing. While other vowels seem able to occur in

nearly all environments in a language, diphthongs are often restricted by syllable shape

or stress. While the question remains of how to integrate such segments into a

phonological framework, there appear to be even more fundamental questions at work. For example, is there even agreement on what a diphthong is phonetically (much less phonologically)? It seems reasonable that if phonetics is concerned with the world of physical movement and physical waves in space, then it seems that phonetics may have a better chance of resolving the issue of what a diphthong is than phonology. At least, this was the thinking initially. Now, it turns out that phonetics and diphthongs is every bit as much a minefield as phonology and diphthongs. Partially this is because many linguists have not been careful in weeding out phonological assumptions from their phonetic studies. This is not the whole picture, however. Due to the computer revolution, more and more field linguists have been able to apply laboratory methods to an increasing number of languages. However, this increase in data seems to have actually made the picture of diphthongs more complex instead of less so. For example, since so much more is known about the speech stream than before, much more is known about how diphthongs vary across languages and across time than before. Eighty years into laboratory phonetics, an agreed upon definition of a diphthong is as elusive as ever. In addition, it seems that many aspects of diphthongs might be either language specific (Peeters 1991) or else might require a combination of phonetics and phonology to fully understand.

Because there has been such disagreement about diphthongs, many linguists have studied them. In fact, there have been so many studies done already that one may question why so little can be definitively stated about the exact nature of a diphthong.

These phonetic studies of diphthongs have been primarily of two types: language specific (e.g. Aguilar 1999, Bladon 1985, Espy-Wilson 1992, Jha 1985) and those that are more theoretically oriented or that try to draw cross-linguistic comparisons based on a survey of several languages, i.e. they are not focused on one specific language (e.g. Lindau 1990a, Peeters 1991, Schouten 2000). This is not to imply that those studies that focus on one specific language also are not concerned with theoretical issues. In the course of these various studies, two competing definitions of diphthongs begin to emerge: those that see a diphthong as being composed of two steady states with a glide in between and those that see a diphthong as merely a vowel with some kind of movement. Nearly all the studies begin with some phonological assumptions, some of which are stated clearly and some that are not. Those that do not begin with phonological assumptions seem to turn to phonology in the end to boost their initially essentially phonetic arguments.

In section 2.2, some (although by no means all) of the language specific studies will be examined. In section 2.3, the more theoretically based studies will be examined. In the case of both kinds of studies, the goal will be the same: to determine if any cross-linguistic definition of a diphthong has been deemed possible by previous researchers. If such a definition or any kind of consensus at all has been reached, then such information will be helpful in resolving the dispute over the exact nature of Sarikoli vowels.

2.2 LANGUAGE SPECIFIC STUDIES

The following sections examine some of the studies done on specific languages. The following sections are organized by similarity of conclusions: languages in which linguists have made similar arguments concerning diphthongs, are presented together.

2.2.1 ENGLISH

In his study of Canadian English diphthongs, Warden (1979) examined the vowels of Canadian teenagers. In doing so, he had two goals in mind: the first was to determine the phonetic realizations of those diphthongs and the second was to examine any possible variation. The study begins by proposing a fairly typical definition of a diphthong, one which Warden borrows from Kopp (1966): 'Each of the diphthongs produces a distinctive pattern composed of the first vowel position from which the glide starts, the glide, and the second vowel position at which the glide terminates' (74-75). In other words one would expect spectrograms that have two steady states with a glide in between, schematized in Figure 2.1. It does not matter whether the first steady state or the second are higher or lower in terms of formant values. What matters is the fact that there are two steady states with a transitionary section in between. However, later in the same article, Warden muddies the waters a bit when he says, '…the diphthong is never long and drawn out but consists of a very rapid change from one vowel position to another' (38). It makes no mention of whether there was a transitionary state between the two steady state portions or not and there is no corresponding picture to clarify the statement. Indeed, he seems to switch to more phonemic definitions rather than phonetic

based ones when he says, 'it cannot constitute the peak of a syllable' (43) and later when he makes a judgment on the diphthong status of a segment based on a reference to a syllable boundary (44). Perhaps this, in and of itself, helps emphasize the fact that judgments as to whether segments are diphthongs or not often begin rather simply but then are later left to a combination of factors taking into account both phonetics and phonology. Warden also noted that vowels in unstressed or closed syllables have a tendency to be less diphthongized, further complicating any potential judgments one might want to make.
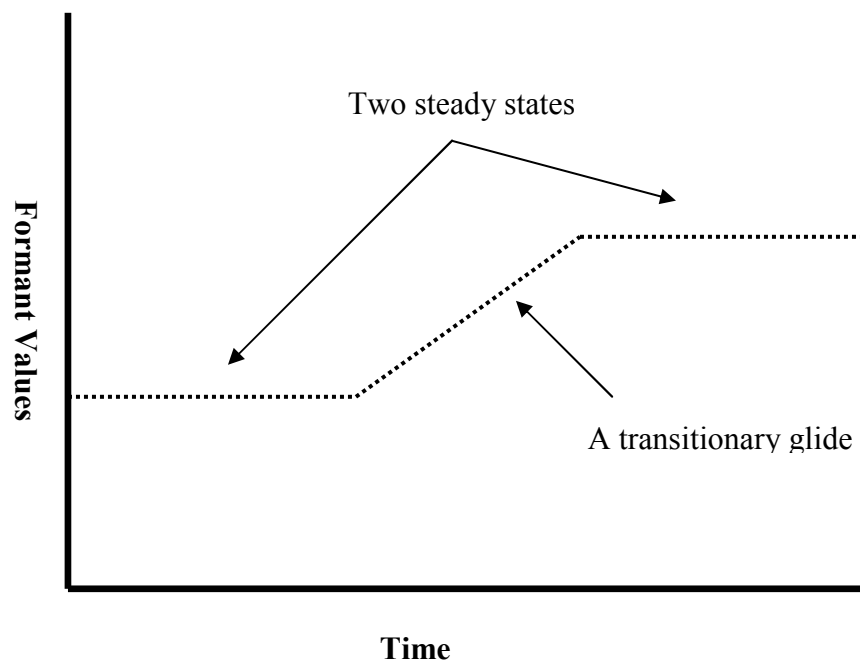
Figure 2.1 Schematized representation of a diphthong with two steady
state portions and a transitionary segment

Rather than Canadian English diphthongs, Gottfried et al (1993) considered American English diphthongs. In their study, they began by examining previous work on American diphthongs, determining in the process that linguists had essentially three different understandings of diphthongs. They used those three proposals as their starting point, evaluating each in turn, in an attempt to determine which (if any) might be the better understanding. The first, the 'onset + offset' hypothesis, defines a diphthong as two vowel targets. These two vowel targets may or may not correspond to simple monophthongs. This definition would be similar to that in Figure 2.1, though the transitionary element is not specifically referred to in this study. That is, the two endpoints are in focus in this definition and little or nothing seems to be said about how the transition is formed from the onset steady state portion to the offset steady state portion. It is possibly assumed that the transition is a slope like that above since this is the only way such transitions are known to occur in English and merely not mentioned for that reason (cf. Weismer 1999). The second view, the 'onset + slope' hypothesis, is based primarily on the work of Gay (1968), in which it was argued that F2 rates of transition are steady in diphthongs. It was further argued that diphthongs can then be classified in terms of their formant values at the onset and the following F2 rate of transition. The third view considered in their article, the 'target + direction' hypothesis, seems to argue for continuous formant change over the length of the diphthong but that there need not be any particular steady state at any given point. (See also Figure 2.2.) After evaluating all three definitions and classifying all the tokens of diphthongs in their

24

own corpus, they found that all three definitions were quite good at describing diphthongs! The first definition (the 'onset + offset' hypothesis) was slightly better than the other two, but all three definitions worked at a rate greater than 90% of the time. This makes it impossible to establish the definitive definition based on their study, but it does point out that diphthongs are a many-faceted object. Perhaps it is impossible to phonetically state that diphthong are always only one thing and not another, or that perhaps diphthongs are a conglomeration of all of the definitions above. Clearly, more work needs to be done before any agreement can be reached.

If the effect of prosodic context on (English) diphthongs is included in the debate about the structure and nature of diphthongs, then agreement is even harder to reach. It is well known that diphthongs often fall short of their second target vowel (Solomon 1984, Weil 2000, Wouters 2002). That is, if a diphthong is said to be composed of two underlying segments such as /ai/, then the diphthong often fails to reach the same formant values as a simple /i/. This is true in all varieties of English (and in many other languages as well). Instead, the second segment formant values will end in the area of a mid-high vowel rather than reaching the actual values for a high vowel. This effect is said to be greater in rapid speech. In addition, such factors as stress, pitch accent, word position, register, and word class have been shown to have effects on vowel articulation (Wouters 2002). If this is the case, then many diphthongs will, indeed differ in their phonemic and phonetic representations. Speakers will see themselves as articulating a diphthong that in fact never reached its second target at all due to some

other prosodic factor. This makes identification of diphthongs in understudied languages even more difficult, requiring some kind of control for prosodic elements.

Another factor affecting whether a diphthong reaches its second element or not is the particular dialect of English being spoken. For example, Southern American English diphthongs often have much shorter second elements and much shorter transitional elements than those found in Standard American English (Weil 2000). In fact, speakers of Southern American English have been said to substitute [aː] for /ai/, a difference that is often mimicked by actors assuming southern accents.

Others who have studied dialects of English have also pointed out this difference between phonology and phonetics. Actually, many segments that are considered as phonemic long vowels in the language are actually noticeably diphthongized phonetically (Minkova 1998, Thomas 2001). This means that even if any one or all of the definitions above are met by a particular segment, a linguist may still not consider the segment as a long vowel for purely phonological or historical reasons. For example, in Received Pronunciation, the vowel in words such as *tree* is noticeably diphthongized but is still considered to be /iː/ rather than a diphthong (Grimson 1970).

This ability to classify an item as a diphthong phonetically but as a long monophthong phonemically is especially important to note in light of the Sarikoli situation being examined in this dissertation. Every other language in the Pamiri mountain sub-branch has been posited as having a phonemic short and long vowel distinction (Sokolova 1953, Sokolova 1967). Sarikoli is the only one posited to have so

many diphthongs but no long vowels. Is it possible that, in light of how linguists have analyzed English vowels, Sarikoli has phonetic diphthongs that ought to be considered as long monophthongs based on historical or phonological reasons? It seems that other researchers have at least opened the doors to such a possibility.

Given that there is often such a large difference between an 'ideal' diphthong and what is actually produced and between phonetic and phonological understandings of the phenomenon, the question then arises: What part of a diphthong is absolutely necessary to be perceived as a diphthong by listeners? Is it the end points or the transition in between the two vowels or something else altogether? The most well known proponent of the view that the rate of change is the most important part of the diphthong is Gay (1968). In his study of the diphthongs /ɔɪ/ and /aɪ/ in English, he found that the F2 rate of transition was more essential to diphthong perception than either the onset or offset steady state. However, this work is now quite dated and has been superseded by more recent work that has been able to both modify and improve upon his study and has also been able to make use of better and more modern equipment. Critics have pointed out that Gay's vowels did not even contain a steady state portion, therefore not giving respondents a chance to recognize the importance of the two steady state portions. If the steady state portions are never there, then there is no contrastive data to prove either their significance or insignificance in diphthong identification. Furthermore, Gay's study has a potential problem in common with all studies of English that have come later: the fact that speakers are limited to identifying

segments that they hear with some known segment of the English language. Therefore, even if segments were perceived as being 'strange' or 'unusual' instances of a diphthong, they were still able to be identified because there was simply nothing else in the neighborhood to choose from. So, for example, if a diphthong that was /ai/ was produced as [aɛ], speakers were still able to recognize and choose the proper diphthong because there simply was nothing else to choose from.

Gay's study touched off a debate that has persisted for many years. This was in part because his study challenged the traditional assumption that a phonemic diphthong in English was composed of three parts: two steady state portions and a glide portion (see Figure 2.1). Gay also claimed that the rate of F2 transition in American English vowels was not dependent on the rate of speech. In slow, moderate, and fast speech the rate of F2 change was the same. He therefore defined diphthongs as two steady states (whose length depended on the rate of speaking) and a transition period (whose rate was not dependent on rate of speech). Such conclusions made it seem as if it was the transitionary part of the vowel that was most important and not the target endpoints. Such a bold challenge to tradition prompted many follow-up studies. Which part of the diphthong was the critical part? Was it the glide or the steady states or some combination of the two? Unfortunately, despite much work, little has been said definitively. For example, respondents who were asked to identify diphthongs that were modified on the computer were able to successfully identify English diphthongs even if they had a long glide (with no steady state portion); they were also able to successfully

identify English diphthongs that had no glide at all, as long as the steady state portions were long (Bond 1978, Bond 1982). Only when both the glide and the steady state portions were short were respondents likely to identify the diphthongs as monophthongs.

To try to update Gay's study and to either confirm or disprove his hypotheses, Bladon (1985) added to and refined upon Gay's methodology. Bladon had phonetically trained respondents to listen to and identify diphthongs that had been modified in various ways. Respondents were all speakers of English but were told not to assume that the diphthongs were found in the English language. Bladon found that when respondents listened to diphthongs that fell short of cardinal positions they were able to identify them as such, i.e. they did not associate them with a target that was not reached. For example, if respondents heard the sound [iɛ], they did not associate it with /ia/ but correctly identified it as falling short of its target. This lends credence to the argument that Gay's study was undermined by the fact that respondents were only allowed to choose from those diphthongs that were phonetically possible in the English language. To further test Gay's work, Bladon asked respondents to listen to diphthongs that had no transitions at all. In other words, they only listened to the two steady state portions of the diphthong. Results show that respondents were able to correctly identify the diphthongs 100% of the time. Finally, respondents listened to diphthongs that only contained transitions (similar to Gay's original study) and were once again to make identifications. In this case, respondents had quite a bit of trouble and only achieved a

44% success rate. In addition, they felt that those times in which they did correctly identify the diphthong were quite difficult or sounded unnatural in some way. These results led Bladon to conclude that transitions serve as a weighting flag and as a temporal pointer but are not important phonetic elements of diphthongs. This leads to the question: Then what of all the diphthongs who never seem to reach their second element? If the steady state portion is so important, then how are speakers correctly able to identify diphthongs that only contain one steady state and a transition? Could it be that the initial steady state is more important than the second steady state? And what to make of work such as Gottfried's (1993) in which it was shown that neither Bladon's nor Gay's definition would work? In the end, it seems that Bladon's conclusions might be too dogmatic to account for the ever-changing nature of a diphthong or the amazing ability of the human brain to decode signals that do not meet their ideal endpoints.

Dolan and Mimori (1986) also sought to examine Gay's claims about fixed rates of F2 transition in diphthongs. Perhaps adding to the problem of the definition of diphthongs rather that reducing it, they compared both American English and Japanese diphthong rates of transition. The problem with this is that Japanese diphthongs are actually vowel clusters by most standard definitions because a syllable boundary intervenes between the two vowels. They found that in both English and Japanese the F2 rate of transition did change depending on the rate of speech. In English, speaker transitions depended on both the amount of time available for the transitions (i.e. rate of speech) and also how far the diphthong had to travel in articulatory space (i.e. the

specific diphthong in question). Similar results were also found in Japanese. (For more on Japanese 'diphthongs', see Hirasaka and Kamata (1981).) Overall, Dolan and Mimori's conclusions were much more similar to Bladon's than to Gay's.

Even more recent work has begun to question the exact value of the endpoints of all vowels -- diphthongs and monophthongs alike. Traditionally, definitions of monophthongs have said that the critical point of the vowel is the middle of it, where formants are generally steady and prominent acoustically. However, recent work has suggested that monophthongs are also inherently dynamic and that perception of them depends on their dynamism and their endpoints more than on their static portions (Nearey 1986, Strange 1989). This means that when linguists describe diphthongs as a vowel with continuously changing formant qualities (Kent 2002, Ladefoged 1996), such definitions are not nearly precise enough. Researchers will have to look again to find the difference between diphthongs and monophthongs.

Harrington and Cassidy (1994) sought to test this dynamic theory of monophthongs and diphthongs in Australian English. In their viewpoint, vowel perception is inherently dynamic with listeners extracting information about the vowel from information that is distributed throughout the speech signal when decoding vowels of all kinds. It did not matter whether the vowels were monophthongs or diphthongs. To test this theory, they removed the middle portion of monophthongs produced by speakers of Australian English and asked respondents to identify which vowels were spoken. They reasoned that if classification of diphthongs is done based on transitions at

end points of the diphthong (i.e. the transition from the previous segment into the diphthong or the movement from the diphthong into the following segment) as some have argued, then respondents should be able to better identify the vowel by the two endpoints than by listening to a small segment of the middle of the vowel. On the other hand, if more traditional target theories (which state that it is the center of the monophthongs that is the critical portion) of vowels are more accurate, then only diphthongs should benefit from the information at the margins of the vowel. In the end, their study seemed to uphold the 'target theory' of diphthongs. While a small number of monophthong identifications benefited from the information at the vowel margins, the number was small. In other words, it did not matter whether respondents were given three spectral slices (from the two margins and the center) upon which to identify their target or were simply given one spectral slice (from the midpoint); results were the same. On the other hand, most all of the diphthong identifications benefited from the inclusion of the material from the vowel margin. Such results indicate that the traditional notion of the definition of a diphthong can be upheld, i.e. the difference between monophthongs and diphthongs is in fact that diphthongs are dynamic while monophthongs are static. However, there is still plenty of disagreement in the literature and many more experiments will undoubtedly be conducted as more is learned about human cognition in general and about speech perception specifically.

Many such similar conclusions are reached in studies of other languages of the world as well. The following sections explore first some of the more similar conclusions

and then examine languages and studies that argued for a different understanding of diphthongs.

2.2.2 MAITHILI

Jha (1985) examined diphthongs in Maithili, an Indo-Aryan language of Nepal and India. While acknowledging that there is widespread disagreement on the definition of a diphthong, he started from the perspective that a diphthong consists of two steady state elements with a transition in between.
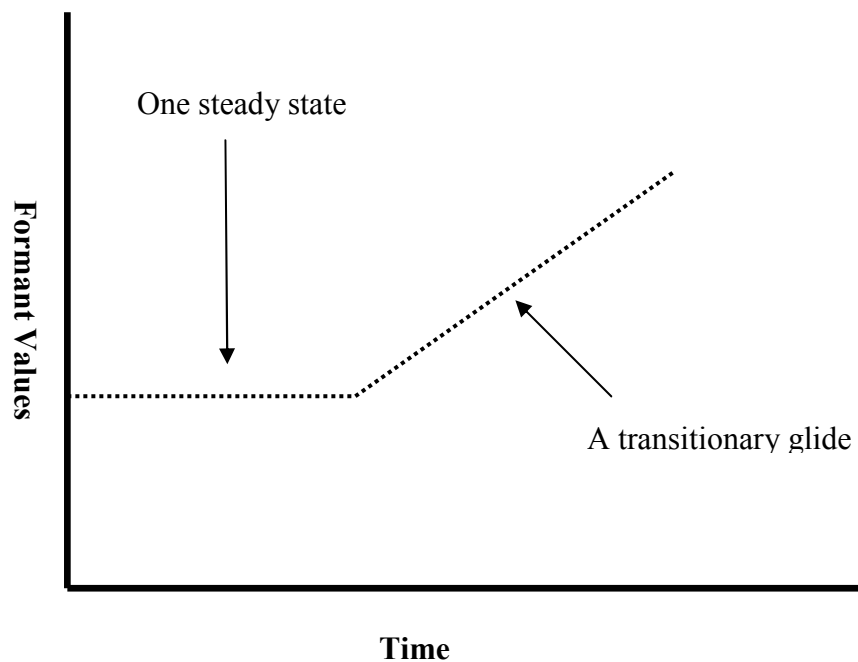
Figure 2.2 Schematized representation of a moderate or rapidly spoken diphthong which fails to reach a second steady state position

In this study, recordings of Maithili diphthongs were made at three different speeds. At slow speeds, the diphthongs did indeed exhibit two steady states with a

33

transition in between. However, in some speech at moderate speeds and in all rapidly spoken speech, the second element of the diphthong was never reached. Movement in that direction was seen in the spectrograms, but the second steady state portion never totally materialized (111). This meant that diphthongs spoken in slow speech did reach the ideal of Figure 2.1, but more rapidly spoken diphthongs were more like Figure 2.2.

These results led Jha to conclude that the important elements in a diphthong (at least in Maithili) are the onset steady portion and the rate of change of F2. Essentially, he said that the second steady state portion is perhaps ideal in diphthong production but is by no means required, actually only being reached in slow, deliberate speech.

2.2.3 SHONA

Pongweni (1983) also seems to feel that the first steady state portion of the vowel is more important than the second. In his study of Shona (a language of Zimbabwe), he never even mentioned the problem of what a vowel might be phonetically speaking, but outlined a methodology that assumed the same pattern as that in Figure 2.1, i.e. a vowel with two steady states and a glide portion in between. When the second vowel in each of his sequences was longer than the first, he argued that it must, therefore, belong to a separate syllable. By arguing that the second steady state portion of the vowel must belong to another syllable by virtue of its longer duration, Pongweni is implicitly accepting Jha's view that the second steady state portion of a diphthong is not as important as the first. All of Pongweni's other arguments for the status of the diphthongs in Shona come from phonology, indicating once again that

phonetics alone does not seem to be able to ultimately determine the issue once and for all of whether a segment is a diphthong or not.

2.2.4 H<small>AUSA</small>

Hausa has been said to have two diphthongs: /au/ and /ai/ (Lindau-Webb 1985). Spectrograms have revealed that Hausa diphthongs have two steady state portions with a transition in between. However, the /ai/ diphthong has been revealed to be realized as [eː] 80% of the time (51). This has lead some to speculate that perhaps this segment is not a diphthong at all but rather ought to be thought of as a long vowel (Newman and Salim 1981 in Lindau-Webb 1985). In fact, this segment has consistently different formant values than the long mid vowel in the language, lending credence to the fact that perhaps this is a long vowel rather than a diphthong at all. This controversy helps to highlight the different approaches that linguists have taken in their approach to phonology vs. phonetics in their interpretation of data. It seems that both researchers were confronted with a similar set of facts, but Lindau-Webb chose to give (apparent) phonological considerations more precedence than the phonetics (though she does not explain why). On the other hand, Newman and Salim chose to follow more phonetic considerations. This, perhaps more than any other, helps to highlight some of the reasons why so little has been definitively said about diphthongs. Even when agreeing on the facts of the case, linguists do not agree on how those facts should be interpreted. This is likely related to theoretical biases and personal preferences, but such never seem

to be stated in the articles on the topic, leading linguists to talk past each other without addressing their underlying assumptions.

2.2.5 PENNSYLVANIA GERMAN

The situation in some dialects of Pennsylvania German is said to be similar to that in Hausa in that both languages have been argued to contain phonemic diphthongs that are then manifested phonetically as long monophthongs (Keiser 2000). In his paper on the subject, Keiser argues not only from a phonetic and phonological standpoint but also from a historical one. He points out that the current diphthong /aɪ/ in Pennsylvania German is a reflex of the Middle High German long, high monophthongs /iː/ and /yː/. Therefore, what once was a high monophthong became diphthongized and is now on its way back to once again becoming a long monophthong. Such processes seem to be quite common historically, particularly in vowels of Indo-European languages (Beekes 1995, de Vaan 2003, Szemerényi 1999). Such changes have been noticed in English, German, Greek and Albanian, just to name a few Indo-European languages in which such changes have been documented (Keiser 2000). These changes back and forth can happen even when the change causes one vowel to encroach on the vowel space of a previously established phoneme in the language, as seems to be the case in Pennsylvania German. This means that such changes can cause major disruptions/changes in the overall phonemic system of a language.

Critically, such changes in vowels may be sociolinguistically motivated (as in Labov's now famous Martha's Vineyard studies) or they may simply be examples of

'drift' (a term coined by Sapir), i.e. languages simply change over time. Although, variationists such as Labov would argue that no sound change is entirely without sociolinguistic motivation. If it is true that a certain amount of variation is allowed among speakers of a particular dialect, then something must cause one variety (i.e. an 'activated variety') to eventually overwhelm the other varieties (i.e. to 'spread'). Nevertheless, after examining a variety of sociolinguistic factors, Keiser was not able to credit any of them with motivating the sound change happening in Pennsylvania German, leading him to conclude that diphthongs were becoming monophthongized due to normal historical processes, i.e. drift.

This study, nevertheless, brings up an important point for those studying understudied languages. It points out that languages are always in a state of change and that variation is common, therefore it is entirely possible that the overall phonology of a language is in a state of drift and that speakers and communities do differ from each other. Such factors must be taken into account as much as possible when making judgments about the phonology in general and about diphthongs in particular. Since diphthongs are particularly likely to change and develop, various perspectives must be brought to bear when making decision about their status in a language. While many other linguists assume the importance of the phonological perspective, Keiser's study effectively points out the necessity of also including sociolinguistics, historical linguistics, and also variationist theories to bear when studying diphthongs.

2.2.6 AUSTRIAN GERMAN

This issue of language drift and sociolinguistic variation in diphthongs has also been studied in Austrian German. In Austrian German, the diphthongs [ae] and [ao] are generally monophthongized to [ɛː] and [ɔː] (Moosmüller 1998). This change is said to have begun in the lower socio-economic classes and to have led finally to a total change among speakers of Viennese German. Since the process began in one socio-economic class and spread to the other classes, this change is seen as being sociolinguistically motivated. This change does not yet seem to be as total as originally implied, however. In Vollmann's (1996) study, he found that register affected how the diphthong was produced. In careful speech, diphthongs were produced, while in more casual speech long monophthongs were recorded. (The same results were obtained by Moosmüller (1998).) In addition, stress also seemed to play a factor as diphthongs were realized in stressed positions while monophthongs were realized in unstressed positions. If it is true that this change from diphthongs to monophthongs began in the lower socio-economic classes, then it seems that the change is still not totally accepted in higher registers, which could account for the differences in his data.

These results indicate that, in languages in which there is purported to be variation or which are claimed to be in a state of transition, slow, clear speech is much more likely to produce the older, more conservative version of the language than rapidly spoken speech. Furthermore, it may be necessary to pay attention to issues of stress placement when examining language change, since this issue also seemed to affect

whether diphthongs were actually phonetically realized or not (cf. Moosmüller 1998). It seems likely, in light of the Viennese speakers' ability to switch back and forth between the monophthongs and the diphthongs, that these segments are still phonologically diphthongs that are being realized as monophthongs in most contexts. This indicates that it is likely necessary to study recordings of diphthongs and monophthongs spoken in various registers and under various circumstances to obtain a true picture of the vowels in a language. While the monophthongization represents a lower register in Viennese, monophthongization could just as easily represent a higher register under other sociolinguistic circumstances. This underscores once again the importance of bringing as many linguistic tools and perspectives as possible to bear on the issue of the status of diphthongs in a language.

2.2.7 ITALIAN

Italian diphthongs have been defined as the presence of one segment marked [+high][+stress] and followed by a vowel marked [-high] (Salza 1988). This can be seen in Figure 2.3. What is interesting about such a definition is that it makes no reference to either of the two definitions mentioned above: it neither concerns itself with whether a diphthong is a vowel with constantly moving formants nor if a diphthong has two steady state portions. It seems to take a different direction altogether.

Italian phonetics seems content to accept something in between either of the two previous definitions as a diphthong. Traditionally, a diphthong was seen as having two phones with two targets but only one syllable nuclei (Salza 1988). No mention is made

of the need to have a glide element in between nor is there any requirement of a steady

state portion. Actually, there is not even any requirement that the first part of the

diphthong be considered a vowel, as can be seen by the +/- value indicated in the initial

segment of Figure 2.3. In Salza's (1988) study of diphthongs, such issues are not even

raised. In his study, he includes one spectrogram of a diphthong versus a vowel cluster.

In the spectrogram of the diphthong, the formants do show continuous and radical

movement, and the second segment does seem to reach a steady state, but the first part

of the diphthong never seems to settle into a steady state. In addition, there is clearly no

separate transitionary element between the beginning and the end of the diphthong.

$$
\begin{matrix}
\text{V} & & \text{V} \\[4pt]
\begin{Bmatrix} \pm\,consonantal \\ +\,high \\ +\,stress \end{Bmatrix} & & \begin{Bmatrix} -\,consonantal \\ -\,high \\ -\,stress \end{Bmatrix}
\end{matrix}
$$
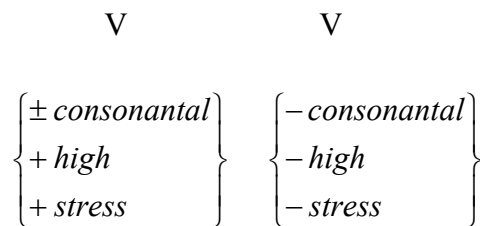
Figure 2.3 Traditional representation of Italian diphthongs

What could account for the fact that Italian linguists have such a different

understanding of vowels than linguists studying other languages? First, it seems that

Italian linguists are asking a different question than some of the linguists studying other

languages. For linguists studying Italian, the fundamental question seems to be: one

syllable or two? diphthongs or vowel clusters? It appears that Italian has a constraint in

40

its phonology that determines whether two contiguous vowels end up in one syllable or two. (That constraint is characterized in Figure 2.3.) On the other hand, Italian does not seem to be concerned with the phonetic differences between long and short vowels. Therefore, any time any vowel-like sequence is syllabified in such a way as to put both vowels in one syllable, it can be called a diphthong -- at least a diphthong in Italian. Therefore, one could say that this definition is one that is more phonology driven than phonetically driven -- even in studies of diphthongs that are otherwise phonetically focused (e.g. Salza 1988).

Such studies can initially disappoint if one is searching for phonetic correlates of diphthongs (as the study in this dissertation seeks to do). However, perhaps people like Salza are making tacit arguments that phonetics and phonology cannot be separated when studying diphthongs. His work does seem to indicate that if one ignores syllable boundaries when studying diphthongs, one may reach wrong conclusions. These Italian studies also show that stress patterns can make a difference in the length of a segment and therefore determine whether it gets placed into a tautosyllabic or a heterosyllabic sequence. This means that while spectrograms are helpful in Italian in determining whether a segment is a diphthong or something else, they cannot be read in a vacuum and must constantly be informed by the phonology. Such an approach may work well in languages in which the phonology has been studied and well established, but may not work in cases of understudied languages.

2.2.8 SPANISH

Like Italian, Spanish is said to have both diphthongs and also vowel-vowel sequences. Therefore, it is not surprising that the basic facts of both languages are quite similar. Like Italian, Spanish diphthongs are also defined in relation to the phonemic notion of syllable boundaries, with diphthongs being defined as two vowels in one syllable, while sequences are two contiguous vowels in two separate syllables. For example: píe [ˈpi.e] 'I/he/she/it cheeped' vs. pie [ˈpi̯e] 'foot', where the only difference between the hiatus and the diphthong is the syllable boundary (Aguilar 1999).

However, outside of this basic distinction many questions remain about the nature of diphthongs in Spanish and little reference is usually made to phonetic properties when such phonological explanations are made (Aguilar 1999, Mauder 1996). The situation is made more complex in that several varieties of Spanish are able to 'reduce' Vowel-Vowel sequences across syllable boundaries into diphthongs, i.e. they are placed in the same syllable. This is a predictable pattern based on word stress, but can create phonetic realizations of diphthongs that otherwise would be considered as Vowel-Vowel sequences phonologically.

Like many of the other languages mentioned above, Spanish vowels differ phonetically depending on whether they are in the onset or offset position in a diphthong (Borzone de Manrique 1976). That is to say that an [i] in an onset position in a diphthong and an [i] (or any other allowable vowel) not occurring in a diphthong are

42

not exactly the same phonetically; their F1 and F2 rates do not match exactly. Those vowels in the onset position seem to have more frication than other vowels and also a greater displacement in frequency. In addition, those vowels in offset positions seem to rarely actually meet their vowel target falling short of reaching the high position, ending instead somewhere around a mid vowel position.

It has been argued that the difference between the sequences and the diphthongs has been said to be one of formant transition rate. Long, slow transitions are said to identify diphthongs while short, quick transitions are indicative of different syllables (Borzone de Manrique 1976). While later work does not dispute such findings, newer research does seem to indicate that diphthongs and vowel sequences also differ in other ways. For example, vowel sequences are longer in duration than diphthongs, have greater curvature in their F2 transition rates, and can be simplified into either diphthongs or long vowels (Aguilar 1999). Of particular interest here is the fact that vowel sequences can be reduced to diphthongs in certain environments. It seems that when durations decrease (i.e. when spoken in rapid speech) that vowel-vowel sequences can be reduced to diphthongs or even to long vowels. This kind of reduction is something that has already been seen in several other languages. When this happens, the only way to recover the fact that the segment in question is a vowel sequence and not a diphthong or long vowel to begin with is to resort to a phonological notion of the differences between the two, once again underscoring the close relationship between phonetics and phonology when it comes to diphthongs. A tendency to reduce vowel-

vowel clusters to diphthongs is also reported to be a historical process in Spanish, indicating that historical considerations must also be married to phonetics and phonology and brought to bear on the problem of diphthongs (Aguilar 1999).

2.2.9 ESTONIAN

In her (now classic) study of Estonian, Lehiste (1967) asked a similar question to the ones that have been asked of Italian. She sought to discover the difference between diphthongs and vowel sequences in Estonian. She, however, took a different approach. Most importantly, she began with phonological assumptions about where syllable boundaries were. Thus, vowel sequences and diphthongs had already been identified based on their phonological properties. In her paper, she sought merely to find any phonetic differences that might be an indication of the phonological assumptions she had made.

She examined the phonetic differences between diphthongs and vowels from several different perspectives. The first was whether the first component of a diphthong was similar to its equivalent short vowel or not. She found that they were quite similar. Secondly, she looked at the second part of the diphthong. Unlike the onset vowel, she found that these offset vowels were not the same as their equivalent short vowel. Finally, she looked at intensity. She found that this was a key factor in separating diphthongs from vowel sequences. This was because vowel sequences had two intensity peaks, while diphthongs had only one. Intensity appears to be a simple and yet telling

difference between diphthongs and monophthongs that has been overlooked by later researchers.

Although the questions regarding Estonian and Italian are similar, i.e. long vowels vs. vowel sequences, the paths that individual researchers have taken to answer the question have been quite different. In fact, Lehiste's older study provides more clues to the linguist hoping to research understudied languages. Although she begins with phonological assumptions, it is possible to use her work to 'work backwards' from phonetic facts to help make determinations about phonological decisions. In this regard, her results showing that intensity peaks can help determine diphthongs versus vowel sequences has great potential consequence to field linguists endeavoring to make such determinations in other languages.

2.2.10 DUTCH

Traditional Dutch grammars make a distinction between 'genuine' and 'pseudo' diphthongs. Genuine diphthongs were seen as vowel clusters that operated as one phoneme, while pseudo diphthongs were vowel clusters (or perhaps vowel + glide combinations, depending on the interpretation of the linguist) that contained two phonemes. (For a comprehensive overview of the history of studies of Dutch phonology, see Zonneveld and Trommelen (1980).) As such, Dutch is an encapsulation of all that is difficult in the study of diphthongs.

This already confusing situation is made more so by the fact that Dutch also contains long vowels that can be diphthongized (i.e. have a release that is a high vowel)

but are considered to be merely long vowels and neither of the two kinds of diphthongs mentioned above (Collier 1982, Gussenhoven 1999). For example, the words /deː/ [deːʲ] 'that (one)' is said to not rhyme with /beːj/ 'I offer' (at least in the Masstricht dialect) (Gussenhoven 1999:159). Further study seems to confirm the fact that the diphthongized long vowels and the true diphthongs are not the same (Zonneveld 1980). For example, diphthongized long vowels that had their second part removed were still able to be identified as long vowels. In addition, the long vowels showed little variation in the initial range of F1-F2 correlates, while the diphthongs varied in their F1-F2 values. It is also worth noting that the phonemic diphthongs undergo an opposite pattern of simplification. Namely, the phonemic diphthongs in the language have allophones that are simply long monophthongs. If the diphthongs can be monophthongs and the monophthongs can be diphthongs, then can the sounds be accurately identified? Is there any real basis for the distinction between long monophthongs and diphthongs in the language?

Phoneticians have endeavored to answer that question. Collier and her colleagues (1982) discovered that there was a real phonetic difference between the two classes of diphthongs. Those diphthongs that were called 'genuine' have 'relatively continuous and gradual changes in formant structure, whereas, the pseudo diphthongs are produced with more abrupt changes in formant structure' (307). This means that the 'genuine' diphthongs look like Figure 2.2 while the 'pseudo' diphthongs look like Figure 2.1. That was not the only difference between the two of them, however. They

also found that the realization of a fixed target is essential for the 'pseudo' diphthongs, but not for the 'genuine' ones. This means that the onset and offset steady states of pseudo diphthongs can be said to be the same as their equivalent simple vowels. However, the onset and offset states of the genuine diphthongs do not correlate specifically with any simple vowel equivalents (318). This is likely because the 'genuine' diphthongs seem to have a monophonemic status in the language and therefore are not simply composed of two simple vowels but are simply separate and unique phonemes that happen to have constantly moving formant values. The 'pseudo' diphthongs on the other hand, are biphonemic. Since they are not a unit, they are composed of two simple vowels instead. Given this set of facts, it does indeed seem that only the 'genuine' diphthongs are in fact true diphthongs. This means that Dutch contains diphthongs that follow the definition of merely having one steady state followed by constantly changing formant values. The second steady state is not required. This is, once again, a judgment based not only on the phonetic evidence but also on phonological and historical considerations.

It has been difficult to find a phonological analysis of Dutch diphthongs that is satisfying on all fronts. Regarding the sequences as two phonemes does not seem to take account of native speaker intuitions that they are single units. On the other hand, regarding them as single phonemes seems to ignore the fact that they are different from simple monophthongs (Cohen 1971). In discussing the problem of phonemic representation, Cohen agrees that phonetics does seem to argue for some kind of

representation that captures the dynamic nature of the segments. On the other hand, it seems that phonologically (psychologically?) only one command is given by speakers to the articulators involved. This would seem to uphold the view that the segments are unitary. In the end, Cohen argues that there would have to be some other new way of representing diphthongs altogether that would be able to account for both phonetic and phonological understandings. It is interesting to note that, 30 years on, the controversy over representation is still unresolved.

Zonneveld (1980) takes the controversy one step further when he goes so far as to state that the phonetic evidence in Dutch is relatively unimportant when analyzing the phonology and morphology. Rather than trying to marry the phonetic and phonological viewpoints (as Cohen had), he argues for separating them and not worrying too much about their relationship. Although he argues from a Natural Generative Phonology perspective, he goes on offer 'a partial phonological analysis of Dutch, apparently paradoxical in terms of the Naturalness Condition, where the phonetic representation of a small and coherent set of sounds is comparatively irrelevant to phonological analysis (266).' If this is the case, then have other linguists simply been wasting their time trying to arrive at a coherent analysis of the phonetic nature of Dutch diphthongs? Most linguists are likely to disagree with Zonneveld and still seek phonetic reality in their phonological analyses wherever possible. In addition, the quest for knowledge of the physical correlates of diphthongs is valuable regardless of the phonological considerations. Zonneveld does not ignore phonetic reality in his analysis.

48

In fact, he spends half or more of the article discussing the phonetic problems associated with Dutch diphthongs before finally making a phonological proposal. Perhaps in the end, he does not so much prove that phonetics is not important in making phonological considerations but merely proves that phonological analyses can be done even when phonetic information is still incomplete or in conflict.

Thus, Dutch is an appropriate language with which to conclude this general survey of language specific studies of diphthongs. Dutch serves as a good example of the thorny problems in analyses of diphthongs and the apparent lack of satisfactory conclusions that are often available.

2.3 CROSS-LINGUISTIC STUDIES

Rather than study a specific language, there have been some attempts to draw more general conclusions about diphthongs based on cross-language analyses. Some of these studies have begun with data from specific languages and then attempted to correlate and assimilate the data across the languages. Others have created artificial diphthongs using modern electronic equipment. In the end, such artificial diphthongs have to be interpreted by speakers of a specific language (usually English), so such studies are also not completely free from the 'tyranny' of being language specific.

In an attempt to determine which aspects of diphthongs are language specific and which aspects are invariant across languages, Lindau, et al (1990a, 1990b) examined diphthongs in Hausa, Arabic, Chinese, and English. (Specific dialects of each language were not specified.) They began with the assumption that diphthongs were

vowel sequences within a syllable, so their study has pulled on phonological information about each of the languages. In addition, they threw out the diphthong /ai/ in Hausa 'because this diphthong is produced as phonetic long [eː]' (1990a:11, cf. Lindau-Webb 1985). This certainly begs the question of what is or is not a diphthong phonetically. It seems that there is a similar circumstance in Cairene Arabic, which was also investigated by one of the co-authors of this study (Norlin 1984). 'Most standard Arabic diphtongs [sic.] /ai/ and /au/ have developed into long /ee/ and /oo/' (188).

If diphthongs can manifest themselves as phonetically long vowels, then perhaps diphthongs are not a phonetic phenomenon at all, but a phonological one. If this is the case, then the assumptions underlying such studies are muddled at best. The stated goal of the study is to find cross-linguistic similarities or differences between diphthongs. Given this goal, it seems that one of the underlying assumptions is that a diphthong may not even have two segments phonetically speaking. If such phonological principles can completely override any phonetic reality, then there is indeed no limit to the power of phonology.

Clearly, however, the authors believe that there is a phonetic reality to diphthongs or they would not have undertaken such a study. In their study, they specifically ask the question whether the rate of transition between the first and second steady state portion of a diphthong is steady cross-linguistically or varies cross-linguistically. This points out another underlying assumption to their article, which they do not specifically state; namely that they have accepted the definition of a diphthong

that assumes two steady state portions joined by a glide (see Figure 2.1). However, some of the diphthongs in their study do not seem to actually achieve a second steady state portion, particularly those spectrograms of the Chinese diphthongs (12). This leads to the conclusion that diphthongs can differ in this way cross-linguistically. In addition, they found that the transition rate between the onset and offset varied dramatically across languages, leading them to conclude that such transition rates were language specific and that no cross-linguistic generalization could be drawn.

Although Peeters (1987, 1991, 2000) has sought to establish a cross-linguistic definition of diphthongs, all of his work has been done among speakers of Germanic languages, namely German, Dutch, and English. He has argued, through a series of perceptual studies, that of the definitions represented by Figure 2.1 and Figure 2.2 neither is truly acceptable. He emphasizes the inclusion of time in the definition of vowels, diphthongs, and vowel clusters, pointing out that they all may express movement but they do so in different ways — ways that are significant within a language system and in ways that are perceptible to native speakers of those languages. For example, even though a phonemic long vowel might be diphthongized phonetically, it will be diphthongized in a way that is perceptually and acoustically different from 'true' diphthongs in the language.

Peeters (1987) declared, 'There are eye-catching differences between falling back diphthongs, diphthongized back vowels and back vowel clusters' (73). These eye catching differences can be seen in spectrograms of comparable clusters and

diphthongs. While the cluster /ao/ had gradual changing amplitude values, the diphthong had much more avalanche like changes in amplitude. Therefore, to discover the difference between similar segments in a language, the language has to be compared internally and the quality of the spectrograms has to be examined relative to the time factor. As the vowel progresses, the angle of formant changes (i.e. whether the changes are abrupt or gradual) seem to be the key to telling true diphthongs from diphthongized long vowels. Peeters readily admits that his research leaves many questions unanswered, but his research does seem promising. Although he has only tested speakers of a few languages, it seems likely that the overall premise is sound. It can be stated that speakers can indeed tell true diphthongs from pretenders, the only problem remains is to discover how they can tell that difference.

It is, indeed, likely that speakers are able to make language internal comparisons and therefore are able to define diphthongs by their unique timing aspects relative to other similar and yet unique phonemes in the language. This means that it is not absolute drift or change that makes a diphthong a diphthong. It is the rate of that change in diphthongs in comparison to other segments in the language that might also exhibit drift. Therefore, if a long vowel becomes diphthongized, it is still recognizable as a long vowel because its rate of drift is slower than that required to be defined as a diphthong. This is something that still remains to be quantified in a systematic way, but the pictures in the spectrograms clearly show these differences in rate of change, making the overall appearance of the picture more important than the actual onset and offset values of the

segments under analysis. It is possible to have a 'true' diphthong in a language and a diphthong-like segment in a language that have identical onset and offset values, but linguists would expect the difference between them to show up in the rate of change in the glide area between them or in the relative weights assigned to those onset and offset steady states. It is not enough merely to study the destination of a diphthong; one must study how it reached its destination, the path it took and how quickly it got there.

2.4 CONCLUSIONS

At the end of the day, what can be said definitively about diphthongs is relatively little. This does not mean, however, that nothing is known about them. Although the data is inconclusive, there are some general guidelines that can be useful in studying language. The first is that diphthongs do seem to change over time. There were many documented cases in which diphthongs turned into monophthongs and vice versa. Secondly, diphthongs seem best to be studied initially on words in isolation. This removes the effects of many prosodic factors such as rate of speech and register. Diphthongs spoken at a rapid rate of speech or in a casual register seem to rarely reach their second offset position (Wouters 2002). Thirdly, although rarely mentioned, intensity seems to be a factor in diphthongs as opposed to vowel sequences. While vowel sequences contain two intensity peaks (often called 'chest pulses' in older literature), diphthongs contain only one. Finally, the rate of transition seems to be important. Long monophthongs can exhibit drift and wander in such ways that they sound nearly identical to diphthongs, but native speakers can still tell one from the

other. In addition, the differences between long monophthongs and diphthongs can be seen in the rate of transition from the onset portion to the end of the segment under study. While this change is not something that has yet been mathematically codified, it can be seen clearly in spectrograms when one visually conducts pattern comparisons. Each of these points will serve as a guideline in the examination of Sarikoli diphthongs undertaken in Chapter 3.

CHAPTER 3

ACOUSTIC ANALYSIS OF DIPHTHONGAL SEGMENTS

3.1 THE PROBLEM

Given that there has been such widespread disagreement over the vowel systems in languages studied as thoroughly as English, German, etc, it should not be surprising that there is little or no agreement as to the nature of the vowel system in a language as understudied as Sarikol Tajik. While some researchers have posited nine diphthongs in Tajik (Gao 1963, Schwarz 1984) others have posited one (Peters 1996). Still others have posited some number somewhere in between these two extremes.

There does, however, seem to be widespread agreement over the number of simple vowels in Sarikoli. Sarikol Tajik has been said to have a balanced seven vowel system (Table 3.1). Although there is some disagreement on the mid-high vowels as to exactly which they are, the difference seems to be the impression of individual linguists rather than any actual disagreement. For example, some linguists always write the mid back vowel as [ɔ] while another always writes it as [o]. This indicates that any given single analysis of Tajik does not contain both of those vowels, so they can be considered simply as two ways of transcribing the same sound. The same is true for [ɛ] and [æ] in that, although both occur in different analyses of Sarikoli, there is no one

analysis in which they will both occur[1]. At this point, whether [o] or [ɔ] is exactly accurate phonetically speaking is not the issue. The important point here is to gain an understanding of the monophthongs in Sarikol before diving into the more difficult task of identifying the diphthongs.

Table 3.1 Commonly accepted monophthongs in Sarikol Tajik

i	ɯ	u
e		o/ɔ
ɛ\æ		a

The most common vowels in Tajik appear to be all of the high vowels, with the mid-central [ɯ] being the most common vowel sound in the language. The back central [o]/[ɔ] is more common than its front counterpart. There are also occasional transcriptions in which one will see [ə], but closer analysis reveals that they are all loan words that have yet to be fully nativized. When speakers are asked to pronounce words that had been transcribed with [ə] in them, they will often then admit that the word is really a Uighur word being used in Tajik or else they will then suddenly try to nativize the word and replace the [ə] with [ɯ].

---

[1] Although Table 1.1 shows Pakhalina and Peters and Peters as having both [ɛ] and [æ] in their analyses of Sarikoli, Pakhalina only includes [æ] as a dialect variation, which will be examined in Chapter 4. Peters and Peters only used [æ] rarely, in less than half a dozen words.

Having established the agreed upon facts about simple monophthongs, a closer examination of the claims concerning Sarikoli diphthongs is possible. Some of the diphthongs posited by researchers are unusual diphthongs that would certainly be marked in the languages of the world (see Table 3.2). Some of the diphthongs posited by previous researchers seem suspicious in their composition and also suspicious in their number. Some have even proposed that Sarikoli contains nine diphthongs, which is possible but would be unusual indeed. Moreover, diphthongs such as [ɯu] and [ɯi] (Gao 1963, 1985) are not impossible but are so marked cross-linguistically that they will certainly will require more support than merely the statement of the linguist that they exist. Indeed, in the diphthongs listed in the UCLA Phonological Segment Inventory, the Stanford Phonology Archive, and in Weeda (1983), covering 674 languages, such diphthongs are not attested even once (Miret 1998). In fact, the vowel [ɯ] appears to very rarely be involved in diphthongs in the studies reported below. The rarity of [ɯ] in diphthongs was reconfirmed by recent studies as well. For example, in her database of 42 languages carefully chosen to represent a cross-section of languages of the world, Sands (2004) found the diphthongs [ɯi], [ɯu], and [iɯ] only occurred in two languages. Such segments do seem to appear in some languages of Vietnam as well, but once again, such segments are still highly marked (Edmondson 2006).

Table 3.2 lists each of the proposed diphthongs in Sarikol Tajik and their number of reported incidences in each of the above studies. Knowing how common

each diphthong occurs cross-linguistically can be a helpful benchmark in forming hypotheses about each of them and their likelihood of turning out to be true diphthongs. Nevertheless, marked segments do occur in languages and the language itself has to be the final arbiter of its own vowel space.

Thus far, definitive proof about the exact vowel inventory of Sarikoli has either been impossible to obtain due to the limitations of the linguists themselves, the limitations of the technology of the period in which their research was conducted, or because of lack of access to native speakers of Tajik. For example, Pakhalina (1971) and Gao (1963) did their research at a time when acoustic studies were possible but were not as easy to conduct (due to the lack of laptop computers) as they are now. Peters and Peters (1996) could have done such a study had they been interested to do so, but their interests were in other areas and they were not trained in how to use such equipment or in how to conduct acoustic phonetics studies. Thus, no precise empirical study of the problem has ever been conducted. Researchers have merely relied on their auditory impressions. And so, the question has remained: How many diphthongs does Sarikoli have?

Table 3.2 A cross-linguistic comparison of the number of diphthongs in three studies
with the diphthongs proposed for Sarikol Tajik by previous researchers

Proposed Sarikoli Diphthong	UCLA (451 languages)	Stanford (197 languages)	Weeda (26 languages)	Total
ai	19	6	7	32
ui	8	4	5	17
ei	7	3	3	13
ou	5	4	4	13
ao	4	7	2	13
oi	6	3	2	11
iu	2	3	4	9
εi	3	1	2	6
εu	3	1	2	6
əu	4	0	1	5
eu	3	1	1	5
ɯi	0	1	0	1
ɯu	0	0	0	0

Note: In the UCLA study, the diphthong [εu] was not in the inventory, but the diphthong [eu] was. They were close enough that they were counted as being the same here. In the Stanford study, it was difficult to separate the semi-vowels [w] and [j] from the vowels [u] and [i]. For the purposes of this study, they were considered the same. So that, if the Stanford study contained the sequence [oj], it was counted the same as [oi], etc.

3.2 METHOD

To answer the question of how many diphthongs Sarikoli has, new recordings of native speakers of Sarikol Tajik were made. Recordings were made in July 2004 in three Tajik villages: Tashkorgan, Burungsal, and Vacha (Figure 3.1). These three villages were chosen because Pakhalina (1971) had stated that there was a slight dialect variation between them. Due to strict border controls in the area (it borders Afghanistan and Kashmir), foreigners are not allowed in two of these villages and any kind of unusual activity is highly suspicious in the other. It was decided, therefore, to hire a local Sarikoli to make the recordings. He went to the three villages and interviewed five speakers in each village. He was instructed to find speakers of many ages and of both genders. In the end, however, all of his interviewees were older men. It seems that many of the younger people had left the villages and gone to the higher grazing grounds for the summer and women were generally too shy to participate (especially with a man doing the interviewing). All the men were herder/farmers of a moderate economic status. That is, they had enough income to survive in their village but would not be considered unusually wealthy farmers. Although this homogeneity has some disadvantages, it also has the advantage that all of the men were basically of the same age and socio-economic status. This means that, if there are any differences in their phonologies, the differences are not likely to be due to socio-economic differences.

The computer seemed too risky to take to such remote villages where electricity was likely to be unavailable or to have sudden power surges. In addition, the Tajik who

went to conduct the interviews was not used to using computers and found it too difficult. In the end, he was equipped with a high quality hand-held tape recorder with a uni-directional microphone and a package of batteries. He did a good job of placing the microphone correctly and the recordings were ultimately of good quality and all were used. Later, the tapes were digitized and analyzed using Praat.

Figure 3.1 Map of Tashkorgan Tajik Autonomous County showing the location of Tashkorgan, Vacha, and Burungsal

Before going to the field, the interviewer was trained briefly in how to conduct the interviews. The purpose of the recordings was explained and a word list was generated. Respondents were told a word in the Uighur language, a Turkish language spoken by many Tajiks, and then were asked to say the word twice in Tajik. The words were, therefore, pronounced in isolation with no frame. The use of the Uighur language meant that only those Tajiks that could speak at least a little Uighur could participate in the study. This would effectively eliminate a few women, but since the interviewer was a man and women were too shy to talk to him anyway, this had little effect on who could participate.

The words that were elicited were chosen based on a variety of factors. Since Gao (1985) had proposed the largest number of diphthongs, almost all of his sample words for those diphthongs were used. There were a few words on his list that no Tajik could recognize during the pre-field preparation time, so those words were thrown out. Otherwise, all of his example words were used. Schwarz (1984) has also proposed a fairly large number of diphthongs, so his example words were used as well. Other words were chosen based on pre-study interviews with Tajiks and through comparisons between the published studies. A complete list of the words used in the study can be seen in on page 64.

A comparison of Tables 3.2 and 3.3 reveals that there are some discrepancies between the two lists. For example, some proposed diphthongs occur in many more of the test words than others. Some of the diphthongs, such as [iu] only occur in one test

62

word. One diphthong that was proposed by Gao, [əu], was not present in any of the test words. This situation was unavoidable. Some diphthongs are more widely used than others and some of the diphthongs that had been proposed were rejected by field informants used in this study. Even when pressed, they insisted that Tajik did not have such a sound. It is hard to tell if this situation occurred because of a mistake by the original field linguist or if there has perhaps been some change in pronunciation over time. In any event, some sounds were simply not able to be included in the study or were not included in as large a number of contexts as would be ideal. In addition, some words were simply not elicited correctly by the interviewer. For example, Tajik has two words for 'mood': [mɯdʒuz] and [χɯi]. It was hoped when he went to the field that the Uighur word he used would elicit the second word. Unfortunately, all the Tajiks in the survey said the first word instead, most likely because that Tajik word and the Uighur word are the same. This meant that some potential instances of diphthongs were lost due to field constraints.

It is difficult to make comparisons across studies. Different researchers have often transcribed the same word in many different ways. For example, the Tajik word 'you' has been alternately transcribed as: [tɔu], [tɔw], [tɛu], [tɛw], [tow] and [tou]. Other words also had similar numbers of transcriptions. Likewise, the diphthongs [ei] and [ɛi] are so close phonetically that it is unlikely that all researchers were able to tell them apart. There also appears to be a process of palatalization that sometimes occurs (at least in some speakers), leading

Elicited words grouped by potential kind of diphthong

(1) a. [ai] [sair] 'to be full'   (4) a. [ou] [tou] 'you'

b. [tsaiz] 'what'   b. [ʁoul] 'ear'

c. [nai] 'no'   c. [zou] 'grain'

(2) a. [ui] [dʒui] 'place'   d. [ðou] 'two'

b. [bui] 'cave'   e. [doulat] 'country'

c. [adʒuib] 'strange'   f. [bou] 'smell/odor'

(3) a. [ei] [peiʃin] 'evening'   (5) a. [oi] [poi] 'to herd'

b. [tʃarein] 'man'   b. [poi] 'yoghurt'

c. [χeil] 'kind/type'   c. [boi] 'rich'

d. [tʃeig] 'to do'   d. [noi] 'flute'

e. [beil] 'shovel'   e. [xoid] 'to read'

f. [deig] 'pot'   f. [vijoid] 'to ride'

g. [deikun] 'farmer'   (6) [iu] [iu] 'one'

h. [waxein] 'blood'   (7) [ɛi] [spɛid] 'white'

i. [baweid] 'to disappear'   (8) a. [ɛu] [njɛu] 'nine'

j. [indeid] 'to stand up'   b. [vrɛu] 'eyebrow'

c. [tʃabɛud] 'pigeon'

some linguists to transcribe 'nine' as [njɛw] while others only transcribed [nɛw]. In addition, many of the [i] and [u] segments transcribed above were transcribed as semi-vowels by some previous researchers but as vowels by others. For the sake of simplification, they are all written as vowels here but this should not be seen as making

a definitive statement as to which kind of segment they are. As a result, the transcriptions that appear in should only be taken as a guideline for the sound that occurs in that particular word.

Words from different speakers and different locations were compared to see if, when spectrograms are considered, any of them reveal any kind of pattern that would be like a diphthong. That is, is there a moment of rapid transition from the onset to the offset portion of the vowel? The previous research examined in Chapter 2 agreed that diphthongs contain a glide portion (i.e. a moment of rapid change) but disagreed about whether there was an obligatory transition target. Some, such as Ladefoged (1996), argued that diphthongs need merely to have some movement. Others thought this definition too broad, as it could not distinguish between long monophthongs and diphthongs. Although previous researchers did not agree on the necessity of the existence of an onset and offset portion, there was widespread disagreement that the glide/transition segment had to be present - even if the exact nature of that transition was not agreed upon. The only researchers to disagree on the necessity of the glide portion were the ones (like Ladefoged) who argued that diphthongs need merely have some movement -- any movement at all. However, other research (as reviewed in Chapter 2) revealed this definition to be too broad and not able to distinguish between long monophthongs and diphthongs. Therefore, the glide portion may be considered a *sine qua non* feature of diphthongs that, if lacking, signals that the segments under consideration must not be diphthongs but, perhaps, 'drifting long vowels'.

In recent years, there has been much discussion in phonetics/phonology about whether the traditional model of discrete categories of segments is best able to account for phonetic data (e.g. Rechziegel 1998, Sole 2003, Taylor 1989). Previous phonetic studies focused on the invariant internal representations of phones. In this view, a segment is composed of invariant correlates (Rechziegel 1998, Scarborough 2005). In this view, a diphthong would be argued to always match with one and only one of the definitions looked at so far in this study. However, recent work in phonetics has challenged the idea that there are invariant phonetic correlates of a phone. For example, it has been established that so-called canonical forms of vowels are rarely reached in normal speech (Boersma 2005, Rechziegel 1998). This term was coined as early as the 1960s to account for the problem of the mismatch between canonical/target vowels (i.e. the articulatory aim of a vowel) and the phenomenon of target undershoot (i.e. actual realization of a vowel). The recognition of the difference between canonical vowel and their actual articulation in everyday production has led some to turn to probabilistic or prototype models of vowels (Kuhl 2000, Sole 2003). Thus, much recent work in phonetics has focused on gradience and variability in data suggesting, 'that phonetic variability itself may be part of the make-up of the phonological categories….By exposure to multiple instances of a category we learn the modal values and the permissible parameters of variation' (Sole 2003:289).

Thus, the concept of prototypes has become increasingly useful in phonetics in helping to define the areas of permissible variation in phonetic data prototype theory

66

was originally applied to color terms and semantic variation (MacLaury 1991). Prototypes associate a word or segment with 'a prelinguistic, cognitive schema or image; and…speakers are equipped with an ability to judge the degree to which an object…matches this prototype schema or image' (Coleman 1981:27). Thus, phonetic events can now be categorized gradiently rather than on absolute set membership. Segments can graded in this way can be said to be more or less like a prototype. This is in keeping with what is known about phonetics in general and diphthongs in particular. For example, several researchers pointed out that diphthongs failed to reach offset targets in their realization of diphthongs. Yes, researchers still felt that the 'best' instantiation of a diphthong would contain such an offset steady state. It seems that most researchers have a model of a prototypical diphthong in mind when conducting research. This can be seen in the widespread underlying assumptions in the studies above and also in the approaches taken to diphthongs. Given that there is such a strong theoretical predisposition to understanding a diphthong but such lack of actual instatiations of that definition need not be troublesome if prototype theory is used. Indeed, the used of probabilistic or prototype driven models of phonetics seems to be the current overall direction/trend (MacLaury 1991, Scarborough 2005, Scarborough 2006, Sole 2003).

Thus, in this study, prototypical diphthongs were defined as vowels that exhibited steep vectors, which indicate rapid movement, in the glide/transition portion and containing on onset and offset steady state. Thus, judgments of whether a segment

67

ought to be considered as a diphthong were made by taking the schema of a diphthong exhibited in Figure 2.1 and comparing spectrograms to that schema. That schema was taken as a prototypical instance of a diphthong and then actual formant results were compared to that prototypical instantiation. It is also necessary to establish prototypes of other segments, such as long vowels. Long vowels prototypically are long in duration and exhibit drift in formant vectors over time as seen in shallow trajectory slopes. Therefore, a segment in consideration here may match either one or neither one of these prototypes. It is possible that some segments will match the prototype and therefore be relatively easy to categorize while other segments may simply be a 'less typical' instantiation of one of the prototypes. Thus, each segment in this study is compared to two prototypes – both the prototype for a diphthong and also the prototype of a long vowel. (All measurements of length are made from the first to the last glottal pulse on the waveform.)

Before proceeding further, it is worth noting that the analysis of speech formants (represented by F1 and F2) to identify vowel space is in accordance with the classic target model of analyzing vowels. A formant is generally defined as 'a resonance of the vocal tract' (Pickett 1998:37). When approaching vowels from the classic target model, vowels are characterized by their location in the vowel space through the identification of F1 and F2 values. F1 identifies tongue height, with high/close vowels having lower F1 values. F2, on the other hand, identifies vowel frontness or backness (also rounding), with front vowels having a relatively higher F2 value. The basic

information provided by formant values is generally accepted in linguistics, but the classic model is not without controversy. Some have argued that listeners must normalize F1 and F2 values to account for the different vocal tract lengths and sizes of individual speakers. For example, speakers with relatively larger heads also have large resonating cavities and therefore have lower overall frequencies in their vowels. Thus, they argue that non-normalized F1 and F2 values are not adequate to characterize the acoustic properties of vowels. Such theories are generally categorized as an elaboration on the classic target model (See Kent p. 109-110 for a discussion of competing theories). In the end, much recent analysis 'lends support to formant pattern[s] as a primary cue for vowel perception' (Kent 2002:110). Moreover, the literature review in Chapter 2 further supports that analyzing F1 and F2 values is still widely accepted and used in most studies of diphthongs. This can be seen in the fact that the classical approach is the one taken by all the studies examined in Chapter 2. Finally the classical approach of describing vowels in terms of a point in the F1 and F2 plane has been bolstered by recent approaches to acoustic phonetics (e.g. Kent 2002, Ladefoged 2001, Pickett 1998, Stevens 1998); that is the approach taken here.

It deserves to be emphasized here that the following results say nothing about the phonology of Sarikoli. It is likely that there are differences between Tajik speaker's phonological understandings of these words and the phonetic descriptions that are offered below. For example, the word 'strange' appears to be a diphthong [adʒɯib] phonetically. However, if native speaker intuitions are any judge at all, then the word is

phonologically /adʒuib/ or perhaps even /adʒujib/[2]. When asked to write the Tajik word (using the Uighur alphabet, the only one available), Tajiks often used one of those two spellings. However, phonetic evidence is clear that the first vowel in the diphthong portion does not reach that back position. Since Tajik phonology is left for a later investigation, this is not of concern to this study.

3.3 RESULTS

This section reports the results for each of the diphthong-like elements reported in (3.1). Each potential diphthong is reported in turn in the order in which they were presented there. Although each of the words under study was examined in each of its 15 instantiations, each instantiation of the diphthongal element was usually very similar or identical to the others. Only when such differences were interesting or deemed to be significant are they mentioned separately.

3.3.1 [aḭ]

Three words were reported to potentially contain the diphthong [ai]: [sair] 'full', [tsaiz] 'what', and [nai] 'no'. Spectrographic analysis of each of these sounds across all fifteen speakers in this study reveals that the elements in these words are not the same. While [sair] 'full' and [tsaiz] 'what' seem to show merely drift, with no rapid change in the glide portion of the vowel element, [nai] 'no' is different and does exhibit steep vectors in the glide portion of the diphthongal element.

---

[2] This second understanding is possibly an influence of the Uighur word 'strange', which is cognate with the Tajik word. The word originally comes from Arabic. The Uighur word is spelled /adʒajip/ in the Uighur orthography, which might account for why some Tajiks feel that there is a semi-vowel in the word.

Formant tracks of [nai] 'no' can be seen in Figure 3.2. The pattern for [nai] 'no' is exactly what is expected from a diphthong, with rapid change in both F1 and F2 exhibited in the glide portion. It can be seen that just after the initial nasal [n], the vowel reaches an onset steady state. The nasal is represented by the low frequency resonances of the nasal passages. Nasals do not normally create release transients but will often cause the transitions between segments to become nasalized. In this case, the initial F1 amplitude of the vowel portion begins low due to the coarticulation of the nasal when transitioning into the vowel. Nasal coarticulation of vowels has been shown to lower F1 (Scarborough 2006) and also to dampen the amplitude of F1, which accounts for the lighter patches in the F1 spectrogram (Pickett 1998). The relatively level onset steady state is followed by steep vectors as the vowel enters the glide portion and then moves to an offset steady state. All in all, a nearly perfect match to the prototype schema of a diphthong.
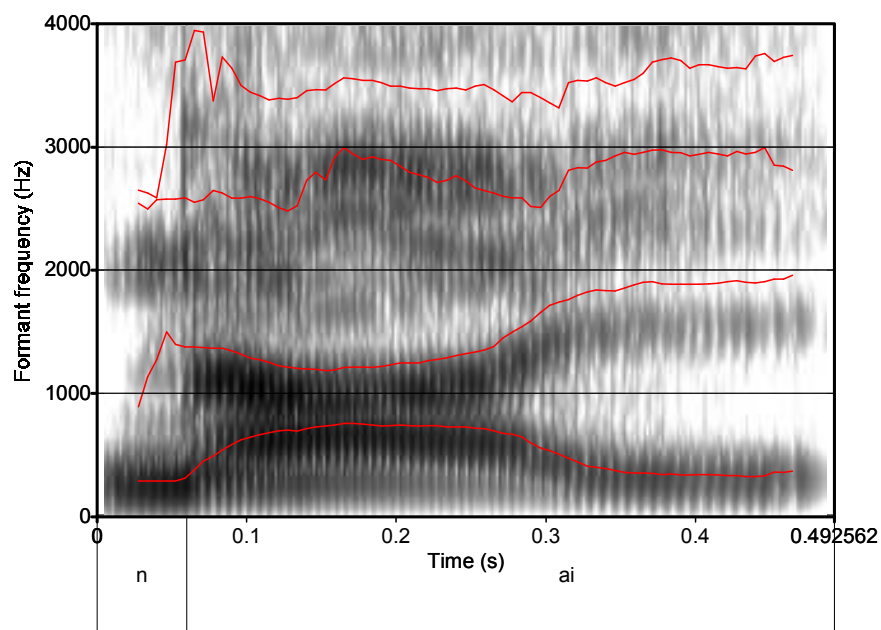
Figure 3.2 Spectrogram and formant tracks of [nai] 'no'

On the other hand, formants for [tsaiz] 'what' show change but they do not show prototypical diphthongal change. (The acoustic differences between diphthongs and semi-vowels are summarized in Table 3.3 for reference.) Figure 3.3 shows a vowel of nearly identical length to that seen in Figure 3.2 (also just over 400 milliseconds) but the shape of those formants is quite different than what was seen in [nai] 'no'. There is some movement in the formants, which is likely what accounts for the perception by some that it is a diphthong, but the nature of that movement does not show locally focused and rapid changes of the F1/F2 vectors in the glide portion of the vowel segment as in Figure 3.2.

Table 3.3 Main acoustic features of diphthongs and semi-vowels compared
(Adapted from Pickett 1998:112)

Features	Diphthongs	Semi-vowels
Timing Oral Articulation	Slow	Medium-fast
Type of Constriction	No narrow constriction	Strong

Spectral
   Intensity
   During
   Constriction
	No narrow constriction	Strong

Spectrum
   During
   Constriction
	No narrow constriction	Low frequencies strong up to about 600 Hz; mid-frequency energy weaker than in diphthongs

Formant transitions	Slow transitions between the two component vowels	Medium-fast transitions appropriate to the place, lateral, opening, or retroflexion
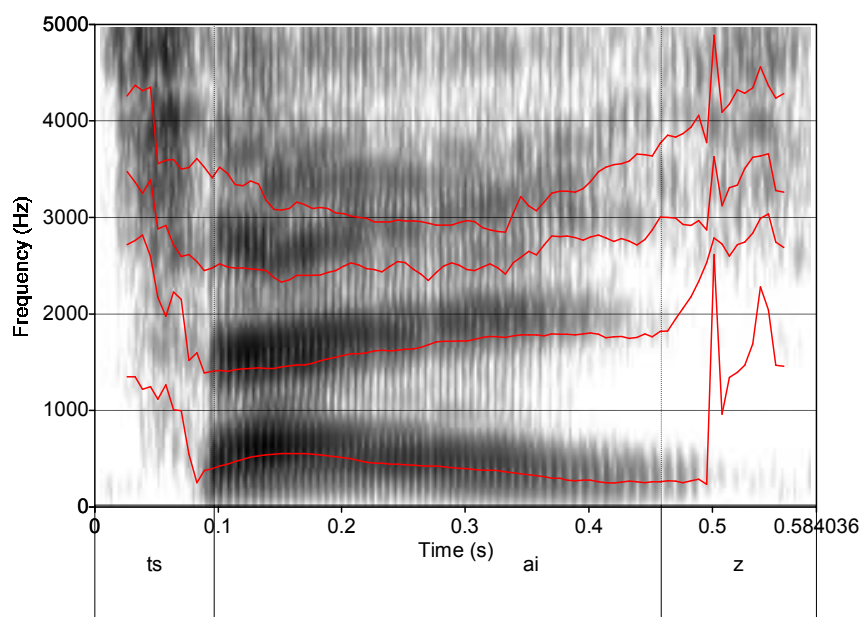
Figure 3.3 Spectrogram and formant tracks of [tsaiz] 'what'

When matched against the prototype of a diphthong, its difference is clear. Thus, it cannot be considered a diphthong. However, when matched against the prototype of a long vowel, the match is good. The segment appears to be a long vowel that drifts but does so with shallow formant trajectories, and indeed does drift into the 'neighborhood' of [i], accounting for why it has been perceived as a diphthong by previous researchers. However, spectrographic results are consistent in revealing a long vowel instead.

Although not pictured here, results for [sair] 'full' were similar to those for [tsaiz] 'what'. While the vowel showed drift, it did not exhibit any abrupt change in the glide portion of the vowel, therefore leading to the conclusion that it best matches the prototype of a long vowel.

Therefore, it can be said that the diphthong [ai] does exist in Tajik, but not all words that had been postulated to contain the diphthong actually did. While one word did indeed include a diphthong, based on spectrographic evidence, the other two did not. Such a result indicates that both [ai] and [aː] exist in Sarikoli. Both the diphthong and the long vowel exhibit a change whereby the end of the vowel moves into the to the vowel space of [i], but they arrive there in a different fashion. Therefore, the difference is not in their respective onset and offsets steady states, but in how they reached those two states. There is no denial that their offsets essentially end in the same place, but the articulatory mechanism that propelled the vowel into that space was different. In the diphthongs, the segments exhibit two steady states with a transitionary glide that

74

exhibits rapid change as seen in the steep vectors of the F1/F2 values. On the other hand, the long monophthongs show movement, but no onset and offset steady states and the vector of the glide segment is level. Just as the manner of articulation in consonants is an integral part of their definition, the manner or articulation in vowels is critical as well.

Further work could determine how different native speakers perceive these sounds to be (if at all), but here the concern is with phonetic 'reality' and not with phonology. Future chapters will examine why such sounds might be so similar and yet different. (For example, there is likely development in progress.) For now, however, the evidence points to two distinct sounds with two distinct articulations in this analysis of contemporary Sarikol Tajik.

3.3.2 [ui]

The proposed diphthong [ui] was also said to occur in three words: [dʒui] 'place', [bui] 'cave', and [adʒuib] 'strange'. The cases of [dʒui] 'place' and [bui] 'cave' each seem indeed to an instance of the category of diphthongs, exhibiting the typical change in the glide portion seen in diphthongs. Both show steep vectors in the glide portion of the vowel and then show a steady state as shown in Figure 3.4. In the case of [dʒui] 'place', after the formant transitions created by the affricate [dʒ], a stable formant pattern corresponding to the onset vowel develops. The affricate portion exhibits the random striations and also the weak low-frequency sounds and the relatively stronger high frequency sounds (Pickett 1998:152). Fricatives (in particular

voiced fricatives) also create longer formant transitions than stops, accounting for the relatively long transition time seen at the beginning of the vowel portion (Pickett 1998:120). Finally, at about 375 milliseconds after the onset vowel has already reached a steady state, it then abruptly transitions to the offset steady state, as exhibited by the steep vector in the glide portion of the diphthong. The offset steady state is not as steady, exhibiting some slight variation in form, but is nevertheless clear. The overall length of the diphthong is quite long at 424 milllliseconds. Due to the match to the prototype schema as seen in its steep glide vectors, the segment under consideration seems best characterized as the diphthong [ui].
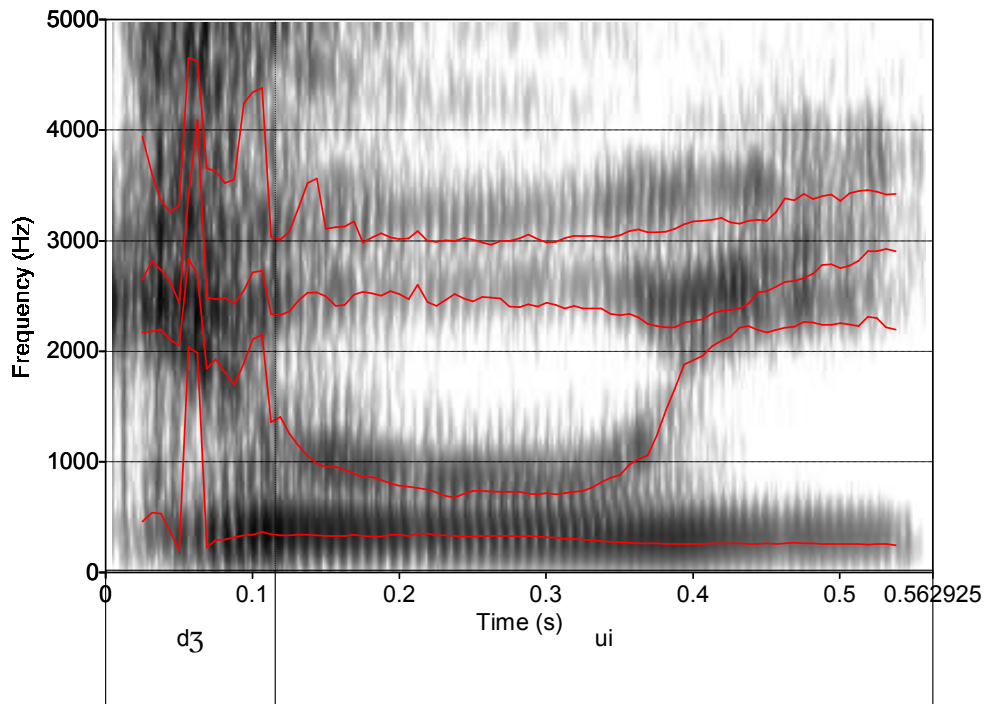
Figure 3.4 Spectrogram and formant tracks of [dʒui] 'place'

Although the case of [ui] might seem straightforward looking at two such words, not all words that were proposed to contain this diphthong were so easy to analyze. In particular, the case of [adʒuib] 'strange' is less straightforward. This is because if this word contains a diphthong, then it seems to be qualitatively different from the diphthong found in both [dʒui] 'place' and [bui] 'cave'. This difference can be seen in Figure 3.5, in which the vowel segment shows a pattern similar to that seen in Figure 3.4 - but not exactly the same. They are similar in that they both seem to have that same rapid transition in the glide portion from the first part of the vowel to the second part of the vowel. They are also both somewhat long, with the segment in Figure 3.5 being 336 milliseconds. The main difference is in which part of the vowel segment is long -- the onset or the offset. In all previous occurrences of Sarikoli diphthongs, the onset has been relatively long and the offset has been relatively short. That is, if the onset and offset length are compared relative to each other, the onset is shorter and the offset is longer. Here in Figure 3.5, however, that is not the case. The onset is noticeably shorter than the offset.
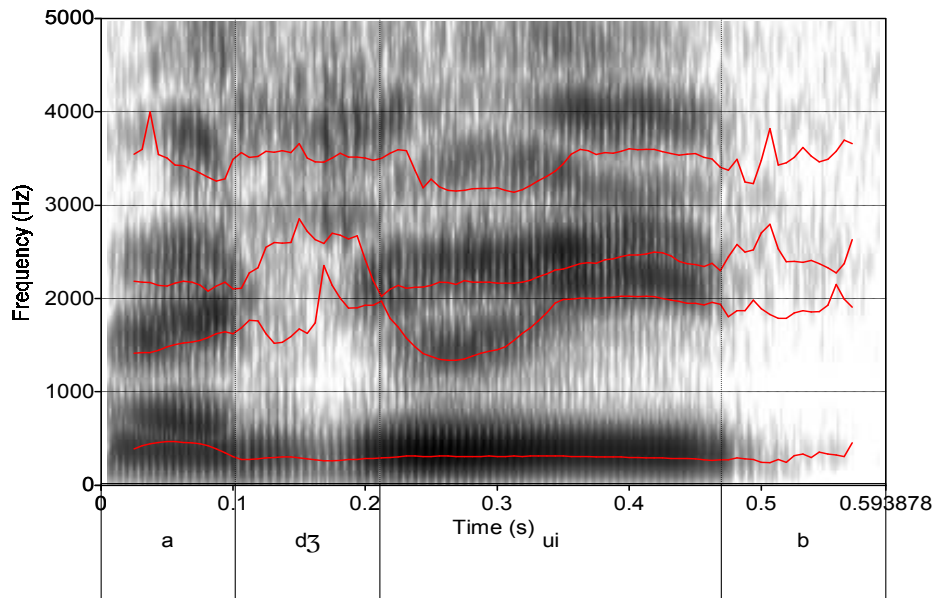
Figure 3.5 Spectrogram and formant tracks of [adʒuib] 'strange'

If this were the only difference, it might be easily overlooked, but some speakers seemed to show an even more marked difference. They seem to have an onset that was so short as to be almost unnoticeable. Such a speaker can be seen in Figure 3.6. In this case, there is a slight dip in F2 at the onset, but it is very slight. In addition, that portion of the vowel is quite short. Interestingly, a larger change can be seen in F3 than in F2 in this case, a formant that has not seemed to play a significant role in Sarikoli vowels thus far. Given an F2 frequency of around 1500Hz it is clear that the onset of this vowel never reaches a back position at all. Its relative instability is probably accentuated by the relatively long transitions brought on by the preceding voiced affricate, but even that does not seem to totally account for the instability exhibited in the onset F2.
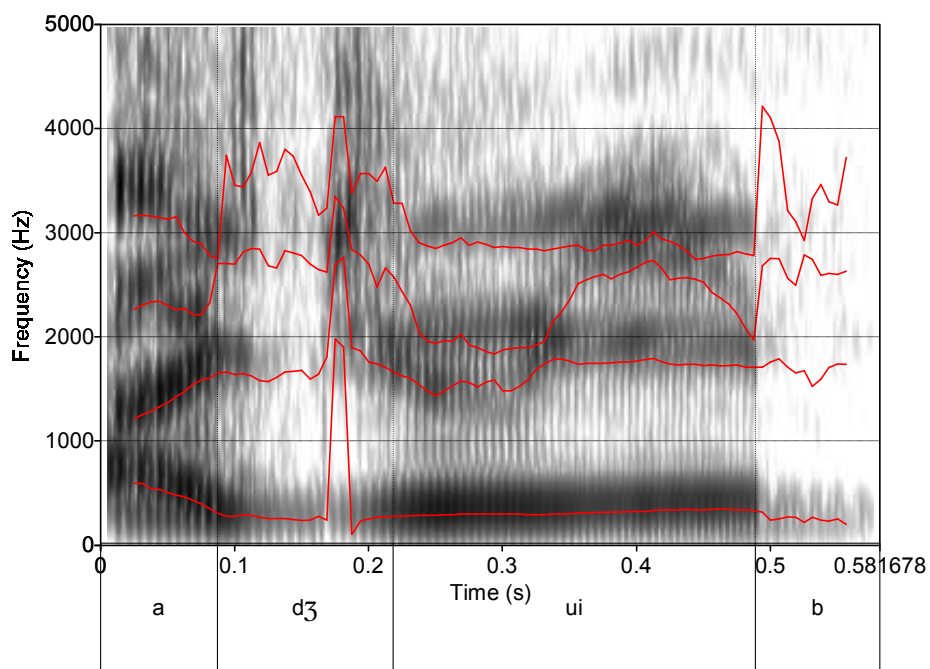
Figure 3.6 Spectrogram and formant tracks of [adʒuib] 'strange' exhibiting little change between onset and offset steady states

What then can be said of this segment? Is it a diphthong or not? In some speakers, it seems rather diphthong-like, but in other speakers, it does not seem diphthong-like. It seems that the most appropriate solution is to call it a less prototypical diphthong for now, but to also acknowledge that it is likely in a state of transition. This word will certainly bear watching in the future. It is possible that a diphthong is gradually being re-segmented into a long vowel by some speakers. It seems that it is, however, too early to already begin to call it a long vowel. Although the glide section seems weak for some speakers between the onset and the offset, it is nevertheless there. In addition, such speakers are currently in the minority, but minority speakers should

79

not be dismissed out of hand. Variation can sometimes provide valuable insight into past or future development of a language (Anderson 1972, Bailey 1996). Therefore, although it is possible that this segment is simply part of one speaker's idiolect, it could be more than this. It is important to remember, however, that some linguists (Gao 1985) have proposed the diphthong [ɯi]. However, in pre-field interviews, no Tajik could think of a word that might have such a sound in it. Could it be that there are such segments in Tajik after all? Perhaps they exist but they are merely transitional forms from the diphthong [ui] to a long vowel. If looked at in isolation, it certainly could present an argument for including the diphthong [ɯi] in the Tajik inventory. Taken alone, however, such a conclusion is premature, especially in light of its minority status.

Given the overall picture provided by the spectrograms, then -- even including the less than prototypical segments in [adʒuib] 'strange' -- it seems that Sarikoli can be said match the prototype of the diphthong [ui]. There are unambiguous cases where the transition rate is clear and the steady states are unambiguous. It might be possible, given more evidence, to posit a transitional diphthong of [ɯi] or perhaps to posit a long vowel [uː], but that will require further research.

3.3.3 [ei]

Of the 11 words from page 64 that were reported to contain the diphthong [ei], results were not uniform. Six of the words show some drift, but nothing like the steep glide portion expected from a diphthong, leading to the conclusion that they exhibit sufficient deviation from the prototype of a diphthong to disallow such a categorization.

Although those words matched the prototype for a long vowel, four of the words seemed to contain simple monophthongs, neither showing any significant drift nor being long. One word, [peiʃin] 'evening' was ambiguous among different speakers and showed a great deal of variation. Each of these results will be examined, with an examination of [peiʃin] 'evening' at the end.
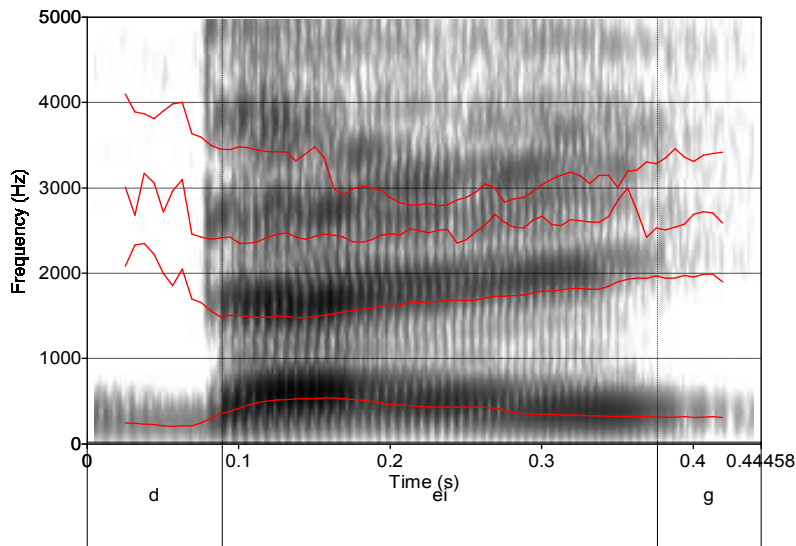
Figure 3.7 Spectrogram and formant tracks of [deig] 'pot'

First, this section will examine the six words from page 64 that appear to be long vowels with releases that are dynamic, but are not diphthongs. Those words are [tʃarein] 'man', [tʃeig] 'do', [beil] 'shovel', [deig] 'pot', [baweid] 'to disappear', and [indeid] 'to stand up'.

The formant tracks in Figure 3.7, [deig] 'pot', provide an example of the typical formant movement of words exhibiting drift. This word is a useful example since the vowel segment is flanked by stops on both sides. Stops usually produce only very short formant transitions in adjacent vowels and have abrupt endings, making it relatively easier to separate the vowel segment from the flanking consonant segments (Pickett 1998:120). The beginning, with the complete lack of energy other than the fundamental frequency (indicating voicing) is the [d]. The end of the word shows a similar lack of energy as the stop [g] is articulated, though there is some residual articulation of the vowel as the energy of the vowel dissipates into the lack of energy indicative of the stop. This leaves a clear and easily seen middle portion containing the vowel under consideration. As can be seen, there is movement in the long vowel section, but the movement exhibits a nearly level vector and does not exhibit any of the rapid movement from an onset to an offset steady state necessary in a prototypical diphthong. Although its formant structure does not indicate a diphthong, its duration is long, with the entire vowel segment of this word occupying nearly 400 milliseconds, quite long for a vowel. Across all 15 speakers of [deig] 'pot', the average length of the vowel was 265 milliseconds. The other five words ([tʃarein] 'man', [tʃeig] 'do', [beil] 'shovel', [baweid] 'to disappear', and [indeid] 'to stand up') exhibited both similar formant patterns and similar vowel lengths, with vowels being between 240 and 300 milliseconds in length. Based on the lack of a glide/transitionary segment in the vowel

portion, it seems that these words do not contain a diphthong, but contain the vowel [eː].
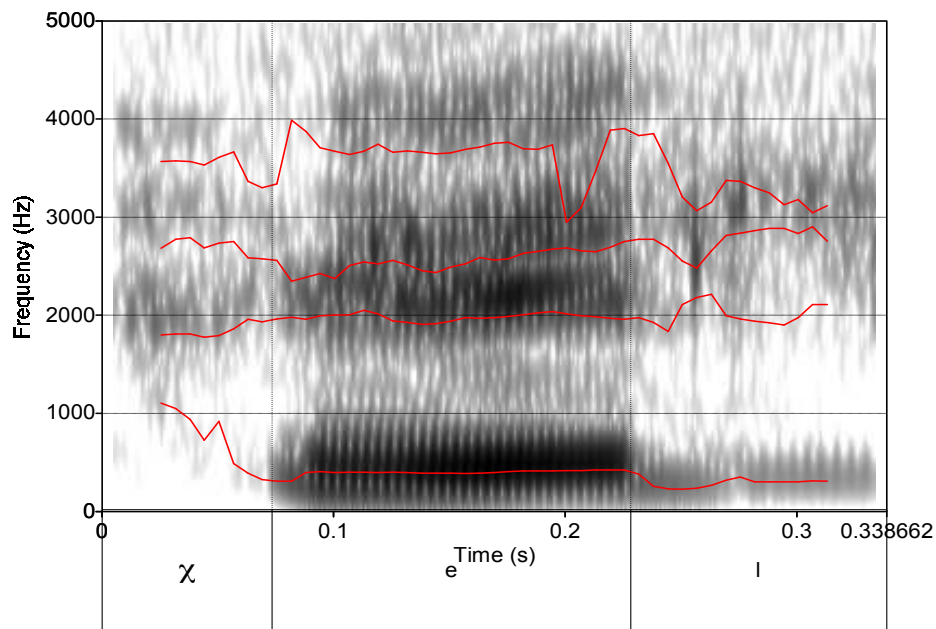
Figure 3.8 Spectrogram and formant tracks of [χeil] 'kind/type'

Although most of the words in this study matched the prototype of a long vowel [eː], three of the words reported to contain the diphthong [ei] seemed, instead, to contain a short monophthong [e]. These three words are [χeil] 'kind/type', [deikun] farmer, and [waχein] 'blood'. These words both lacked the rapid movement required to be a diphthong and the length required to be a long vowel.

A typical example of words of this type [χeil] 'kind/type' is shown in Figure 3.8. This spectrogram, shows a fairly clear demarcation between the onset and offset

consonants and the vowel under consideration. The onset [χ] shows the random striations of a voiceless fricative in the mid- and high-frequency areas and lacking any transient upon release (Pickett 1998:131). Likewise, the end of the word shows some of the weaker formant structure along with a glottal pulse, which is the [l]. Laterals generally take their formant positions from adjacent phones, as can be seen here. In particular, laterals in final positions often are allowed to gradually die out in intensity (Pickett 1998:110). The portion that contains the vowel shows strong, level, short formant structure. In this case, the entire word is 338 milliseconds, with the vowel portion being 155 milliseconds in length, much shorter than the lengths observed for the vowels in the six words above. Across all speakers for [χeil] 'kind/type', the average length of the vowel portion was 144 milliseconds. The other three words ([deikun] farmer, [waχein] 'blood', and [qiv tʃeig] 'to call') showed similar vowel lengths. Thus, these four words are both qualitatively and quantitatively different than the six words above. They are qualitatively different in that their formant structures are level, showing neither drift nor any rapid glides in the vowel portions. They are quantitatively different in that the average lengths of the vowels under consideration were much shorter. Thus, these vowel segments ought to be considered as simple monophthongs.
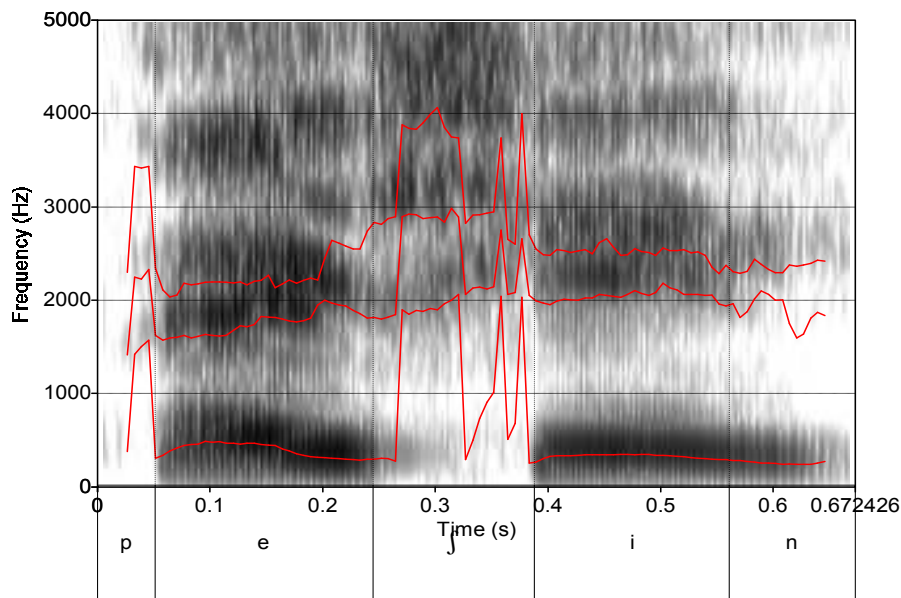
Figure 3.9 Spectrogram of [peiʃin] 'evening' showing a monophthong

Finally, the word [peiʃin] 'evening' deserves to be discussed in some detail because results across speakers were not identical. Among the 15 speakers included in this study, 12 of them seem to use a monophthong in this word. A typical spectrogram is shown in Figure 3.9. That spectrogram has been segmented to show the boundaries of each of the phones. The section labeled [e] is the section under consideration here. The word begins with the absence of energy characterized by the voiceless consonant [p]. The relatively low frequency striations (at less than 3000Hz) is indicative of the aspiration associated with [p] (Ladefoged 2001:51). Generally, voiceless stops will have relatively little coarticulation effect on following vowels as the positioning of the lips for such vowels normally takes place during the aspirated portion of the voiceless stop (Ladefoged 2001:51). It can be seen that the vowel uttered here shows some drift, but

not as much as some of the previous long vowels. Also, the length of this vowel is only 163 milliseconds, which is in the range of the short vowels rather than the long vowels studied above. If this were typical of all speakers, it would be fairly straightforward to declare the vowel in this word simply a short monophthong. A simple, short monophthong, however, was not typical of all speakers in this study.

Of the three remaining speakers in this study, it appears that the vowel they uttered in [peiʃin] 'evening' was a vowel followed by a semi-vowel (cf. Table 3.3). Instead of the relatively brief vowel segments with level vectors and no identifiable glide portion seen in Figure 3.9, these speakers exhibit a vowel segment that seems to have two portions, with the offset portion being a semi-vowel. That is, the segment that had been reported to contain a diphthong in [peiʃin] 'evening' seems to be [ej]. Figure 3.10 shows a spectrogram for one of the three speakers whose vowel portion seems to include a semi-vowel. Like the other speakers, these three speakers exhibit vowel segments that also begin with a vowel [e]. However, this part of the segment has a longer duration, here 205 milliseconds, when including both the segment labeled [e] and [j] in Figure 3.10. Duration alone indicates that something is different about these speakers' pronunciations than those examined earlier.
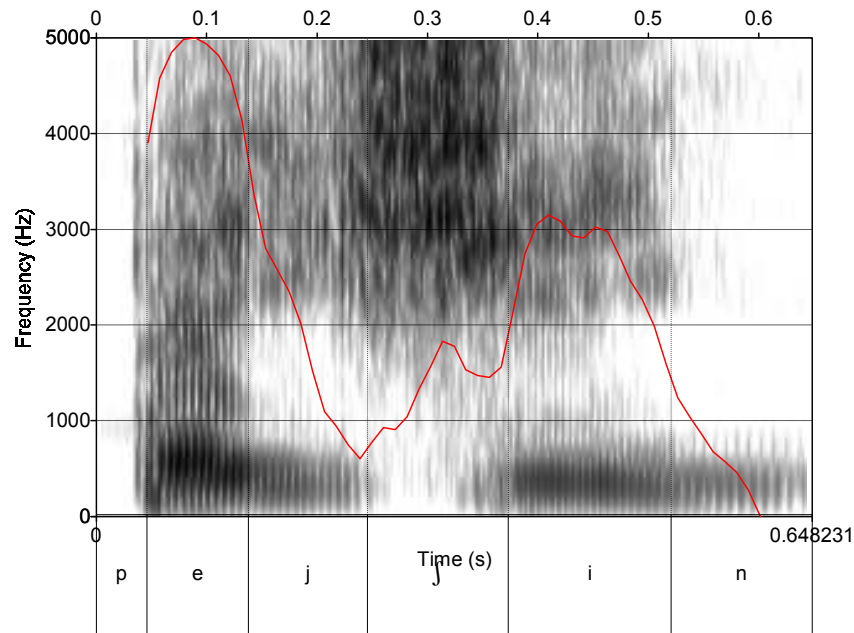
Figure 3.10 Spectrogram of [peiʃin] 'evening' in which the second
part of the vowel decreases in intensity

Even more telling, however, is the decrease in intensity in the second part of the
vowel section. This loss of can be seen in both the spectrogram and also in the intensity
contour that has been superimposed over the top of the spectrogram. The peak of
intensity is in the beginning part of the vowel, which is to be expected. In fact, the
intensity is so high, that the intensity contour can barely be seen near the top of very
black formants at the top of the spectrogram. In the next section, that labeled [j], there
is a drastic loss of intensity and also a noticeable lightening of the formants in the
spectrogram, both indicating a decrease in the energy involved in the constriction
necessary to form [j].

This massive decrease in intensity/energy is one of the major reasons why this part of the segment must be considered as [j] and not as [i]. As is well known, [j] and [i] are such similar segments that it is difficult to tell them apart. They are even more difficult to differentiate when diphthongs are under consideration. Obviously, there are differences between vowels, semi-vowels, and diphthongs, but these differences are merely differences of degree. That is to say that they do not have any of the easy to detect differences seen in (most) consonants, such as the large differences in manner of articulation. Often, even the place of articulation is not even differentiated when analyzing vowels, semi-vowels, and diphthongs. Their differences are more subtle and are not always immediately obvious to either listeners or researchers. One definition categorizes [j] as 'a narrowing in the oral cavity that is more constricted for a high vowel' (Stevens 1998:530). The narrow constriction produces low first formant frequencies, often in the range of 250 – 300Hz, but is not so constricted as to produce the turbulent noise indicative of a fricative. This constriction results in a loss of amplitude, which in turn leads to a loss of intensity, particularly at the higher frequencies (Stevens 1998). The following definition can also offer some guidelines on telling such semi-vowels from their vowel counterparts (and also highlight the pitfalls of doing so):

> The semivowel glide consonants, [w] and [j], when combined with
> vowels, are similar to diphthongs; the differences are that the glide
> consonants are produced with a constriction that is greater than the

closest vowels and the articulatory movements to and from the glide

constriction are faster than the movement between the two vowels of a

diphthong. (Pickett 1998:102)

Thus, the difference between a semi-vowel and a vowel is a matter of degree in which

the articulatory movements are faster than vowels but slower than other consonants.

This slowness and therefore longer period of constriction is indicated by the loss in

intensity seen here.

Pickett argues that the difference between semi-vowels and vowels is that the

semi-vowels should have a lower intensity of F1 and F2 and often the other formants,

F3 and above, will disappear altogether (1998). In particular, semi-vowels are expected

to have strong energy below 600 Hz, but to show weak mid-frequency energy (Pickett

1998:112). This loss of energy in the mid-range can be seen in Figure 3.10 and is

indicative of a semi-vowel rather than a vowel. However, the spectrogram here

continues to show strong energy and amplitude in the F3 range, perhaps indicating that

Pickett's definition is not as useful as some of the others explored here. Indeed, semi-

vowels are difficult to identify solely through the intensity contour because close

vowels are less sonorous and less intense than more open ones. The intensity contour

combined with the disappearance of mid-range energy helps lead to the conclusion that

some speakers are articulating a vowel + semi-vowel in this word.

In this case, the only piece of evidence dictating against the interpretation of the

second segment as a semi-vowel is the somewhat longish nature of its duration. In this

case, the segment labeled [j] in Figure 3.10 measures 83 milliseconds, somewhat long for a semi-vowel. Although it is true that it would be more satisfactory if it were shorter, its duration is not unexpected for semi-vowels and is not unexpected for the rate of speech of this speaker. Semi-vowels are categorized as having 'medium-fast' oral articulations and 'medium-fast' transitions (Pickett 1998:110). In fact, transitions from a glide into a vowel have been known to measure 100 milliseconds or more (Stevens 1998). Thus, transition time combined with articulation time can sometimes be relatively long for semi-vowels as opposed to overall articulation times for other consonants.

In summary, spectrograms of some speakers of Sarikol Tajik pronounce the initial segment of [peiʃin] 'evening' as a vowel followed by a semi-vowel. Those who do so, produce spectrograms that reveal a vowel-like structure that lacks mid-range energy and has a radical decrease in amplitude and intensity with a relatively (when compared to vowels) short span of duration in the second part of the segment. These speakers are not the majority, however, so there may be some instability in this vowel that will need to be monitored in the future.

The three speakers who produced spectrograms containing [ej] were not from one location. In fact, the three speakers were from each of the three different locations under study here. More study will need to be made on this specific word to see if there is any explanation as to why these speakers would be different and to see if there is any particular path of development happening. For now, they are presented merely as those

that are different than the majority and the reason behind such a difference is left for later research.

In considering all instances of [ei] under consideration here, none of the words that had been suspected of containing this segment actually matched the schema of a diphthong. Instead, just over half of the words match the schema of long vowels with level formant vectors indicative of long vowels. A few words, however, revealed themselves to be simply monophthongs. Such segments were not long and also did not contain the glide portion indicative of a diphthong. Finally, one word exhibited widespread variation. It seemed to contain a simple monophthong for most speakers, putting it in the second category, but a few speakers produced a vowel + semi-vowel sequence instead. Truly this is a large variety of results for such a small handful of words.

3.3.4 [ou]

There were six words in this study that were examined to determine if they contain the diphthong [ou]. Of those six words, it now appears that none of them match the schema of a diphthong but instead all of them contain a long vowel followed by a semi-vowel. This is difficult to tell in some of the cases since it is always difficult to tell if a segment is a true vowel or only a semi-vowel. In addition, if the preceding vowel is a monophthong that is showing drift and decreasing in intensity as it moves towards the next segment, the problem is even greater. Nevertheless, strong arguments can be made for the vowel + semi-vowel conclusion.

The extent of this identification and segmentation problem can be seen in a spectrogram of [tou] 'you' in Figure 3.11. An examination of that spectrogram reveals an intensity peak at the first part of the vowel, which begins right after the random striations caused by the aspiration associated with the voiceless stop [t]. In this case, the transient portion of the stop is short, followed by an intense period of frication, in which the amplitude of the formants increases. Aspirated [t] generally will produce waveforms in higher ranges above 3600 Hz (Kent 2002:147). This aspiration can be seen here in the range mostly between 2000 - 3000 Hz. After the aspiration, there is a peak in intensity corresponding to maximum amplitude of the vowel. After the peak, however, there seems to be a steep and constant decrease in intensity to the end of the word. This loss of energy in the mid- and upper-ranges is reflected in the loss of formant structure and also in the decrease in intensity. In addition, since there is no rapid movement indicative of a glide portion of a vowel in any of the formants, it is clear that there is not a diphthong here. But that is only part of the dilemma.
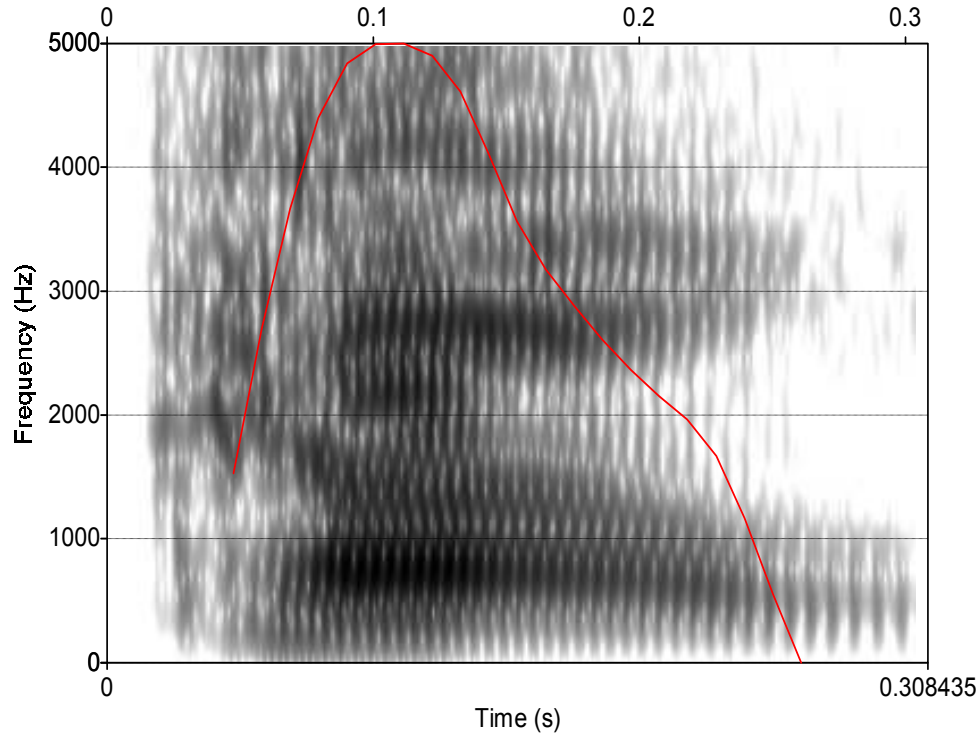
Figure 3.11 Spectrogram with intensity contour of [tou] 'you'

The spectrograms reveal that there is a vowel that has a transition to the semi-vowel [w] in it. Recall that the loss of mid and upper level energy is often indicative of semi-vowels. As the oral cavity becomes more restricted to form the semi-vowel, the energy is increasingly impeded, at least in comparison with vowels. In Figure 3.11, there is a loss of mid-range energy at approximately 137 milliseconds. However, there seems to be a second point at which intensity also drops off dramatically. This is the point at which the vowel begins to be rounded, as can be seen by the downward bending formants. This indicates a decrease in formant values, which in turn indicates a rounding of the lips. This has been formulated into the 'Lip-Rounding Rule' by Pickett

93

in which he states that all formants are lowered by lip rounding. (Pickett 1998:42).

Indeed, [w] has been said to constrict the lips in such a way as to lower both the

frequency and the intensity of F1 and F2 and to cause a great reduction in intensity of

F3. In addition, [w] can be formed in any tongue position and will take on the tongue

position of adjacent vowels (Pickett 1998:105). This makes the division between the

vowel and the adjacent semi-vowel very difficult and perhaps even meaningless given

the high ability to co-articulate the vowel and the semi-vowel. If, however, the criterion

of 'loss of F3' is adopted to distinguish the vowel from the semi-vowel, then the

resulting vowel measures nearly 200 milliseconds, indicating that it would be a long

vowel.

Since people do not speak with phones divided into segments and packaged for

ease of division, other points of division could be said to have their merits as well. For

the sake of the present investigation, the most important point here is that there is no

spectrographic evidence of anything resembling a diphthong. Nevertheless it can be said

that the interpretation that seems most consistent with the spectrographic evidence --the

intensity contours, what is known of semi-vowels in general, and what is known of

vowel lengths in Sarikoli in particular -- is that the segment in Sarikoli 'you' seems to

be [to:w].
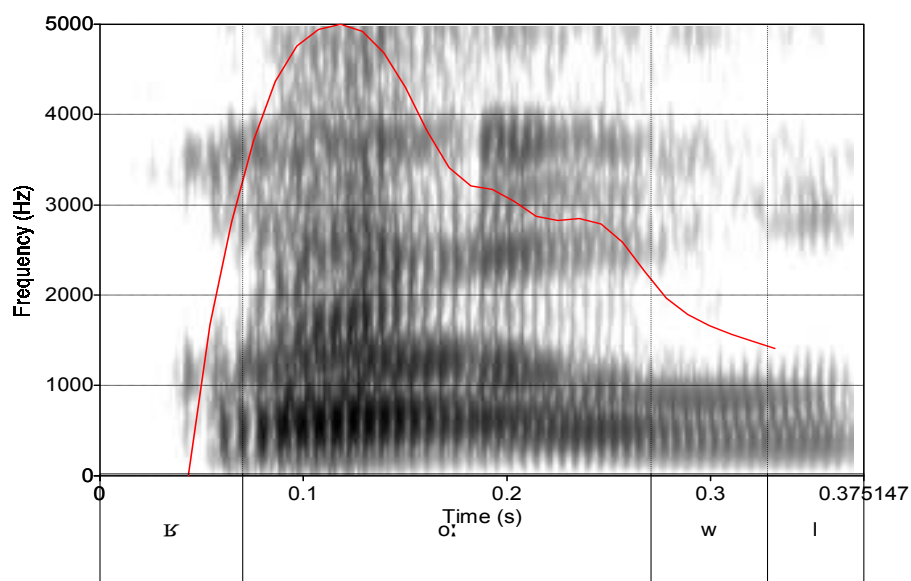
Figure 3.12 Spectrogram with intensity contour of [ʁoul] 'ear'

Leaving behind [tou] 'you' and examining other similar words yields like results. That is to say, [ʁoul] 'ear', [zou] 'grain', [ðou] 'two', [doulat] 'country', and [bou] 'odor' all appear to contain the long vowel [oː] followed by a semi-vowel [w]. This can perhaps be seen most clearly in a spectrogram of [ʁoul] 'ear' as shown in Figure 3.12. That spectrogram begins with the voiced fricative [ʁ], exhibited by striations at higher frequencies. In addition, back fricatives often cause the F2 and F3 values of following vowels to initially begin fused and then separate in frequency (Kent 2002:160). This is the case here in which F2 and F3 begin fused but then separate. The spectrogram has been divided into phones showing [oː] to have a duration of 200 milliseconds. After the vowel is the semi-vowel [w], which has been demarcated following the 'loss of F3' rule mentioned previously. Where F3 disappears is said to be

95

the beginning of [w]. As is typical of semi-vowels like [w], there is also a corresponding loss of energy in the mid- and upper-ranges. F1 and F2 continue to merge as [l] is articulated, which exhibits a further decrease in intensity but a slight strengthening of mid-level energy. In fact, this increase in mid-level energy around the area of F3 or F4 is indicative of a lateral (Pickett 1998:110).

Given this dramatic drop-off in intensity and the related disappearance of mid- and upper-level formant structures, this vowel segment seems best interpreted as a vowel + semi-vowel [oːw] and not as a diphthong. In any event, the spectrograms clearly do not show any of the characteristics of a diphthong. If this were a diphthong, the glide portion between the onset and offset steady state would be clear and prominent. Here, the change is more of a trailing off of intensity, leading to the rejection of the diphthong hypothesis for these segments.

Thus, none of the segments examined in this study reveal a diphthong [ou]. While it is sometimes difficult to know exactly where the vowel ends and the semi-vowel begins, it seems that all the words considered here contained a long vowel [oː] followed by a semi-vowel [w]. Such a conclusion is consistent with the length of the segment analyzed, the formant shape, and the lack of intensity exhibited in the [w] portion.

3.3.5 [oi]

Six words were examined to test for the diphthong [oi]. In this case, the results were unlike the previous segments analyzed in that all six words from page 64 do

appear to contain a diphthong. This is illustrated in Figure 3.13, which shows a spectrogram of [xoid] 'to read' segmented into phones. This word will suffice as an example of the very similar situation exhibited by the other words in this group.
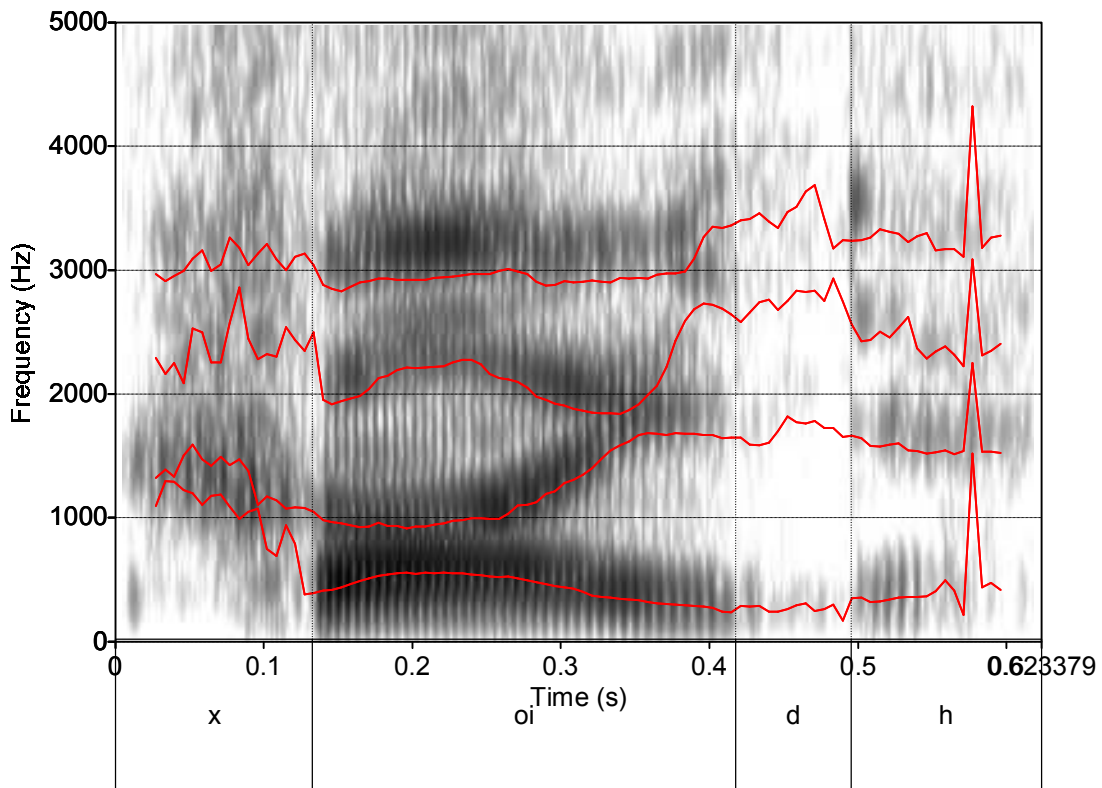
Figure 3.13 Formant tracks of [xoid] 'to read'

In the case of [xoid] 'to read', the beginning of the formant tracks shows the random striations of the fricative [x] followed by the onset vowel [o]. Velar fricatives are generally characterized by noise peaks below 2,000 Hz, with velar fricatives exhibiting higher frequencies than uvular fricatives (although this difference may not

97

appear in all languages) (Gordon 2002). Velar fricatives also seem to have little effect on vowel formant transitions in many languages (Gordon 2002), but here there seems to be a lowering of F1 and a raising of F2, similar to the pattern often found in velar stops (Pickett 1998). Here, however, the onset vowel is seen followed by the glide/transition portion of the diphthong indicating a change in constriction in the vocal tract as the offset [i] is formed. In fact, the movement is a nearly prototypical example of a diphthong schema. F1 lowers and F2 raises until finally a leveling off is reached and the offset steady state is parallel to the base. Finally, the voiced plosive [d] is indicated mostly by silence but with diffuse high frequency bands and continuing glottal pulse, to indicate voicing. Finally, the [d] is heavily aspirated on release. This can be seen in the [h], which retains the basic formant structure of the offset vowel in the diphthong, but with lower intensity (Pickett 1998).

In the formant structure of the long monophthongs examined to this point, the movement in the vowel was clearly qualitatively different from what is seen in this spectrogram. In those long vowels, the transitions in vocal tract shape were slow. In this example, however, the changes in vocal tract constriction from the onset to the offset portion of the diphthong are relatively rapid. This diphthong also shows a relatively longer initial segment and a relatively shorter off-glide segment, consistent with all cases of unambiguous diphthongs that have been examined thus far. In addition, the off-glide segment shows no appreciable loss of intensity, eliminating the consideration that

it might be a semi-vowel. Taking all this into consideration, the inevitable conclusion is that the segment in Figure 3.13 is a diphthong [oi].

Although there is no doubt as to whether words like [xoid] 'to read' contain a diphthong, [vijoid] 'ride' deserves some special consideration due to its large concentration of vowels, semi-vowels and diphthongs. If this transcription is correct, then the differences between these segments ought to be able to be seen in the spectrograms representing them. Indeed, the differences between the segments can be seen quite clearly. After the initial fricative [v], as indicated by the overall weak energy and the fairly diffuse spectra (although stronger above 5,000 Hz) and glottal pulses gaining in energy (Kent 2002:166). The word reaches an intensity peak on the vowel [i]. Since this is the stressed syllable in the word, and one of the ways stress is identified in Tajik is through intensity, this is to be expected. Next, there is a dramatic decrease in both intensity and formant contours in the region of [j]. This loss in intensity is likely due to the relatively faster constriction required for the semi-vowel than for the adjacent vowels. Indeed, the primary difference between semi-vowels and vowels is that the rate of transition for semi-vowels is faster than that in vowels (Kent 2002:178, Pickett 1998:102). Much greater intensity builds in the oral cavity in the formation of semi-vowels due to the increased and relatively rapid constriction. The rapid constriction also causes a corresponding decrease in mid-frequency energy (Pickett 1998:112). In this case, there is a noticeable weakening in F2 during the semi-vowel portion. In

Figure 3.14, it can be seen that the intensity curve does indeed, temporarily drop off and then rebounds with the onset of the vowel in the next syllable. This is followed by the articulation of the diphthong, with the glide vector of rapid movement that is indicative of it. Finally, the voiced plosive [d] has some co-articulation of the diphthong, but ultimately is characterized by only the glottal pulse indicating voicing. Typically, F2 will rise and F3 will lower the formant frequency when transitioning from [o] to [d] (Kent 2002:154), which can also be seen in Figure 3.14.
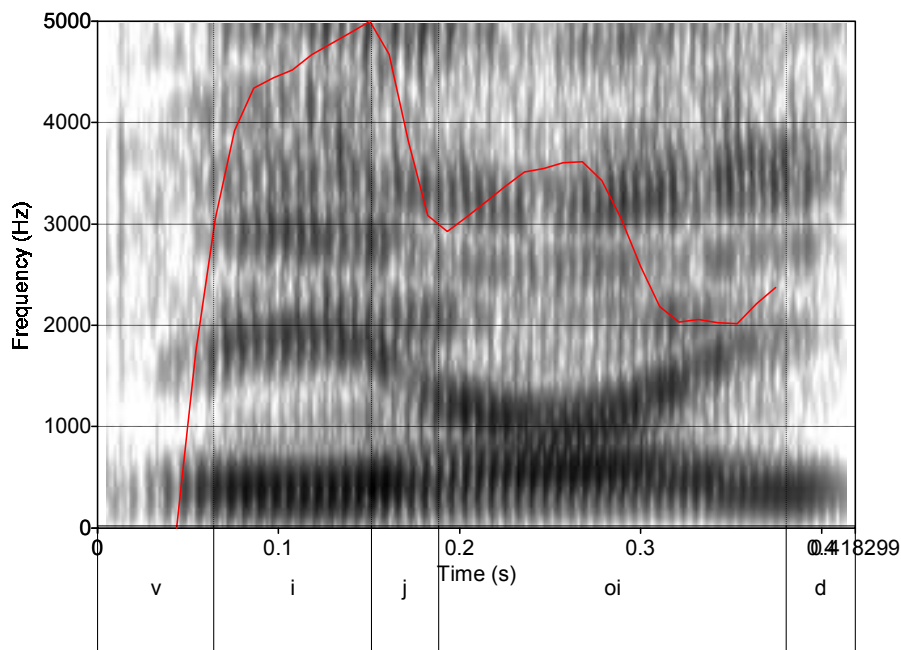
Figure 3.14 Spectrogram and intensity contour of [vijoid] 'ride'

Nonetheless, not every speaker of this word was as unambiguous as this one. In fact, there were two speakers whose articulations of [oi] did not appear to be diphthong-

like at all. Both speakers moved to a point of articulation consistent with [i], but there seemed to be no transition state to that vowel. Instead, the change was abrupt, with no glide between [o] and [i]. Therefore, these segments lacked the relatively level formants of the long vowels seen in other words, but they also lacked the glides seen in diphthongs of other speakers. If the segments are in the process of being interpreted into two monophthongs, then it is also possible that the word is in the process of being resyllabified as well. One of the speakers was from Vacha and one from Tashkorgan, so it does not appear to be a regional variation. Since both speakers were also men of about the same age and socio-economic status, no immediate differences can be detected between these two men and the others in the study. The ultimate resolution of why these two are different will have to be left for a different study.

For the vast majority of speakers and in the case of every word examined in this study for evidence of the diphthong [oi], spectrographic analysis revealed near prototypical diphthongs. Each instance of the six words under study revealed vowel segments with steep glide vectors from the onset steady state to the offset steady state. Those two speakers who were exceptions to this rule, were exceptions only in their pronunciation of the one word [vijoid] 'to ride' and were the same as the other speakers in every other way. This leads to the conclusion that these words do in fact contain the diphthong [oi].

3.3.6 [iu]

Investigation of the diphthong [iu] is troublesome because it has been said to appear in only one word: [iu] 'one'. This alone makes its existence suspicious since a diphthong is expected to have a wider distribution than one word. Nevertheless, pre-study interviews did not reveal any other words that purportedly contained this sound nor have any other researchers proposed it in any of their transcriptions. Given the lack of its occurrence in the data and the inability of native speakers to provide examples, it should not be surprising to find that spectrographic analysis did not indicate a diphthong. Instead analysis bears out previous suspicions and reveals a long vowel [iː] plus a semi-vowel [w] in all speakers.

Figure 3.15 shows a spectrogram in which [iː] gradually gains energy, indicated by gradually darkening formants, until quite suddenly the opposite happens. F2 suddenly disappears altogether and F1 continues, albeit much more weakly than before. Even with the apparent partitioning off of this second part of the vowel, the [iː] portion still measures 221 milliseconds. On the other hand, the second segment is only 87 milliseconds by contrast. The second segment contrasts in other ways as well. For example, the energy of the segment is weak and shows very weak mid- and high-level energy. This decrease in energy reflected in the weak formants, combined with its relatively rapid transition into the second portion (Pickett 1998) make this second segment clearly a semi-vowel and not a vowel. This alone would be enough to reject the hypothesis that this word is composed of a diphthong, but the lack of match with the

prototypical glide portion between onset and offset portions also indicates that this segment is not a diphthong. Thus, based on three pieces of evidence, lack of match to the prototype, lack of coherent formants due to lack of energy, and rapid transitions, this segment is best considered to be [w].
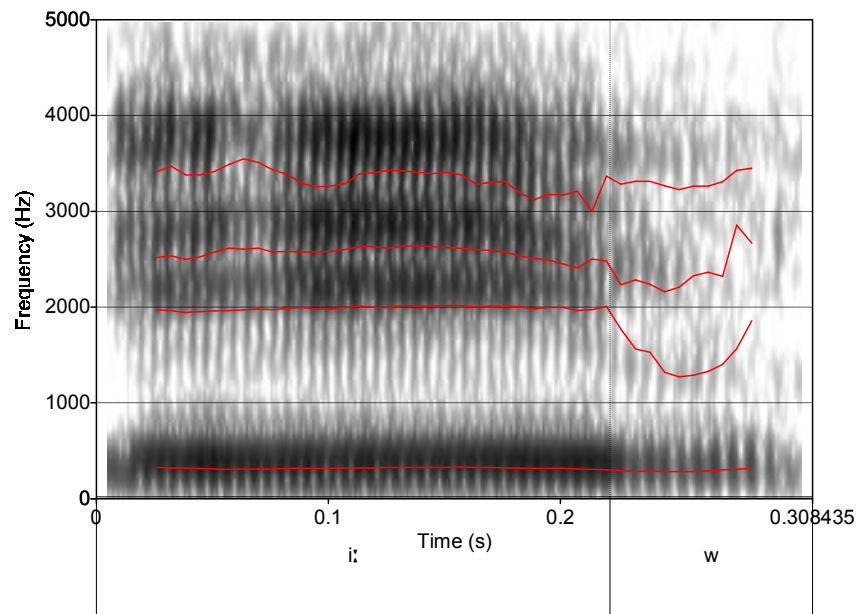
Figure 3.15 Spectrogram and formant tracks of [iu] 'one'

Since there are no other words to test, it can only be concluded that there is no diphthong [iu] in Sarikoli. Instead, words reported to contain this sound (if there are any others) seem to more closely match the schema for a long vowel [iː] followed by a semi-vowel [w].

3.3.7 [ɛi]

Like [iu], [ɛi] is a difficult diphthong to investigate due to the paucity of words in which it has been said to occur. In addition, 'white' has been variously transcribed by different linguists as [spæjd], [spejd], [speid], [spɛjd], and [spɛid]. Therefore, it should not be surprising to find some variation in speaker pronunciations. Analysis reveals that all speakers are uttering a long vowel [ɛː], but some speakers seem to have relatively steep formant vectors toward the end of that long vowel than others, leading to the impression that this segment is a diphthong. However, while the segment approaches such an interpretation, it seems that the movement does not match the schema for a diphthong. On the other hand, the nature of the movement is different than what Sarikoli long monophthongs have exhibited thus far. This raises the question: How slow is too slow to be diphthong? How much of a difference is there between a long vowel with drift and an actual diphthong?

Before approaching those more fundamental questions, it is perhaps prudent to look at the actual formant patterns in question. In Figure 3.16, it can be seen that the vowel segment does exhibit movement. It begins with F1 around 600Hz and ends with F1 around 400Hz. This indicates that the vowel begins at a mid-low height around what one would expect from [ɛ] but ends in a very close articulation, much more what one would normally expect from [i] (Pickett 1998:44). Moreover, F2 also shows a great deal of change. It initially has a value around 1400Hz but ends with a value around 2000Hz. This indicates that the vowel does begin in a rather central location but ends at a very

front point of articulation. Moreover, the vowel is quite long, measuring 221 milliseconds. The movement in the formants does match the prototype of diphthongs, but the change is more rapid than has been seen in long monophthongs up to this point. In addition, the long monophthongs seen in this study thus far have all exhibited a fairly level F1. Even when F2 showed movement F1 generally remained fairly level. This movement certainly explains why it would be easy to interpret the vowel as a diphthong simply from hearing it, but the movement of the formants in most speakers does not warrant such an interpretation -- at least not at present. If this were the only evidence, then one would likely conclude that this segment matches the schema for a long vowel with drift, but the situation is even more complex.
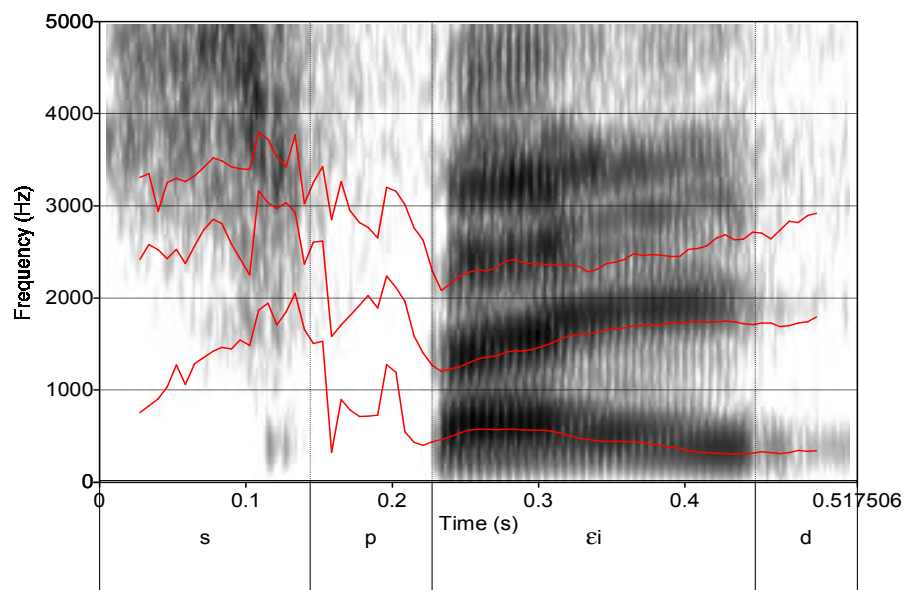
Figure 3.16 Spectrogram and formant tracks of [spɛid] 'white'

To make a complex situation even more difficult, it seems that articulation of this segment shows wide differences among individual speakers. Most important to this discussion, some speakers seem to be articulating something very closely approximating a diphthong. For example, Figure 3.17 shows the formant contours of a speaker from Burungsal. In this, his second articulation of the word, the vowel actually achieves an off-set steady state. His first articulation, however, did not show such a steady state. It merely showed long, level formants like those seen for other speakers. Even in this articulation, the F2 slope is greater than that previously seen in long monophthongs, but certainly less than the prototype of a diphthong. Given this ambiguity, it seems possible that some speakers are producing diphthongs in this position at least some of the time. Such speakers appear, however, to be in the minority for now.

This situation shows that the difference between a long monophthong with 'considerable' drift and a true diphthong are only matters of degree. It is possible that the degree is so small that different speakers interpret the vowel in different ways and actually articulate it in slightly different ways. It is highly likely that this segment is in a state of transition and does exhibit a fair amount of variation from speaker to speaker, which is why different linguists have interpreted the segment in different ways. If the current pattern continues, this segment should continue to change from a diphthong to long monophthong.[3] It seems, however, that the transition is not quite complete.

---

[3] Based on the change already seen in other related languages, it is clear that the change is going from the diphthong to the simple vowel and not the other way around. See Chapter 5 for more information.
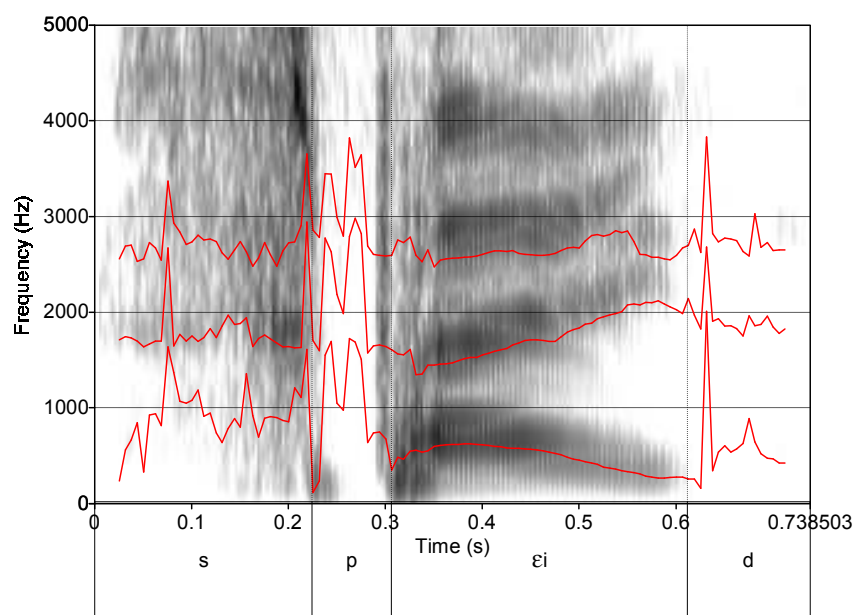
Figure 3.17 Spectrogram and formant contours of [spɛid] 'white' sho
wing a nearly prototypical diphthong

Given the lack of definitive evidence provided by the previous analysis of [spɛid] 'white', one might wonder if the situation might become clearer if there were more words to investigate. Because of the pre-field interviews which could not elicit any more words containing this diphthong, and because the diphthong does not occur in most linguists' analysis of the language, and because of the inconclusive results of the spectrographic evidence, it seems premature to call this segment a diphthong at this point. Thus, for now, this segment will be considered as a long monophthong [ɛː].

On the other hand, perhaps such an analysis begs the question. Perhaps when posing the question of whether this segment is a diphthong or a long monophthong, the best answer is just 'yes.' It is different segments at different times -- sometimes even

different to the same speaker at different times. This acknowledges reality, but it somehow also fails to capture a generalization. Thus, perhaps the most accurate statement is that generally speaking, this segment best matches the schema for a long monophthong [ɛ:] but some speakers during some articulations actually utter [ɛi]. At least based on spectrographic evidence, this is the most that can be said. More will be said in Chapter 5 about why this situation might exist.

3.3.8 [ɛu]

Unlike the previous two diphthongs, the last diphthong under consideration here [ɛu] has been reported to occur in more than one word. In this study, three words [njɛu] 'nine', [vrɛu] 'eyebrow', and [tʃabɛud] 'pigeon' were taken into consideration. Analysis indicates that all three of these words do not contain a diphthong but actually contain the vowel [ɛ] and the semi-vowel [w]. Although, like some of the speakers in the previous section, some speakers had more diphthong-like pronunciations than others, potentially indicating that this segment is also in a state of change. Certainly a change from a diphthong to a vowel + semi-vowel would be small indeed, merely a change in intensity and speed of articulation. A close vowel such as [u] is also less intense than its open vowel counterpart which could account for the loss of energy in the spectrograms, but the divergence of energy/intensity between the first part of the vowel segment and the end is greater than what can be accounted for merely by the difference in two vowels -- lending weight to the conclusion that this is a semi-vowel. Secondarily, the rate of the transition is rapid. Recall that the rate of transition in a

diphthong from the onset steady state to the offset steady state is slower in diphthongs than in transitions to from vowels to semi-vowels. Thus, diphthongs exhibit a glide/transition portion while vowels to semi-vowel transitions are relatively rapid and therefore do not have such long transitionary portions (Pickett 1998). All of these points can be seen in analyses of specific words below.

A typical spectrogram for words analyzed for the diphthong [ɛu] can be seen in Figure 3.18. In this analysis of [tʃabɛud] 'pigeon', the vowel segment under analysis begins after the essentially silent [b]. Voiced biliabial stops are characterized by low frequency glottal vibrations. They can also be seen in the lowering of the formants in the transition from the stop to the vowel. [b] seems to lower all formants in the transition phase of following vowels. (Kent 2002:155). From the intensity of the formants, it can be seen that the second part of the vowel segment (the part labeled [w] here) is much less intense than the first. This drop in intensity is caused by the relatively constricted position of the tongue in the oral tract required to form [w]. The constriction blocks energy, which is seen in the lightening of the formants in the spectrogram. The rounding of the lips required in the articulation of [w] also results in the lowering of F2 (Pickett 1998:42). In addition, it is clear that there is no transition segment between [ɛ] and [w] that would be required of a diphthong. Finally, [ɛ] is not as long as has been seen in previous long vowels. Here [ɛ] is only 91 milliseconds and lacks any sign of drift. All this points to the conclusion that this phone is a simple monophthong [ɛ] followed by a semi-vowel [w]. Still, while all instances of [tʃabɛud] 'pigeon' were

109

fairly clear in supporting this conclusion, other words were not so clear or straight-forward.
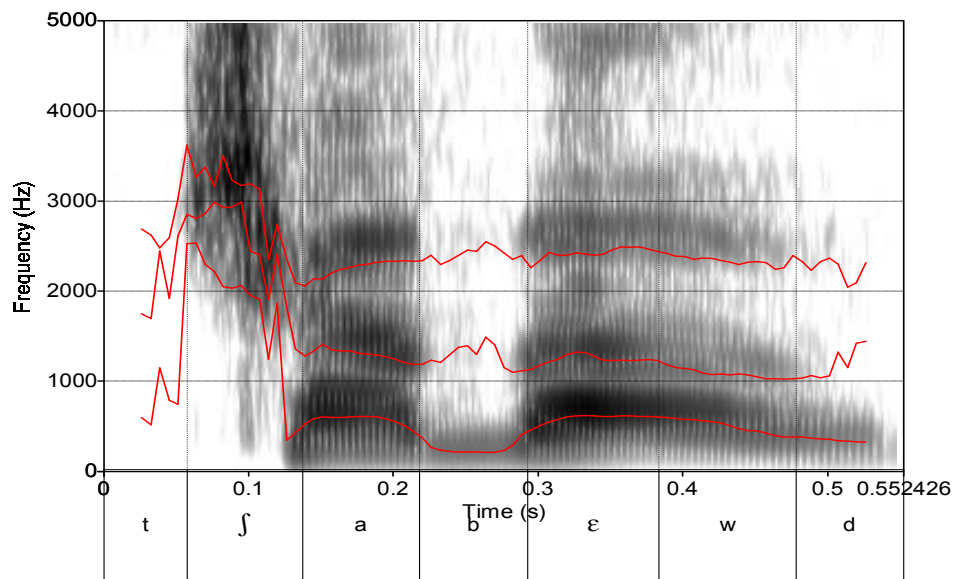
Figure 3.18 Spectrogram and formant tracks of [tʃabɛud] 'pigeon'

Complicating what at first appears to be a straightforward situation is the variation in several speakers' pronunciations of [neu] 'nine', which showed considerable variation. For instance, some speakers exhibit a noticeable decrease in intensity during the articulation of the last part of the vowel segment. This was indicated in both the intensity contour and also in the random striations that formed in the formants, or in some cases, the disappearance of the upper formants altogether (cf. Figure 3.19). On the other hand, some speakers seem to have a nearly classic diphthong

pattern, complete with the all-important glide portion and the on-set and off-set steady state.
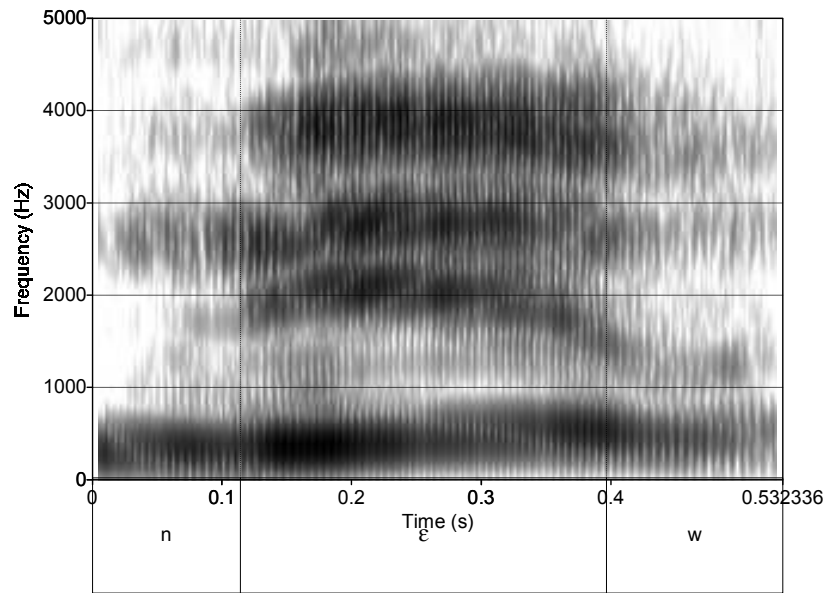
Figure 3.19 Spectrogram of [neu] 'nine'

A good example of both problems can be seen in Figure 3.19. After the initial [n], indicated by the strong murmur intensity and some mid-frequency energy (Pickett 1998:140). Typically alveolar consonants produce transition formants in which F1 is lower and F2 is higher in the transition into following vowels (Kent 2002:136). In this case, however, the nasal is palatalized, which obscures this effect. The articulation of the [ɛ] is clear and then there is rapid movement to the final segment, but this movement does not seem to match the schema for diphthongs. In addition the relatively close constriction of the segment is indicated by the noticeable weakening in energy in

111

the formants of the last segment. Over the last segment, there is a leveling off of the formants, which might be indicative of an off-set steady state. However, it is also clear from the lack of intensity in the formants that are present and the loss of energy that this segment does not carry the same intensity as [ɛ].

Although it is certainly not clear, the inclination here is to identify these phones as diphthongs and attribute the weakness of the last segment to word attrition experienced by a word spoken in isolation. Arguing against this is the weakness of the formants and the fact that an unambiguous case of a monophthong followed by semi-vowel already exists in the word [tʃabɛud] 'pigeon'. The main barrier to accepting a similar interpretation (of monophthong + semi-vowel) in [nɛu] 'nine' and [vrɛu] 'eyebrow' is the abrupt change in F2 -- a change that had not been present in previous segments that were seen to have a semi-vowel in them. Given that both sides of this argument have their strengths and weaknesses and given that this is an attempt (in this chapter) to look at language synchronically, this will be deemed a less prototypical diphthong. Reasons for why this segment might be weak and is likely in a state of transition are explored in Chapter 5.

3.4 CONCLUSIONS

Based on spectrographic analysis of Sarikoli words purported by various linguists at various times and places to contain diphthongs, it appears that Sarikoli Tajik does contain some diphthongs, but not as many as previously supposed by some. This result is not surprising given the unusually large number of diphthongs that had been

proposed for the language. What is surprising, however, is how many long vowels were revealed. Although every other Pamiri mountain language has been reported to contain both long and short vowels, Sarikoli has never been reported to contain such a difference. No linguist who has ever studied Sarikoli has proposed a short and long vowel distinction before. On the other hand, it is well attested in the literature that all other Pamiri mountain languages have both long and short vowels in them (Morgenstierne 1938, 1974, Pakhalina 1960, 1969, Sokolova 1953, 1967). All literature printed about Sarikoli in either English (with the exception of Schwarz (1984) who based his work on a Chinese source) or Russian has been based on the work of T.N. Pakhalina, so that all other sources were only passing on what they had inherited. It now seems extraordinary that no one really seemed suspicious about this unusual state of affairs. Why should Sarikoli be so unusual among the languages of this area? No one seemed to raise the question. Equally unusual has been the Chinese approach to Sarikoli, which posited an amazing nine diphthongs in the language (Gao 1985). Despite the fact that no language in the world has (yet) been attested to contain so many diphthongs, this was also accepted with little or no hesitation. The most recent study conducted by Peters and Peters (1996) did not propose such a large numbers of diphthongs. On the other hand, not being linguists, they were unaware that all other Pamiri languages contained both short and long monophthongal vowels, leaving them out of the controversy altogether.

Apparently, long vowels in Sarikoli display so much drift that long vowels were all believed to be diphthongs by previous researchers. Recognizing the difference between Sarikoli diphthongs and long monophthongs reveals an answer to the long standing mystery of why only one language in the whole family did not have a distinction between long and short vowels. One of the reasons this mystery has persisted is that it does have the same long and short monophthong distinction as all of its neighbor languages but Sarikoli's long monophthongs exhibit such large drift that they were previously improperly analyzed. Only with the modern tools of spectrographic analysis can such a mystery finally be solved.

Another reason why the mystery of the lack of long vowels in Sarikoli has persisted for so long is revealed in the great variety of results among speakers in this analysis. There is a large amount of variation among Sarikoli speakers, which has made it difficult to decipher whether speakers were uttering a long monophthong or a diphthong. Results also show that some speakers' idiolects are in such a state of flux that even within one idiolect speakers will sometimes use a long monophthong while at other times using a diphthong. This lack of consistency is not at all surprising, but has only been recently accepted in linguistics as natural. Formerly linguists always searched for consistency across the language as a whole or even accepted the data from one speaker as indicative of the state of the entire language. Only recently has the field of linguists changed sufficiently to recognize that change and transitionary states are normal and are worthy of study. Previously, this lack of acceptance of inconsistency

and transitional forms led linguists to ignore transitionary forms, insisting that a form be either one thing or the other - in this case either a diphthong or a long monophthong. Now, linguistic theory has advanced to the point where linguists are able to recognize less and more prototypical varieties of diphthongs and to recognize that diphthongs and monophthongs are merely ends on a continuum. While a tool as simple as prototype theory might not seem revolutionary anymore, it has opened up paths of analysis previously unavailable to linguists. While the path might no be much wider than before, it is a path that is wide enough to solve a question whose solution had previously been a mystery to linguists.

Table 3.4 Sarikoli Vowels in light of spectrographic evidence

Simple Monophthongs			Long Monophthongs		Diphthongs
i	ɯ	u	iː		ui
e		o	eː	oː	oi
ɛ		a	ɛː	aː	ai

To the question of why Sarikoli lacks long vowels when other Pamiri mountain languages all have them, it can now be said: It is like other Pamiri mountain languages. It does have long vowels. The only difference is that Sarikoli long vowels exhibit considerable drift and therefore might sound like a diphthong initially. Given this, then what is the current state of Sarikoli vowels? As seen in Table 3.4, Sarikoli currently

exhibits a relatively balanced inventory of short and long vowels. The basic vowel inventory is composed of a balanced seven vowel system. In addition, that system is mirrored in a system of five long vowels. Only the vowels [ɯ] and [u] do not have a long vowel equivalent. Finally, there are three diphthongs, [ui], [oi], and [ai]. Given the unstable nature of some speakers pronunciations and the current gap in the long vowel inventory, it can predicted that [ui] will likely develop into a long monophthong given more time and then will fill in the 'gap' in the long vowel chart (see discussion in 3.3.2.), providing a nearly perfectly balanced system. The diphthong [ui] might develop into either the long monophthong [uː] or perhaps even [ɯː]. Time will tell, and more on this path of development will be explored in Chapter 5. For now, this segment will be regarded as a diphthong, but a less prototypical one.

The Sarikoli vowel system as presented in Table 3.4 is relatively unmarked. The simple vowels are balanced and contain neither an exorbitantly large nor small number of vowels. The only somewhat marked segment [ɯ] is unusual but is completely logical as the only central vowel in the language. The most unmarked case would likely be [ə], but [ɯ] is not so unusual either. The fact that Sarikoli has long monophthongs at all moves it from being a language marked within its language family to being unmarked. Since all other languages in the family are widely known to contain long vowels, the presence of long vowels is neither marked nor remarkable. The diphthongs that remain are also not marked. According to Miret's (1998) synthesis of previous research, [ai] is the most common diphthong in the world, [ui] is the second most common, followed by

[oi] which is the sixth most common. Thus, some of the more marked diphthongs that were initially included in this analysis were proven at last to not be diphthongs at all. In the end, the spectrographic evidence supported the conclusion that Sarikol Tajik is not that unusual after all. In fact, it is much like its neighbor languages and much like other languages of the world.

Now that Sarikoli has been observed to be quite unmarked and in light of the spectrographic evidence, some of the test words used in this study can be given a new understanding. Some of those that were understood to contain diphthongs can now be seen as long vowels while others truly are diphthongs. Still others were deemed to be vowels followed by semi-vowels. To provide clarity, all the words tested in this study are re-written in light of the conclusions drawn from the spectrographic evidence on page 118. It ought to be re-iterated that these are still not phonemic representations but simply more accurate phonetic representations in light of the new evidence in this dissertation.

In conclusion, Sarikoli appears to now be a relatively unmarked language in terms of both world-wide linguistic norms and also in terms of other Pamiri mountain languages. New tools of phonetic analysis combined with applications of prototype theory and variation theory have led to a deeper understanding of the language and removed some of the mystery surrounding this language. While previous researchers were not able to explain why Sarikoli lacked long vowels or why it appeared to have so many diphthongs, these tools have provided solutions that were not previously available

due to the lack of computer technology or the lack of development of contemporary linguistic theory.

Words in study re-written in light of spectrographic evidence

(9) a. [aː] [saːr] 'to be full'
b. [aː] [tsaːz] 'what'
c. [ai] [nai] 'no'

(10) a. [ui] [dʒui] 'place'
b. [ui] [bui] 'cave'
c. [ui] [adʒuib] 'strange'

(11) a. [ej] [pejʃin] 'evening'
b. [eː] [tʃareːn] 'man'
c. [e] [χel] 'kind/type'
d. [eː] [tʃeːg] 'to do'
e. [eː] [beːl] 'shovel'
f. [eː] [deːg] 'pot'
g. [e] [dekun] 'farmer'
h. [e] [waχen] 'blood'
i. [eː] [baweːd] 'to disappear'
j. [eː] [indeːd] 'to stand up'

(12) a. [oːw] [toːw] 'you'
b. [oːw] [ʁoːwl] 'ear'
c. [oːw] [zoːw] 'grain'
d. [oːw] [ðoːw] 'two'
e. [oːw] [doːwlat] 'country'
f. [oːw] [boːw] 'smell/odor'

(13) a. [oi] [poi] 'to herd'
b. [oi] [poi] 'yoghurt'
c. [oi] [boi] 'rich'
d. [oi] [noi] 'flute'
e. [oi] [xoid] 'to read'
f. [oi] [vijoid] 'to ride'

(14) [iːw] [iːw] 'one'

(15) [ɛː] [spɛːd] 'white'

(16) a. [ɛw] [njɛw] 'nine'
b. [ɛw] [vrɛw] 'eyebrow'
c. [ɛw] [tʃabɛwd] 'pigeon'

This does not mean that all controversy surrounding Sarikoli diphthongs has been resolved, however. There still remain (at least) two critical issues to be solved. The

first is whether or not Sarikol Tajik dialects vary in their pronunciation of some of these vowels. Previous researchers have reported that there are three dialects of Tajik that vary primarily in their pronunciation of vowels. Although the vowel system has been examined in terms of the language as a whole, are there regional dialectal differences? Can the spectrographic evidence either confirm or deny these differences? These issues are taken up in Chapter 4.

Finally, there is the issue of how some of the current segments in Sarikoli Tajik came to be that way and how they are developing. Why are some diphthongs prototypical while others are marginal? Why do some idiolects show such vast variation? Why are results across speakers so variable? In other words, there are issues of development in the language that can perhaps be explained from a historical perspective. Some of these issues related to vowel system development are explored in Chapter 5.

CHAPTER 4

ACOUSTIC AND STATISTICAL ANALYSIS OF DIALECT VARIATION

4.1 PREVIOUS PROPOSALS

Tajiks, like other people of the world, are aware of dialect variations in their language. They are quick to point out that people from one area or another speak 'funny' or 'oddly'. Linguists who have previously studied Tajik of China have likewise detected dialectal differences between speakers from various locations. One must be careful when talking about 'dialects' with Chinese linguists because their definition of a dialect is quite different from the world standard. Chinese linguists usually define a dialect as two languages that have some sort of historical relationship and for some sort of political reason are considered to be the same language or simply the language of one place as opposed to the language of another place. This is the case with the work of the Chinese scholar Gao (1985), who identifies what he terms 'dialects' in his work, but actually are not dialects by world standards at all but two mutually unintelligible languages. In fact, the 'dialect' that he refers to is Wakhan Tajik, which is indeed spoken in Xinjiang but is not mutually intelligible with Sarikoli. Therefore, his work on 'dialects' cannot be included here. Peters and Peters (1996) make no reference to dialect variations and so their work will also not be included here.

The most reliable source on dialect variation, T.N. Pakhalina (1971), identifies three regional variations in Tajik. One of the ways in which these dialects are said to vary is in terms of how diphthongs are realized. She claims that diphthongs differ in each of three areas that she terms Central, Near Eastern, and Far Eastern. 'Central' corresponds to the variety of Sarikoli spoken in Tashkorgan, 'Near Eastern' refers to Sarikoli as spoken in Vacha and surrounding regions and 'Far Eastern' refers to the Sarikoli of the Burungsal area (cf. Figure 3.1). Her proposal (displayed in Table 4.1) identifies three different regional varieties. Although she proposed segments that contained a vowel and semi-vowel, her transcription has been adapted to fit the definition of diphthongs adopted thus far in this dissertation.

Table 4.1 Pakhalina's proposed dialect correspondences

Central (Tashkorgan) 塔什库尔干		Near Eastern (Vacha) 瓦恰乡		Far East (Burungsal) 布龙夏村
[ɛi]	~	[æi]	~	[ai]
[ɛu]	~	[æu]	~	[au]

Pakhalina identified all three locations as having diphthongs, but those diphthongs differed primarily in the dimension of the height of the onset vowel. While Tashkorgan diphthongs contain the mid-close vowel [ɛ], Vacha diphthongs contain the mid-open vowel [æ]. Finally, Burungsal speakers are said to use the most open vowel of all with a low central [a].

Given the anecdotal evidence that corroborates Pakhalina's proposal, it is worth investigating to see if some of these differences are real or not. Do these places differ only in their onset vowel or do they also exemplify some other differences in how the diphthongs are manifested? Such questions are critical if one considers the development of a language over time. In particular, it is sometimes possible that dialect variations show the course of potential future development that a language might undergo (Bailey 1996, Chambers 2001). So that if one particular dialect has a change while another does not, it is possible that one lect is less like its historical predecessor than another. If this is the case, it is possible that the conservative variety that has not yet changed might change in that direction in the future. Before such assessments about any potential change or whether one dialect is more 'conservative' while another is more 'innovative' can be made, the basic facts of the case need to be set forth first. Therefore, before considering such notions, the basic phonetic facts must be established.

To establish those basic phonetic facts, data was recorded from five speakers in each of Pakhalina's three proposed dialect areas for a total of 15 speakers. These are the same recorded data that were used in Chapter 3. However, the main purpose of Chapter 3 was simply to assess whether speaker articulation matched the schema for diphthongs, but little or no actual measurement of formants was employed to determine if speakers differed in vowel height. Indeed, to my knowledge, Pakhalina had no equipment upon which to make such measurements either; her assessments were based totally on her own ear. Although she was highly skilled, she simply did not have access to the kind of

equipment now available -- equipment that can either confirm or falsify her proposal concerning lects of Sarikol Tajik.

4.2 METHOD OF EVALUATION

In order to evaluate Pakhalina's proposal concerning dialectal differences in Sarikoli Tajik, spectrograms of words containing her proposed diphthongs were examined for vowel height differences.

Immediately the problem arises that two of Pakhalina's proposed diphthongs turned out to not be diphthongs under spectrographic analysis. In the case of [ɛi] and [ɛu], they were actually long vowels. However, the diphthong [ai] did match the diphthong schema. This is not overly problematic for the kind of analysis being conducted here however. The essential question here concerns the height of those vowels -- no matter whether they are an onset vowel in a diphthong or now understood to be a long vowel. There was likewise, a set of words originally proposed to contain the diphthong [ei], that were eventually determined to contain long vowels showing drift, but that does not bar them from analysis either. Since the diphthongs [ei] and [ɛi] are so similar, it seemed prudent to also include those words in this analysis. In reality, different researchers have used both symbols in different places and times, leading to the conclusion that more concrete and systematic study of those words is also warranted.

When Pakhalina's proposals are combined with the transcriptions offered by other linguists like Gao (1985) and Peters and Peters (1996), the words that are likely to

exhibit some kind of dialectal variation number 13. They are summarized on page 125. The columns labeled 'Central', 'Near Eastern' and 'Far Eastern' are based on Pakhalina's transcriptions in her Sarikol - Russian Dictionary (1966). In this case, the offset portion of the diphthong was transcribed as a vowel, but it should be noted that she normally transcribed a vowel followed by a semi-vowel (as did all Soviet linguists). No other changes have been made in her transcription. These transcriptions were not tested for vowel height in Chapter 3; they were only tested to see if they matched the schema of a diphthong or were instead a long monophthong. Therefore, the diphthongal status can be regarded as set, but the height of the vowel cannot be regarded as definitive at this time. Finally, in the last column of on page 125, the English translation is provided.

Occasionally, a particular speaker did not provide the word that was expected during the elicitation procedure. They either misunderstood the word or provided a synonym or perhaps used a Uighur word instead of a Tajik one. Since the assistant doing the elicitation did not know what would be done with the data after he collected it, he did not notice any difficulty with different speakers providing different words. For this reason, [tʃarein] 'man' was excluded from this evaluation altogether, although it has been reported to exhibit variation across dialects. There were simply so many different responses that there was not enough consistent data to compare across speakers. This variation is also why the *N* values for each word evaluated below are not always the same. This simply means that one or more speakers provided a different

word, or in the case of a verb might have provided the same word but from a different part of the verb paradigm. This changed the vowel and made that particular repetition of the word not able to be used in this study.

Sarikoli words likely to exhibit dialectal variation

	Segment	Central (Tashkorgan)	Near Eastern (Vacha)	Far Eastern (Burungsal)	
(1) a.	[ai]	[sɛir]	[sæir]	[sair]	'to be full'
b.		[tsɛiz]	[tsæiz]	[tsaiz]	'what'
c.		[nɛi]	[næi]	[nai]	'no'
(2)	[ɛi]	[spɛid]	[spæid]	[spaid]	'white'
(3) a.	[ɛu]	[nɛu]	[næu]	[nau]	'nine'
b.		[vrɛu]	[vræu]	[vrau]	'eyebrow'
c.		[tʃabɛud]	[tʃabæud]	[tʃabaud]	'pigeon'
(4) a.	[ei]	[dɛig]	[dæig]	[daig]	'pot'
b.		[tʃɛig]	[tʃæig]	[tʃaig]	'to do'
c.		[bɛil]	[bæil]	[bail]	'shovel'
d.		[bɛid]	[bæid]	[baid]	'to.disappear'

To evaluate whether the vowels in these words were truly different, each word was examined using Pratt. Spectrograms were generated and F1 and F2 values were measured. In Praat, when a segment of speech is chosen, Praat analyzes average formant values of the selected portion and returns a linearly interpreted average of

formant values in the chosen section. In cases where Chapter 3 revealed a diphthong, the onset portion was composed of the first glottal pulse of the vowel to the beginning of the glide portion. In cases in which Chapter 3 revealed long monophthongs with drift, the entire vowel was composed of first glottal pulse of the vowel to where the drift began. Praat then computed an average value for F1 and F2 for the entire vowel. Once these values were computed, results for F1 and F2 values between regional varieties were compared using an ANOVA to see if the results were deemed to be statistically significant ($p \leq .05$). That is to say, statistically speaking, were the regional results different enough to be considered different values?

If these dialects do actually exist, then Tashkorgan and Burungsal will likely be at two endpoints along a continuum with Vacha in the middle. This is because of their relative geographic positions. Furthermore, if speakers in these areas do differ in their pronunciations, then F1 values in Tashkorgan ought to be of the lowest average values while those in Burungsal would be of the highest average values. That is to say, since lower F1 values indicate that a vowel is articulated in a relatively closed manner, Tashkorgan ought to have the lowest F1 values. Conversely, Tashkorgan ought to have the highest F2 values, since F2 values indicate the relative frontness or backness of the vowel under study. This is summarized in Figure 4.1. Tashkorgan vowels are labeled as having relatively low F1 and relatively high F2 values. This is because Tashkorgan is expected to contain vowels that are relatively higher and more front in the vowel space than the other lects. Therefore, a low F1 (relative to the other lects) will indicate a

vowel articulated at a place higher in the oral tract. On the other hand, a high F2 (relative to the other lects) will be an indicator that Tashkorgani vowels are farther front in the oral tract than the other lects.

F1 Relatively Low		F1 Relatively High
F2 Relatively High		F2 Relatively Low
Tashkorgan	Vacha	Burungsal
More closed and		More open and
front articulation		back articulation

Figure 4.1 Expected Differences in Diphthongs in Dialects of Sarikol Tajik

It is possible that relative differences in F1 and F2 values of various locations are not statistically different. In this dissertation, such an outcome will not necessarily necessitate drawing the conclusion that there is no difference in the dialects spoken in various locations. The results of the ANOVA will be balanced with the mean scores for F1 and F2 values and will also be considered in light of other words reported to contain the vowel under study. That is to say that the ANOVA results are certainly given credence, are reported, and are taken into consideration, but they alone are not considered to be the final authority on whether the differences in dialects is real or not. The ANOVA evidence will instead be taken as one part of the overall solution to the puzzle with the descriptive statistics such as mean F1 and F2 values and the shape of spectrograms also contributing to the solution.

Indeed, the tension between statistical significance and linguistic significance is not new. In their now classic work on statistics in linguistics, Woods, Fletcher, and Hughes (1986) identify this problem by saying, 'We must keep in mind always the difference between statistical and scientific significance, and we should remember that the latter will frequently have to be assessed further…' (129): the statistics are helpful in pointing out whether differences are real, but statistics must be evaluated by humans in light of overall linguistic knowledge and the field under investigation. Therefore, just because the ANOVA might reveal that there is not as yet any statistically significant difference between two segments; it does not necessarily mean that there is no difference worth noting or examining in the data. In fact, those working in the field of computational linguistics have discovered that systems based on only statistical methods have not been powerful enough to solve problems of natural language processing, leading many to determine that statistics alone are not enough to determine the importance of individual lexical items (Hatzivassiloglou 1996). Others writing concordance programs have also determined that statistical significance metrics can provide valuable insights into what might be linguistically significance, but that linguistic significance is based ultimately on what is 'useful' or 'natural' to users of the linguistic system (Atwell 2006). Even introductory textbooks state, 'A statistically significant effect is not necessarily practically significant' (Lane 2006). Others have said, 'Statistically significant changes, however, can be observed with trivial outcomes' (ACP July/August 2001). Thus, statistical significance will certainly be given heed here,

but statistical significance and linguistic significance are not always identical. Linguistic significance will also be weighed based on all the data available at this time. It is sometimes possible that changes in linguistics begin small and are not presently statistically significant, but may become so in the future or may be significant in light of the system as a whole. Thus, both concepts of statistical and linguistic significance will be employed here.

4.3 RESULTS

This section examines each of the words from page 125 in turn. Each word is evaluated spectrographically for F1 and F2 values in accordance with the procedure outlined in Section 4.2. After obtaining measured values for F1 and F2, each word is evaluated to determine if reported dialect variation in diphthongs can be confirmed or denied.

4.3.1 SEGMENT [ai]

Three words that were believed to contain the diphthong [ai] were examined for dialect variation. They were [sair] 'full', [tsaiz] 'what', and [nai] 'no'. In this section, each word will be examined in turn and then finally a conclusion will be drawn about the range of dialect variation among Sarikoli speakers.

Table 4.2 Formants of model vowels (Pickett 1998)

Model Vowel Front	F1	F2	Model Vowel Back	F1	F2
i	250	2150	u	250	800
e	400	2000	o	400	900
ɛ	550	1850	ɔ	550	1000
æ	700	1700	a	700	1100

Beginning with [sair] 'full', measurements did reveal an average difference in F1 and F2 values between each of the three locations. In fact, the differences are exactly as predicted in Figure 4.1. Average values were computed by measuring sounds in the way described in section 4.2, entered into SPSS, and then an average was computed. That is to say that Tashkorgan exhibited relatively lower F1 values and relatively higher F2 values than the other two locations. In addition, F1 values rose gradually among the three locations while F2 values decreased until the other extreme was reached in Burungsal. All results are reported in Table 4.3.indicating that the Tashkorgani vowel is a higher more front vowel than in the other two locations. For example, Tashkorgan, Vacha, and Burungsal had average F1 values of 468Hz, 498Hz, and 544Hz respectively, a gradual increase. On the other hand, Tashkorgan, Vacha, and Burungsal average F2 values computed to 1932Hz, 1808Hz, and 1845Hz respectively. Nevertheless, the differences between the three locations were not statistically significant based on a three

way ANOVA[4]. But, when the differences in F1 values between Tashkorgan and Burungsal were calculated, the value was close to being statistically significant with $p = .055$, indicating a trend.

If the average values of the formants in all three locations are compared with model vowel formants (cf. Table 4.2), the results of the ANOVA seem to be verified. Model formant values should only be taken as a very general guideline since absolute formant values can vary by gender, body size, and perhaps are even language specific. Therefore, '[t]he values in the formant-frequency tables should not be taken prescriptively, but rather as averages around which considerable variation can occur' (Kent 2002:113). The values offered in Table 4.2 are those from Peterson and Barney's (1952) study of 76 speakers of American English. This study is now considered classic and is often quoted in introductory texts. More recent studies have agreed 'reasonably well' with their results, and where differences between Peterson and Barney have existed, the differences may be due to dialectal variation (Kent 2002:113). In the case of the actual mean formant values listed in Table 4.2, the values have been adjusted to allow easier memorization. 'The 150Hz steps are used for F1 and for the F2 range of the front vowels from [i] to [æ]. The 100Hz steps are used for the F2 steps between adjacent back vowels and 150Hz steps are used for F1' (Pickett 1998:44). Since the model formants are used only as the most general guideline anyway, the adjustment

---

[4] An ANOVA is a statistical test used to compare the mean values of two or more groups. The advantage of using an ANOVA rather than multiple t-tests is that it reduces the chance of encountering a Type I error (the probability of finding something by chance) GEORGETOWN_UNIVERSITY. 2005. Research methods and statistics resources. October 1, 2006.
<http://www.georgetown.edu/departments/psychology/researchmethods/index.htm>..

made from the actual mean formant values does not affect the usefulness of the chart and increases its ability to be memorized. It is important to remember throughout the study here that the model formant values are offered only as a very general guideline and are not the final arbiter of decisions made regarding particular vowels and diphthongs.

Table 4.3 Statistical results for [sair] 'full'

(a) Descriptive Statistics for F1

| | N | Mean | Std. Deviation | 95% Confidence Interval for Mean | | Minimum | Maximum |
				Lower Bound	Upper Bound		
Tashkorgan	10	468	34	444	493	430	539
Vacha	7	498	75	428	568	368	574
Burungsal	10	544	82	485	603	423	652

(b) Descriptive Statistics for F2

| | N | Mean | Std. Deviation | 95% Confidence Interval for Mean | | Minimum | Maximum |
				Lower Bound	Upper Bound		
Tashkorgan	10	1932	105	1857	2008	1694	2068
Vacha	7	1808	304	1527	2089	1479	2153
Burungsal	10	1845	138	1746	1944	1707	2189

Table 4.3 - *Continued*

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	.055	.580
	Vacha	.666	.409
Vacha	Burungsal	.381	.922
	Tashkorgan	.666	.409
Burungsal	Tashkorgan	.055	.580
	Vacha	.381	.922

Turning attention to the next word to be analyzed [tsaiz] 'what', a similar situation is revealed (Table 4.4). Once again Tashkorgan shows an average F1 relatively lower than the other two locations. As can be seen by the minimum and maximum value statistics, however, there is dramatic overlap in average values between the three locations. Nevertheless, the maximum F1 values do increase steadily from Tashkorgan to Vacha to Burungsal, adding some further evidence for the overall trend of an increasing F1 value across the three locations. This difference in F1 indicates that Tashkorgani speakers (average 535Hz) are using vowels that are articulated more front in the oral cavity than in the other two locations. Likewise, Vacha (average 543Hz) speakers' vowels appear to be more front than Burunsgal (average 616Hz) speakers' vowels. The ANOVA reveals that the difference in average F1 values among each

location is statistically significant. Technically speaking, the difference in F1 between Vacha and Burungsal does not quite reach the required cut-off point of $p \leq .05$, but the value is so close at .054 that it can be considered as being a strong trend.

F2, however, does not at first seem to reveal any significant difference between the three locations under study here. If only average second formants are considered, then the difference is certainly not significant (p lies between .409 and .922). Despite all this however, there seems to be a trend in the minimum and maximum F2 values, in which Tashkorgan is the lowest and Burungsal is the highest. This is a trend that is not reflected in the mean values of the vowels and is therefore not reflected in the ANOVA results, since those are based solely on mean values. However, the evidence provided by minimum and maximum scores also should not be ignored. Therefore, it can be seen that there is a trend in this data, but this trend is not (yet) statistically significant.

Table 4.4 Statistical results for [tsaiz] 'what'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	9	535	37	506	564	482	600
Vacha	10	543	89	479	607	394	703
Burungsal	10	616	51	580	653	557	710

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	9	1743	88	1539	1947	1071	1916
Vacha	10	1783	76	1610	1956	1419	2098
Burungsal	10	1776	55	1651	1901	1593	2085

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	**.036**	.951
	Vacha	.965	.930
Vacha	Burungsal	.054	.998
	Tashkorgan	.965	.930
Burungsal	Tashkorgan	**.036**	.951
	Vacha	.054	.998

Table 4.5 Statistical results for [nai] 'no'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	704	53	665	742	604	782
Vacha	10	742	101	669	815	623	869
Burungsal	10	764	94	697	832	630	896

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	1485	89	1421	1548	1353	1612
Vacha	10	1497	145	1392	1601	1119	1585
Burungsal	10	1499	169	1377	1620	1280	1793

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	.304	.975
	Vacha	.616	.981
Vacha	Burungsal	.844	1.000
	Tashkorgan	.616	.981
Burungsal	Tashkorgan	.304	.975
	Vacha	.844	1.000

136

Given this set of data, there does appear to be variation of some dimension, though not exceeding the level required to be significant. Perhaps additional examples would push the results into significance. That is, F1 values in Tashkorgan are relatively lower than those in other locations, meaning that the vowel is more closed in Tashkorgan than in the other two towns. In addition, the F2 values are also lower in Tashkorgan than in either Vacha or Burungsal, indicating that the vowel in Tashkorgan is relatively more central. This result is contrary to expectation given previous reports. Perhaps this indicates that height is more critical to perception than frontness or backness. In other words, perhaps F1 contributes more to perception than F2. Such speculation falls to the realm of other studies, however, leaving the results here simply to note that they are contrary to expectation.

By now, the situation for [nai] 'no' ought to be familiar. Like the previous two words examined here, [nai] does show a trend where Tashkorgan demonstrates lower F1 values than the other two values, but that change is not statistically significant. Nevertheless, minimum and maximum values do show a persistent increase in F1 values from Tashkorgan to Vacha to Burungsal, indicating variation.

The situation for F2 is also nearly identical to that seen above. F2 values generally increase as one travels from eastern to western locations, but the change is very small. In addition, the lowest value for F2 in Vacha clearly crosses over into the range of the Tashkorgan F2 values. Even though Vacha's overall mean F2 is higher than Tashkorgan's mean F2 value, Vacha's minimum and maximum scores are lower

than Tashkorgan's. Ironically, this is what was originally expected, but not what has been seen thus far in the data. As segments increase in F1, indicating a more open articulation, they are also expected to move farther back in the mouth, as would be indicated by a decrease in F2. This is what is happening here with Vacha. Vacha's segment is relatively more open and central than that in Tashkorgan. The only difficulty here is that this is not what has been seen in the trend thus far. For example, Burungsal segments do seem to be more open than those in the other two locations, but they do not seem to move to a more central location. If anything, they do the opposite and also move more forward in the mouth, though also more open.

Given the ambiguity of the data thus far, it might be illuminating to examine a scatter plot of the F1 and F2 values for all three words together. Looking at the scatter plot in Figure 4.2 does provide some clarification on the issues above. It is clear, for example, that the values for all three locations overlap considerably. On the other hand, vowels from Burungsal are less common in the extreme upper left-hand corner of the plot, indicating that Burungsal Sarikoli speakers' pronunciations contain a vowel that is indeed lower than their Vacha and Tashkorgan counterparts. Although the difference is not absolute, there is a trend. It also appears that Tashkorgan speakers do have a propensity to cluster near the upper left corner of the scatter plot, which would indicate that more of them are using a vowel that is higher than the other two locations. In the end, however, these are all only trends -- nothing can be said absolutely.

Log F2-F1
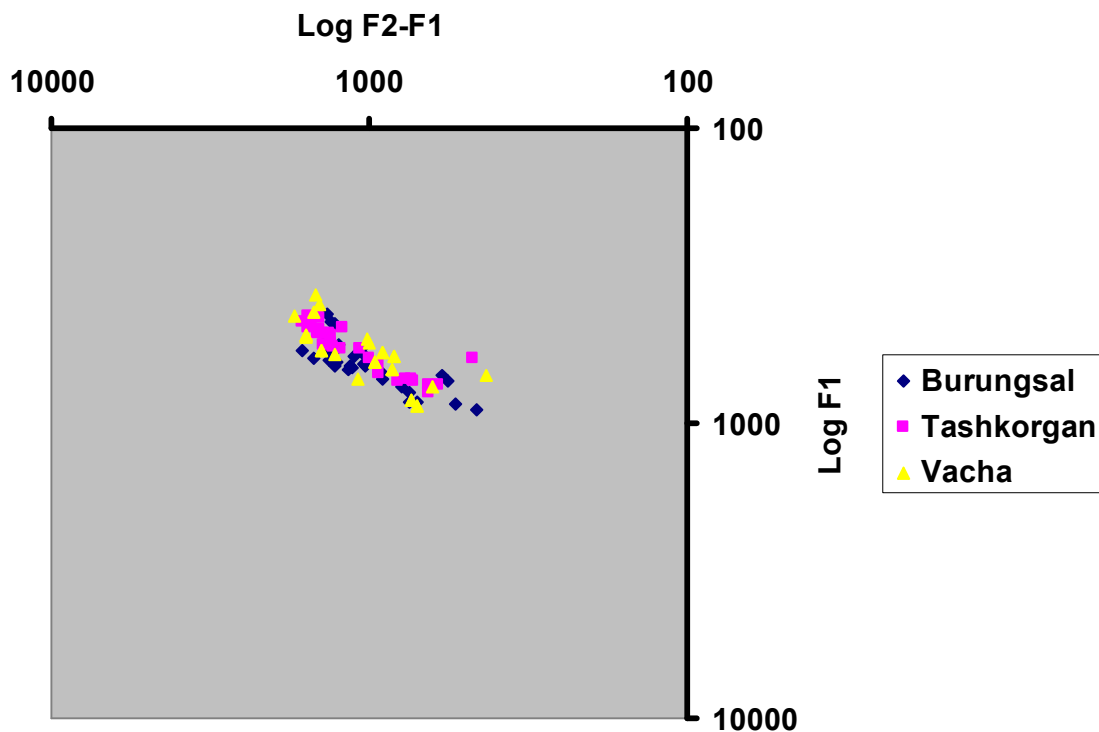
Figure 4.2 Scatterplot of F1 and F2 in [sair] 'full', [tsaiz] 'what', and [nai] 'no'

Even after the analysis of the scatter plot and the ANOVA, results still show a great deal of variation and the answer remains, do these dialects differ in the way proposed by T.N. Pakhalina or not? If the data of location is cross-referenced with the data from individual words, it seems that there is some regional variation, but that alone does not account for all the variation in the data. The real difference is that different words cluster either higher or lower depending on the word as well as the location. That is, all speakers of [nai] 'no' uttered a vowel lower and more central than all the speakers of either [tsaiz] 'what' or [sair] 'full'. If each cluster of each individual word is considered, then Burungsal speakers do tend to cluster on end, though the trend is not

139

absolute. In general Burungsal speakers tend to utter a lower, more central vowel than speakers from either Tashkorgan or Vacha. This can be seen if the data in Figure 4.3 is combined with the information in the scatter plot in Figure 4.2. First, examining Figure 4.3, it is clear that each of the words under review does not occupy the same part of the vowel space. The word [nai] 'no' occupies a space indicating a lower more central vowel than is contained in either [tsaiz] 'what' or [sair] 'full'. Next, the plots for [tsaiz] 'what' overlap considerably with those for [sair] 'full', but the plots for [sair] 'full' have a tendency to cluster in the closer and higher part of the vowel space. Finally, if it were possible to combine this information from that in Figure 4.2, it would show that in each word cluster the plots for Tashkorgani speakers cluster in positions indicating higher more front vowels. That is to say, although all speakers pronounce [nai] 'no' with a lower more central vowel than the other words, Tashkorgani speakers produce that word with the highest and most front pronunciations of any of the three locations. Likewise, the vowels in [tsaiz] 'what' or [sair] 'full' are higher and more front among all speakers, but Tashkorgani speakers utter higher more front vowels than speakers from other locations.

In light of this, it does seem that there is a continuum of variation with Tashkorgan on one end and Burungsal at the other. Previous accounts of this dialect variation failed to account for the immense diversity involved and failed to highlight the gradient nature of that relationship. In addition, previous analyses did not realize that all

pronunciations of [nai] 'no' are lower and more central than the vowel used in the other words.

This account also adequately explains the ANOVA results above. In those results, the differences between Tashkorgan and Burungsal were significant or nearly significant for both [sair] 'full' and [tsaiz] 'what'. However, there were no significant differences found between locations for [nai] 'no'. It appears now that [nai] 'no' is quite a different vowel from that found in the other two words and seems to have a more homogeneous pronunciation than either [sair] 'full' or [tsaiz] 'what'. Recall also from Chapter 3 that it was argued there that both [sair] 'full' and [tsaiz] 'what' mosr closely match the schema for long monophthongs showing drift while [nai] 'no' more closely matches the schema for a diphthong. It appears here that variation in diphthong pronunciation is less varied than variation in the pronunciation of long monophthongs.

To summarize the results so far, it is perhaps best to say that there are some trends indicated in the data, but no absolutely clear demarcation can yet be drawn between the three locations. If the statistical powers of the ANOVA alone are relied upon, then the tokens under investigation can be said to have been drawn from one population of speakers and not from different populations. However, as the scatter plot shows and as the descriptive statistics show, there are definite trends in the data that seem to provide some tentative confirmation of Pakhalina's proposals. Perhaps analysis of further segments will also provide some cogent data.

Log F2-F1
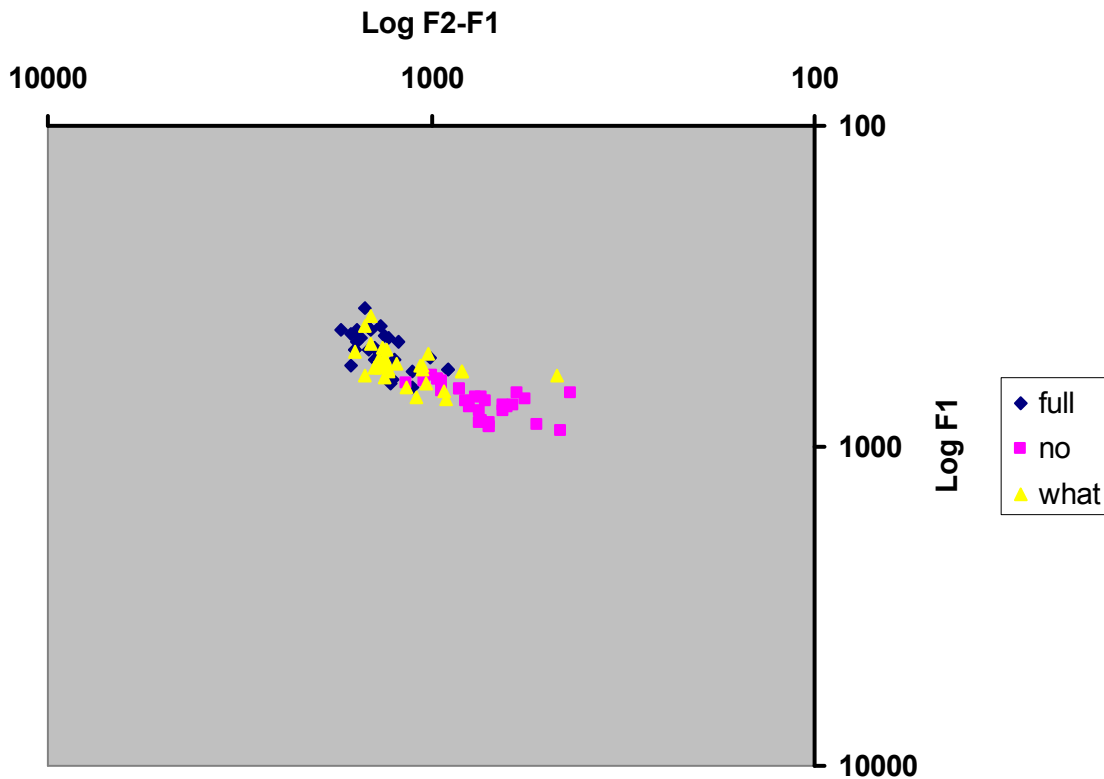
Figure 4.3 Scatterplot of F1 and F2 in [sair] 'full', [tsaiz] 'what', and [nai] 'no' sorted by word

4.3.2 SEGMENT [ɛi]

If attention is now turned to results of the analysis of the segment [ɛi], some special difficulties are presented. These difficulties arise because the analysis of results in Chapter 3 showed that this segment was not a diphthong at all but a long vowel. Thus, what this section will examine is not whether or not this segment is a diphthong but rather whether the segment shows geographical variation or not. Even though this segment is no longer understood here as a diphthong, whether or not it varies by

location is important to making statements about its potential path of development. Thus, this segment will be further examined here even though it did not fit the prototype of a diphthong in Chapter 3.

Recall from page 125 that only one word was claimed to contain this diphthong. Results for that word [spɛid] 'white' are presented in Table 4.6. Descriptive statistics for F1 show a trend whereby Tashkorgan shows a relatively low F1 value, followed by Vacha in the middle and then with Burungsal at the other end of the spectrum, showing the highest F1 mean value. It is also interesting to note that Vacha exhibits a great deal more variation than speakers in the other two locations as reflected by the Standard Deviation scores, the 95% Confidence Interval and also the minimum and maximum scores. Descriptive statistics for F2 also reveal a pattern in which Tashkorgan has a relatively higher F2 value while Burungsal records the lowest F2 values. Looking at these statistics alone, it appears that there is a trend in this segment in which Tashkorgani speakers tend to utter a higher more front vowel than speakers at the other two locations. Burungsal speakers appear to utter a segment relatively lower and more back than in the other two locations.

On the other hand, even though there is a trend in the data, such differences are not actually all that great. The results of the ANOVA show that Tashkorgan and Burungsal do differ statistically along F1 values, but no other values are statistically significant. In addition, if the mean values for F1 and F2 are compared to the model values for vowels in Table 4.2, it can be seen that the differences between vowel spaces

143

is indeed small. In particular, the F1 values for Burungsal at 598Hz fall quite a bit short of the model value of 700Hz needed to be either [æ] or [a]. Meaning that considering the segment in Burungsal to be [a] is perhaps a bit extreme. Nevertheless, the trend is real and it is being revealed in the data. Although the trend is not yet significant enough to definitively declare the segments different, the trend also cannot be ignored.

Table 4.6 Statistical Results for [spɛid] 'white'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	461	35	436	487	415	542
Vacha	10	530	109	451	608	323	635
Burungsal	10	598	43	566	629	537	661

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	1891	230	1726	2056	1241	2010
Vacha	10	1808	88	1744	1871	1629	1945
Burungsal	10	1734	137	1636	1833	1540	1972

Table 4.6 - *Continued*

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	**.001**	.120
	Vacha	.119	.529
Vacha	Burungsal	.122	.610
	Tashkorgan	.119	.529
Burungsal	Tashkorgan	**.001**	.120
	Vacha	.122	.610

This trend in the data can be seen even more clearly in the scatter plot in Figure 4.4. In this case, all the speakers from Tashkorgan cluster together in a fairly small area, indicating little variation among speakers from Tashkorgan. On the other hand, Vacha speakers show a wide amount of variation, but perhaps only because of one speaker being unusual. Most speakers from Vacha and Burungsal show intermingled distribution, leading to the conclusion that their two varieties are nearly identical for the time being at least as far as this segment is concerned. The scatter plot taken together with the previous analyses leads to the conclusion that there are differences in articulation between speakers of the three areas, but that difference is generally small and not as large as might be indicated by Pakhalina's transcriptions. Nevertheless, the data are clearly indicative of a trend -- a trend that mirrors what was expected. Namely,

145

speakers in Tashkorgan are on one end of a continuum, uttering a higher more front vowel than speakers in other two locations. Although speakers from Burungsal and Vacha overlap a great deal, there is nevertheless a trend there as well in which speakers from Burungsal occupy the other end of the continuum from speakers in Tashkorgan, uttering vowels that are generally lower and more central or back than speakers in other locations.
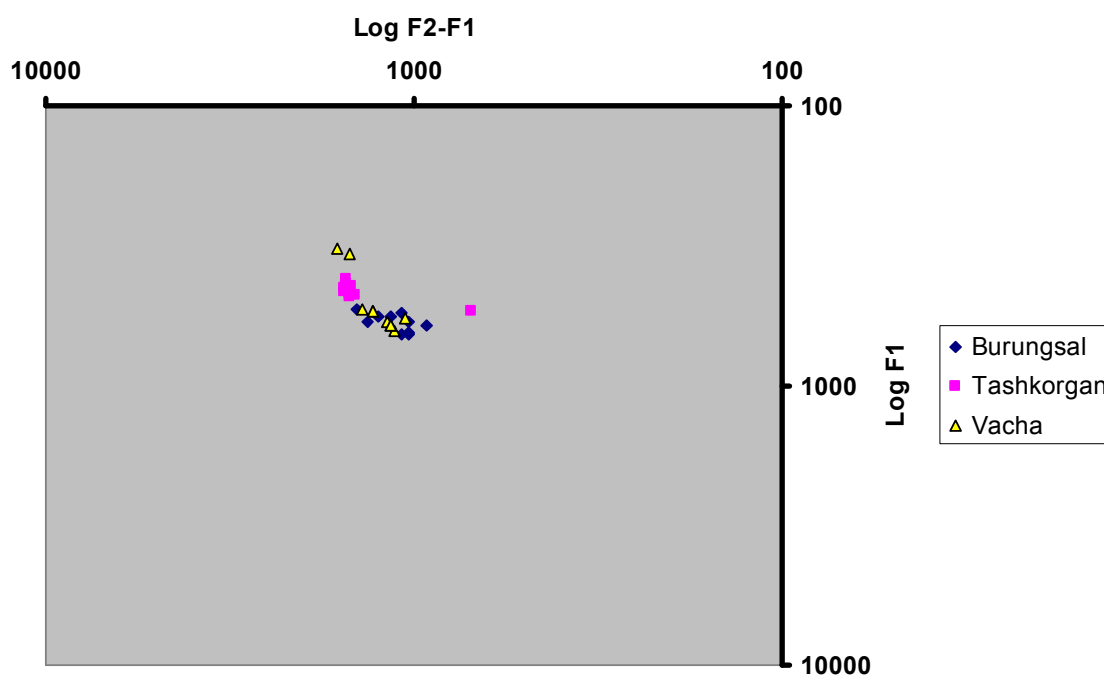
Figure 4.4 Scatter plot of F1 and F2 in [spɛid] 'white'

4.3.3 SEGMENT [ɛu]

The next segment examined is [ɛu]. Results of Chapter 3 showed that words previously thought to contain this diphthong actually contain a vowel followed by a semi-vowel. As such, the main point to be considered here is whether the vowel in this pair shows variation across dialects. The semi-vowel is not under consideration and was not included in measurements taken of the three words in this study that contain this segment. Each of these words will first be examined in turn. Then, a comparison will be made across words and across locations to reach a conclusion about regional variation.

Examining statistical results in Table 4.7 for [nɛu] 'nine' reveals a trend similar to that found above. F1 does increase gradually as one moves from the east in Tashkorgan to the west in Burungsal. F2 does not show as clear a pattern as seen previously since Vacha had a higher mean F2 value than the other two locations. Both Vacha and Burungsal show a great deal of variance in their F2 values as shown by the standard deviation. Tashkorgan speakers exhibit much more homogenous pronunciations overall. The only difference found to be statistically significant was that between F2 values of Vacha and Burungsal.

Table 4.7 Statistical results for [nɛu] 'nine'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	408	44	376	440	352	468
Vacha	10	436	50	400	472	376	531
Burungsal	10	443	78	387	500	370	637

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	2045	112	1964	2125	1903	2250
Vacha	10	2213	169	2092	2334	1943	2447
Burungsal	10	1992	171	1869	2114	1733	2285

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	.426	.745
	Vacha	.579	.066
Vacha	Burungsal	.964	**.012**
	Tashkorgan	.579	.066
Burungsal	Tashkorgan	.426	.745
	Vacha	.964	**.012**

148

Analysis of [vrɛu] 'eyebrow' does not reveal any very surprising results at this point. There is a definite trend in F1 in which mean F1 values increase when beginning in the east and moving west. In this case, all speakers exhibited less variance than in the previous word however, with standard deviation scores being much lower than those for [nɛu] 'nine'. In addition all F1 values were higher in this case than in the previous case. In [nɛu] 'nine', all F1 values were in the neighborhood of the model formant value for [e]. In [vrɛu] 'eyebrow', however, F1 values were more in keeping with model values for [ɛ] and even bordering upon the model values for [æ]. The fact that this vowel is more back in the overall vowel space than that in [nɛu] 'nine' is also confirmed by overall lower F2 scores, indicating that this vowel is more central than that in [nɛu] 'nine'. Standard deviation values were also quite low indicating more homogeneity among speakers. F2 values here showed the same counter to expectation trend as above; with Vacha exhibiting a higher mean F2 than Burungsal.

In this case, the ANOVA found that differences in F1 values between all three places to be statistically significant, but the F2 values did not meet the test of significance at the $p \leq .05$ level. The mean F1 values do indeed cover a large amount of vowel space and do seem to nearly coincide with those segments previously proposed. It is worth pointing out that the mean F2 difference between Tashkorgan and Vacha is nearly significant, but does not quite meet the standard.

Table 4.8 Statistical results for [vrɛu] 'eyebrow'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	511	28	490	532	464	555
Vacha	10	545	63	500	590	484	665
Burungsal	10	619	48	584	654	566	707

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	1451	94	1384	1519	1277	1598
Vacha	10	1569	82	1510	1628	1386	1705
Burungsal	10	1492	130	1399	1586	1307	1733

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	**.000**	.689
	Vacha	.307	.058
Vacha	Burungsal	**.009**	.274
	Tashkorgan	.307	.058
Burungsal	Tashkorgan	**.000**	.689
	Vacha	.009	.274

Turning attention now to the last word in this section to be analyzed, it can be seen that the vowel segment in [tʃabɛud] 'pigeon' occupies a position intermediate between the vowels found in the other two words. Like the previous two examples, F1 values do increase when moving from Tashkorgan to the other two locations. This time, however, the mean F1 for Vacha and Burungsal are nearly identical. In addition, the values for F1 place it somewhere between the model values for [e] and [ɛ], but certainly nowhere in the vicinity of [æ]. This indicates that the vowel in [tʃabɛud] 'pigeon' is overall lower than that in [nɛu] 'nine' and higher than that in [vrɛu] 'eyebrow'.

Somewhat surprising here are the F2 values. Contrary to expectation, they actually increase when moving from Tashkorgan to Vacha and then to Burungsal. Furthermore, they are sufficiently low as to question whether the vowel is even a front vowel at all. In addition, if this is truly a back vowel and not a front vowel, then the data fit into the overall pattern more elegantly. It appears that Tashkorgan has a tendency to pull vowels towards the extremes of the vowel space. Thus, in front vowels, Tashkorgan Sarikoli speakers will have a higher F2 value, pulling the vowel forward to the more extreme position. Likewise, with back vowels, Tashkorgan Tajiks will pull the vowel back to the more extreme position. Speakers from Vacha and Burungsal always utilize vowels in the more neutral center positions, no matter whether they are front or back.

None of these results were deemed statistically significant under analysis of variance. This is not surprising considering what a small range of vowel space the mean values represented.

Table 4.9 Statistical results for [tʃabɛud] 'pigeon'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	506	45	473	538	425	590
Vacha	10	558	91	492	624	401	705
Burungsal	10	558	53	519	596	457	639

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	1179	100	1107	1251	1024	1325
Vacha	10	1292	108	1215	1370	1153	1436
Burungsal	10	1258	107	1181	1335	1096	1457

(c) Post-Hoc Scheffé Result s

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	.234	.263
	Vacha	.233	.075
Vacha	Burungsal	1.000	.776
	Tashkorgan	.233	.075
Burungsal	Tashkorgan	.234	.263
	Vacha	1.000	.776

It is has become quite clear at this point that the primary difference between these words is not a regional variation but simply that the three words contain vowels that occupy different places in articulatory and acoustic space. It seems that the vowel in [nɛu] 'nine' is different from that in the other two words. This can be seen clearly in Figure 4.5 in which all the values for [nɛu] 'nine' cluster in the lower right section of the plot. Moreover, within that plot, there is no apparent pattern to the way different locations pronounced the vowel in that word. The other two words, [vrɛu] 'eyebrow' and [tʃabɛud] 'pigeon', clearly cluster together in the upper left section of the scatter plot. Even here, however, these words have very few interspersed values in regard to F1 and F2. While there is considerable overlap in the height of the two vowels, nearly all instances of [tʃabɛud] 'pigeon' contain a vowel that is articulated farther back in the mouth than that in [vrɛu] 'eyebrow'. In the case of [tʃabɛud] 'pigeon' and [vrɛu] 'eyebrow', there does seem to be some trend in geographical differences. Examining the higher F1 values reveals that there are many plots from both Vacha and Burungsal in that area but none from Tashkorgan, meaning that Tashkorgani speakers trend toward a higher vowel with a concomitant lower F1 value than speakers in the other two locations.
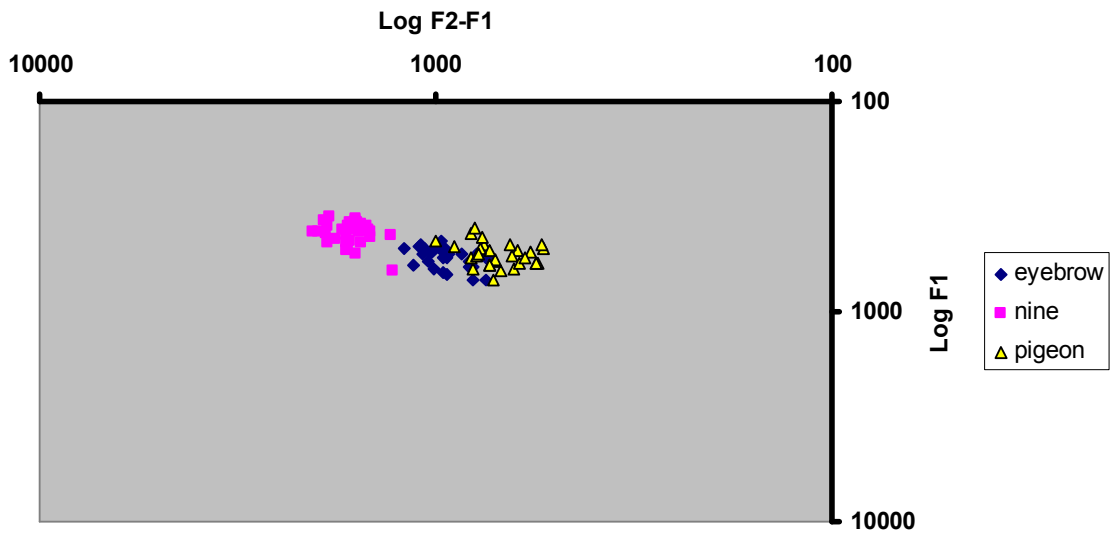
Figure 4.5 Scatter plot of F1 and F2 values of [nɛu] 'nine',
[vrɛu] 'eyebrow' and [tʃabɛud] 'pigeon'

These differences between the actual vowels in the words was confirmed by use
of an ANOVA comparing the mean F1 and F2 scores of all three words without regard
to location. Results show that there is a significant difference between [nɛu] 'nine' and
the other two words (cf. Chapter 3). Comparisons of [vrɛu] 'eyebrow' and [tʃabɛud]
'pigeon' did not reveal a statistically significant difference between those two words
when F1 was considered. However, when F2 was considered, the difference between
them was statistically significant. Since a vowel depends upon both F1 and F2 for its
formation, it is reasonable to assume that if the vowel should differ along either F1 or
F2 then it ought to be considered a different vowel. Post-Hoc Scheffé tests on F2 also
grouped the words into three different groups, once again reflecting the statistically
large difference between each of the words. In this case the statistical tests help to

154

underline a trend that was already clear in the linguistic data. These vowels do have differences in their articulation.

Table 4.10 ANOVA results for [nɛu] 'nine', [vrɛu] 'eyebrow' and [tʃabɛud] 'pigeon'

Dependent Variable	(I) Word	(J) Word	Sig.
F1	nine	eyebrow	**.000**
		pigeon	**.000**
	eyebrow	nine	**.000**
		pigeon	.572
	pigeon	nine	**.000**
		eyebrow	.572
F2	nine	eyebrow	**.000**
		pigeon	**.000**
	eyebrow	nine	**.000**
		pigeon	**.000**
	pigeon	nine	**.000**
		eyebrow	**.000**

There is no *a priori* reason to expect that each of these words would contain vowels that are statistically different. All previous researchers have agreed that these words contain the same vowel (at least within a certain dialect region). Therefore, these

results are surprising. There is also nothing in the environment of the vowels in these words that would be expected to have an unduly large influence on the point of articulation of the vowel. That is to say that the vowels do not occur in identical environments, but the environments are similar and none contain any kind of consonant that would normally have any unduly large lowering or backing effect on its neighboring vowel.

What of the dialect variation? Like the preceding segments studied here, these segments do seem to show some slight variation across dialects, but it is small. It is certainly not of the dramatic nature that would warrant the kind of transcription differences that Pakhalina proposed. Nevertheless, her argument is not totally without merit either. There is a trend, albeit one that is statistically insignificant and one that cannot clearly be seen in the scatter plot. Given such a tenuous trend, it would be best to say that it is possible to define trends, but one that will require more attention in the future. Is it possible that the dialect variation was greater 50 years ago when Pakhalina did her research? It is possible, but there is no way to successfully evaluate that since there were no instrumental analyses done at that time. All that can be confirmed is that, synchronically speaking, there are differences in the three regions, but those differences are currently small. Speakers from Vacha have a tendency to utter more generally lower and more central vowels in these three words than speakers from Tashkorgan. These differences are not generally great enough to move the vowel into the space of a

neighboring vowel and are not statistically significant, but the differences are observable.

4.3.4 SEGMENT [eɪ]

If many of the previously examined diphthongs showed trends, but trends that were not strong enough to be statistically different, then what will the outcome be if the diphthong [ei] is considered? In this case, there were four words considered: [deig] 'pot', [tʃeig] 'to do', [deig] 'shovel', and [beid] 'to disappear'. Each of these words will be analyzed and examined in turn.

Beginning with [deig] 'pot', results are quite similar to those previously seen for other diphthongs. Descriptive statistics for [deig] 'pot' can be seen in Table 4.11. F1 does show an increase from Tashkorgan to Vacha to Burungsal. In this case, the difference between F1 was also deemed statistically significant when comparing Tashkorgan and Burungsal. These differences would indicate a vowel in the range of [e] in Tashkorgan, a vowel approaching [ɛ] in Vacha, and a low [ɛ] moving into [æ] in Burungsal.

F2 results for [deig] 'pot' are likewise similar to results seen in previous words. Once again F2 values are lowest in Burungsal, but the difference is not statistically significant. Nevertheless, taken with the differences in F1 values, it does appear that there is a trend in which Tashkorgan speakers have a tendency to use a higher and more front vowel than speakers in the other two locations.

Table 4.11 Statistical results for [deig] 'pot'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	468	39	439	496	423	526
Vacha	10	522	83	463	582	364	616
Burungsal	10	549	72	498	601	462	669

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	2017	61	1974	2061	1912	2100
Vacha	10	2005	107	1928	2083	1859	2197
Burungsal	10	1903	188	1768	2038	1650	2197

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	**.039**	.166
	Vacha	.213	.979
Vacha	Burungsal	.674	.232
	Tashkorgan	.213	.979
Burungsal	Tashkorgan	**.039**	.166
	Vacha	.674	.232

Table 4.12 Statistical results for [tʃeig] 'to do'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	461	40	432	490	412	544
Vacha	10	510	55	470	550	436	609
Burungsal	8	541	73	480	603	486	658

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	2022	111	1942	2102	1743	2138
Vacha	10	1890	141	1788	1991	1589	2051
Burungsal	8	1916	226	1727	2105	1684	2347

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	**.022**	.398
	Vacha	.175	.206
Vacha	Burungsal	.516	.942
	Tashkorgan	.175	.206
Burungsal	Tashkorgan	**.022**	.398
	Vacha	.516	.942

Table 4.13 Statistical results for [beil] 'shovel'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	468	39	439	496	423	526
Vacha	10	522	83	463	582	364	616
Burungsal	10	549	72	498	601	462	669

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	2017	61	1974	2061	1912	2100
Vacha	10	2005	107	1928	2083	1859	2197
Burungsal	10	1903	188	1768	2038	1650	2197

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	**.039**	.166
	Vacha	.213	.979
Vacha	Burungsal	.674	.232
	Tashkorgan	.213	.979
Burungsal	Tashkorgan	**.039**	.166
	Vacha	.674	.232

Table 4.14 Statistical results for [beid] 'to disappear'

(a) Descriptive Statistics for F1

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	493	47	459	527	392	585
Vacha	10	549	88	486	612	388	643
Burungsal	10	580	69	530	630	485	672

(b) Descriptive Statistics for F2

	N	Mean	Std. Deviation	95% Confidence Interval for Mean		Minimum	Maximum
				Lower Bound	Upper Bound		
Tashkorgan	10	1923	79	1866	1980	1745	2031
Vacha	10	1872	110	1793	1950	1718	2134
Burungsal	10	1791	100	1719	1863	1680	1992

(c) Post-Hoc Scheffé Results

(I) Location	(J) Location	Sig. F1	Sig. F2
Tashkorgan	Burungsal	**.035**	**.020**
	Vacha	.229	.510
Vacha	Burungsal	.613	.201
	Tashkorgan	.229	.510
Burungsal	Tashkorgan	**.035**	**.020**
	Vacha	.613	.201

161
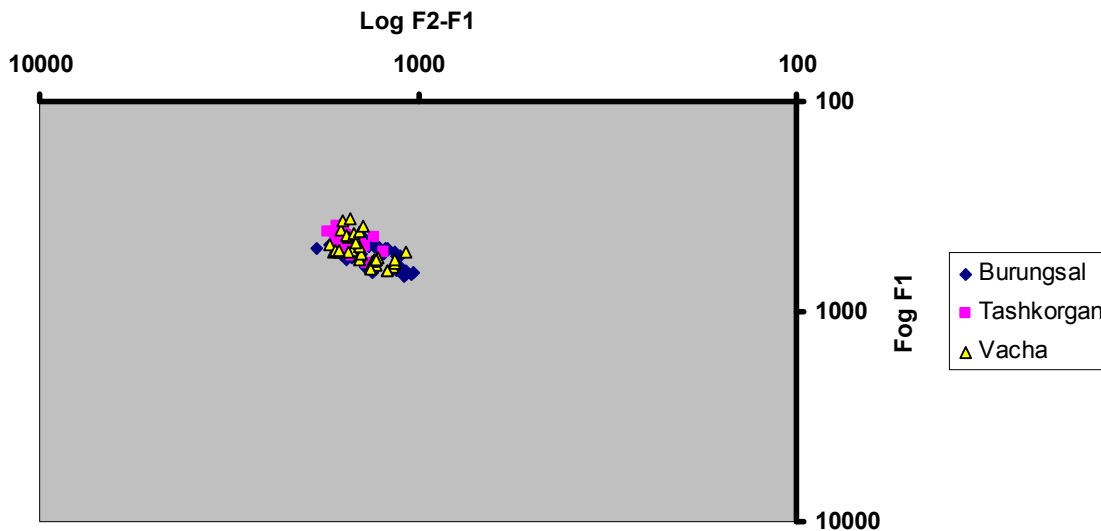
Figure 4.6 Scatter Plot of [deig] 'pot', [tʃeig] 'to do', [beil] 'shovel',
and [beid] 'to disappear' sorted by location

This trend remains true for all three of the other words tested in this section. All three words contain a vowel that is highest and farthest front in Tashkorgan, at a middle level in Vacha, and is relatively low and less front in Burungsal. Likewise, all the results were only deemed statistically significant when comparing the F1 values of Tashkorgan and Burungsal. F2 values for [tʃeig] 'to do' were not quite consistent with other trends in that mean values for Vacha were lower than for those in Burungsal. The only statistical exception to this was in [beid] 'to disappear' in which F2 values were also deemed statistically significant when comparing Tashkorgan and Burungsal. All these results can be seen in Table 4.12-Table 4.14.

Finally, when all the data for each lexical item is combined and sorted by location (Figure 4.6), it can be seen that the important variable is location. When such a

scatter plot was constructed with each word as the grouping variable, results were random, without any trend or pattern. In Figure 4.6, however, it can be seen that, although there is considerable overlap among data points, values for Burungsal cluster near the lower right-hand quadrant, indicating an overall trend toward containing lower, more central vowels. Likewise, most values for Tashkorgan cluster in the upper left-hand section of the scatter plot, indicating a trend towards higher, more front vowels. As expected, values for Vacha have a tendency to group in the middle. However, statistics indicate that such differences are not consistently statistically significant, and the scatter plot seems to also support such a conclusion.

4.4 DISCUSSION AND CONCLUSIONS

In light of the results of each of the segments analyzed above, what conclusions can be drawn? Does Sarikol Tajik truly have three lects that differ primarily in the height and frontness or backness of onset vowels in diphthongs? Temporarily setting aside the question of whether these segments are diphthongs and concentrating solely on the differences in the segments themselves, the answer seems to be in the affirmative. Although the differences may not yet be strong enough to be declared statistically, the differences are seen in the descriptive statistics. It appears that Sarikol Tajik has an eastern dialect in Tashkorgan, a western dialect in Burungsal, and a transitional dialect in Vacha. All of the results are summarized on page 164 in which results are summarized by means of IPA symbols. Although actual mean values for some segments fell between the iconic values for some segments, the IPA symbols are

163

still a useful shorthand for summarizing all the data. The symbols in the data on page 164 reflect that there are actual differences between the three regions, but those differences are not as great as the transcriptions originally chosen by Pakhalina would have suggested. In addition, Pakhalina's transcriptions have also been modified to reflect the determination about whether the vowel segments in question were diphthongs or not.

Summary of dialect analysis

	Segment	Central (Tashkorgan)	Near Eastern (Vacha)	Far Eastern (Burungsal)	
(5) a.	[ai]	[seir]	[sẹir]	[sɛir]	'to be full'
b.		[tsɛiz]	[tsẹiz]	[tsæiz]	'what'
c.		[næi]	[nẹi]	[nai]	'no'
(6)	[ɛi]	[spẹːd]	[spɛːd]	[spẹːd]	'white'
(7) a.	[ɛu]	[new]	[nẹw]	[nẹw]	'nine'
b.		[vrẹw]	[vrɛw]	[vrẹw]	'eyebrow'
c.		[tʃabẹwd]	[tʃabɛwd]	[tʃabɛwd]	'pigeon'
(8) a.	[ei]	[deːg]	[dɛːg]	[dẹːg]	'pot'
b.		[tʃeːg]	[tʃɛːg]	[tʃẹːg]	'to do'
c.		[beːl]	[bɛːl]	[bẹːl]	'shovel'
d.		[bɛːd]	[bẹːd]	[bẹːd]	'to.disappear'

It is well known that dialect variation does not occur in tight, well-separated areas. Dialect variation can be messy and inexact, making it difficult or impossible to create isoglosses from actual data in hand (Bailey 1996, Chambers 2001). Instead, there seems to be a pattern of lexical diffusion in which sound patterns originate at a source and change gradually over time via a process of analogy (Phillips 1999). Bybee argues, 'The study of the diffusion of sound change in the lexicon contributes to a better understanding of the nature and causes of sound change' (Bybee forthcoming). Bybee further argues that sound changes affect the deepest levels of mental representation and provide an ongoing record of the changes that are in progress. This record of changes is produced because sound change is regular but its spread is diffusional showing earlier in some lexical items than in others (where frequency of use is considered an important variable) (Bailey 1996, Bybee forthcoming).

It is therefore useful to examine what the data are revealing about the current status of sound change in these lects of Sarikol Tajik. When examining the data, it seems that it might be best to consider Tashkorgan as one lect in contra-distinction to the lect spoken in Burungsal while the lect spoken in Vacha can be seen as a transitionary lect between the two. Such a conclusion seems warranted by the statistical data examined above. Firstly, there was a trend quite clear in the descriptive statistics whereby vowels were higher and more front in Tashkorgan and lower and more back in Burungsal. In the eleven words examined above, this difference was found to be statistically significant eight times. Even laying aside the question of the statistically

significant difference between these two areas, the actual difference as measured by mean F1 and F2 scores was absolute, with differences being reported 100% of the time. Most of the time, the statistically significant difference was in F1 values rather than F2, perhaps indicating that the salient difference between the segments is the height and that the relative frontness or backness is simply a reflection of that difference in height. Secondly, the differences between Tashkorgan and Burungsal can be seen in the scatter plot diagrams. At either extreme, the differences become absolute. That is to say that in the scatter plots in the section indicating lower more central vowels there are only plots representing Burungsal - those sections never contain plots representing Tashkorgan. The converse is also true: in plots representing higher more front vowels, there are always clusters of plots from Tashkorgan but never any from Burungsal. There are still individual speakers that do not match the overall trend completely, but the trend is clear.

As for Vacha, it appears to be a transitional area between two more well-established dialect areas. It is well known that speakers in transitional areas often show variable pronunciations and may use one pronunciation on one occasion and another on another occasion (Chambers 2001). This would explain why speakers from Vacha often show a very large standard deviation, use vowels that contain minimum and maximum scores that cross over onto the other two areas, and why Vacha speakers plots in the scatter plots are found in the same areas as both Tashkorganis and Burungsalis. Even though there is a great deal of crossover, Vacha speakers' scores are never at either extreme. This is expected of neighbors that need to accommodate both

Tashkorgani speakers and also Burungsali speakers. This accommodation often causes speakers of transitionary dialects to move freely from one dialect to another even if their own overall dialect is somewhere between the two dialects on their flank. This would help to account for why there are so many more outlying values for formants spoken by Vacha speakers than by those spoken in other areas.

Given that there is variation among the three areas, is it possible to tell where the sound change began? In other words, which dialect is the most conservative, reflecting an older form? Which dialect is showing greater changes? Is it possible to identify the origin of the change? Given the data, the origin of the change must be either in Tashkorgan or in Burungsal, but which is it? This is the topic of Chapter 5.

CHAPTER 5

HISTORICAL AND DEVELOPMENTAL ANALYSIS

5.1 THE PROBLEM

Having once established both spectrographically and statistically the nature of the segments under consideration in the words in this study, it should then be possible to analyze their developmental patterns. That is to say, it should be possible to analyze whether Sarikoli embraces change more or less rapidly than the other languages in the language group. Chinese scholars working on Sarikoli have often remarked (in personal conversations) that Sarikoli is more conservative to change than the other languages in the family, but is this truly the case? Such a decision about whether Sarikoli is more or less conservative than the other Pamiri mountain languages must be based on many phonological and historical factors. Nevertheless, it does seem beneficial to begin to offer some preliminary hypotheses about the path of development of Sarikoli Tajik at least in light of the diphthongal evidence offered here. Any such conclusions based solely on examining one aspect of the language are necessarily tentative, but such an analysis should at least serve as a beginning point for further studies.

To determine whether a language is more progressive or more conservative in embracing change than its neighbors and to plot its general path of development, it is

necessary to know both the present state of the language and as much about the history of the language as possible. It is also helpful to know as much as possible about the other languages in the group/family as possible. In this case, the ancient predecessor of the Pamiri mountain languages, Avestan (330 BC 225 AD), is a language that has been well documented and well studied. It is also unique in using a generally straightforward and very consistently phonological alphabet. Because Avestan was (and is) a sacred language of the Zoroastrian religion, the writing system was codified at an early stage. This enabled priests in far-flung places to be able to look at the sacred texts and know how to pronounce them correctly. Not only does such an orthography serve worshippers well, it has been a great boon to linguists too.

Avestan also has other advantages as a language of historical and phonological comparison. Firstly, it belongs to the oldest strata of Iranian languages. Linguists generally divide Iranian languages into three eras, called respectively 'old', 'middle', and 'new'. Avestan belongs to the oldest of these groups (Rezakhani 2001). This allows for a view across thousands of years of time. Secondly, Avestan belongs to the same sub-branch of Iranian languages as Sarikoli, indicating a close relationship between the two of them. Persian languages are divided roughly along east-west lines, with Sarikoli being a part of the eastern branch. Avestan is the major eastern Iranian language in the Old Persian era. Finally, much is known about Avestan. There are other more modern antecedents to Sarikoli from the Middle Iranian era (e.g. Sogdian, Bactrian, and Khotanese), but not as much is known about these languages as about Avestan. It may

prove useful as more is learned about these other languages to use them as a basis of comparative analysis, but that time has not yet come. Thus, because much is known about Avestan and many studies of the language have been conducted (see the introduction to de Vaan 2003 for a lengthy bibliography of relevant studies), Avestan makes a convenient and useful starting point for this study. In addition, those who have attempted any kind of etymological or historical study of any Pamiri mountain languages or of Sarikoli in particular have used Avestan as their basis of comparison (Morgenstierne 1974, Pakhalina 1960, Pakhalina 1969, Sokolova 1967).

Among those linguists who have used Avestan to study the history of the Pamiri mountain languages, all of them seem to have been most heavily influenced by the Structuralist or Neogrammarian school of linguistics. This can be seen by the way in which they present their analyses and results. For example, they generally present a table of sound correspondences that account for the facts of the languages in study, but provide little or no explanation of why the changes might have taken place. That is to say, they present their data in terms of splits and mergers in particular vowels, an approach consistent with the Structuralist tradition (McMahon 1999). An example of such an analysis can be seen in Table 5.1. The sound changes that Pakhalina proposes in each language do not seem to have any motivation but simply exhibit linguistic drift or regular sound change over time. When necessary, some rules are offered to account for the influence of surrounding segments, but most sound changes are offered as being unaffected by the particular segments surrounding them. In addition, there is very little

170

interest expressed in either explaining a change in terms of either rules or socio-linguistic phenomenon nor are any predictions made about potential future paths of development of any of the languages under analysis. Indeed, the data is presented in such as a way as to suggest that there perhaps really is no particular organizational pattern to the change nor is it possible to predict the future based on such a wide variety of apparently unpredictable sound changes.

Table 5.1 Correspondences of the major Pamiri mountain languages and Avestan

Sarikoli	Shughnan	Roshani	Bartangi	Avestan
		o	ɵː	
e	iː	eː, æ	eː	ai, eː,iː, i
ɛj, i	eː	iː	iː	ai, eː
a	aː	aː	aː	aː
o	ɛː	eː	eː	aː
u	oː	oː	oː	au, oː
ɛw	ʊː	ʊː		au, oː
ə	uː	uː	uː	aː, a
ɯ	u	u	u	aː, a
i	i	i	i	iː. i
a	a	a	a	ao, eː, iː, i

This figure is adapted from: (Pakhalina 1960, Pakhalina 1983)

During the time in which Pakhalina was conducting her major studies of Sarikoli in Moscow, a revolution of sorts was happening in linguistics in America and secondarily in Europe. The theories of Chomsky were gaining ever more widespread acceptance among linguists around the world. This caused a major shift in emphasis

among work that was generally being done at that time. Chomsky, as is well known, is most primarily concerned with studying languages from a synchronic viewpoint and considers the study of one idiolect in an ideal speaker-hearer relationship adequate for linguistic analysis. There were many other linguists not following the Chomskyan paradigm at that time (most notably Labov and other so-called 'variationists'), but their numbers were comparatively few and their influence on linguistics as a whole was less than that of Chomsky and his followers. This created a situation in which, 70 years after the first reliable data on Sarikoli had become available, some fundamental questions had never even been broached. Questions such as: Given what is known about the past and the current state of variation among these languages, is it possible to predict their future path of development? or Is it possible that not all the changes are as uniform as they might appear in the data presented thus far? or Are some of the languages more closely related to each other than others?

Some of these questions were raised by a few linguists, but their advocates were never that interested in Sarikol Tajik. For example, Labov directly confronted the neo-grammarian notion (called the 'regularity hypothesis') that sound change must be exceptionless (Labov 2001a:xiv). In fact, such a hypothesis is in direct contradiction to what dialectologists have always known: there is much variation in the data and changes are not exceptionless (Chambers 2001, Davis 1990). Labov says that 'The proposed resolution of the Neogrammarian Controversy was essentially that changes in phonetic realization of a category were regular, but that changes in category

membership showed lexical diffusion' (Labov 2001b). Such a statement is not meant to negate the fact that there are often overwhelming trends in the data, but trends should not be construed as being exceptionless. On the contrary, to truly account for reality, it is best to simply analyze the trends and identify them. McMahon commented, '[T]he regularity hypothesis need not be discarded, but only be modified; we can no longer make the claim that all sound changes are instantaneous, phonetically gradual and lexically abrupt instead, we must allow some changes to operate by gradual diffusion' (McMahon 1999:228).

While Labov concerned himself with identifying sociolinguistically motivated sound changes, others who would consider themselves to be variationists placed more emphasis on non-sociolinguistically motivated sound change - what has come to be known as 'drift' (Bailey 1996, Chambers 2004b, Fasold 1975). Thus, there are two types of change to be aware of when studying variation (although the following terms are unique to Bailey, the ideas are not):

1. CONNATURAL (INTERNAL) - changes that take place without contact with other
language systems

2. ABNATURAL (EXTERNAL)- changes that take places as a result of language
contact

Both types of change are considered completely natural and normal parts of language development[5]. The idea of considering both types of change to be completely

---

[5] Here, the word 'development' is not meant to imply that one lect is better than any other but merely meant to designate the state of a language at a given point in time.

natural can be said to be somewhat unique to this particular sub-set of variationists in that most variationists seem to consider themselves to be sociolinguists rather than historical linguists or dialectologists. This self-eponym is clear in the fact that Labov called his work *Studies in Sociolinguistics* (2001a) and not *Studies in Language Variation*. Moreover, the articles contained in *The Handbook of Language Variation and Change* (2004b) are overwhelmingly occupied with sociolinguistic concerns and the entire book is dedicated to Labov. On the other hand, even recent books on historical linguistics seem to by and large follow the Neogrammarian example of considering all or nearly all language change to be exceptionless (e.g. Beekes 1995, Lehmann 1993, Szemerényi 1999). This indicates that there is a 'blind spot' among linguists in general in which the well-established knowledge of variationists is by and large not impacting the methodology of historical linguists, though they would likely benefit from said knowledge. There are exceptions. For example, the works of Bailey (1996), Chambers and Trudgill (2001), and McMahon (1999) are notable here as each of them consider themselves to be working in different fields (general linguistics, dialectology, and historical linguistics respectively).

Given the desire to study variation and change among such linguists, several new approaches to categorizing and understanding data had to be developed. No longer would the all encompassing charts of the Neogrammarians suffice. Such tables (like those in Table 5.1) failed to show the kind of variation that was possible and failed to show if there was any particular relationship among the particular languages being

studied. Likewise, the rules formulated in the generative tradition were also too powerful and too cumbersome to allow for the observable variation in the data. Therefore, to account for both the observed difference and also highlight the observed similarities and trends in the data, implicational scales were adopted. Implicational scales reveal 'structure in variability' (Rickford 2004). Although their use seems to have declined significantly in recent years, they are still a useful method for analyzing data and can also be used in a modified way to help identify the source of a sound change. The classic (and highly-oversimplified) example of how this works is provided here as a basis for understanding (Bailey 1981:77). For example, if there are four lects which vary from each other in only one feature, they can be represented as in Table 5.2.

Table 5.2 Potential variants in a dialect bundle

3	2	1	0
a	a	a	a
	b	b	b
		c	c
			d

In this representation, lect 0 has all of the changes under study. Lect 1 has sounds changed *a, b*, and *c* but not *d*. Lect 2 has sound changes *a* and *b* but not *c* or *d*. Finally, lect 3 shows only sound change *a*. In such a case, there are two ways to identify which of these lects is the origin for this sound change (although they are really two

ways of saying the same thing). The first is to identify the most developed lect as the center of the change. The second way is to look to see which lect can imply all of the others. In this hypothetical example, lect 0 is considered as the source for all of the others. Item *d* has not yet had time to spread to the other lects and so is considered the newest of the four items. In addition, if a particular lect has item *d* it is also sure to have all the others. This makes the pattern: d ⊃ c ⊃ b ⊃ a. If a lect has a particular feature, it will also have the features to the right on the implicational hierarchy but will not have those to the left. It deserves to be emphasized once again that the change may not necessarily proceed in neat concentric circles from a point of origin.

In addition to implication scales, maps of dialect variation are often still utilized as well. The motivation for this is to show that regional variation might skip over one area and to emphasize visually that results are not consistent from one house or village to the next but show variation. In contrast to the approach taken here, traditional models have relied on the 'wave-theory' and models of isoglosses to study language change. In the wave-theory model, languages develop along lines that provide for relatively neat lines between dialects. It is now felt, however, languages change may but may not, indeed probably will not, result in neat isoglossic boundaries, i.e. a bundle of isoglosses. As such, linguists working in the variationist school of thought still make use of maps showing dialectal difference, but interpret them differently. Maps are used to show geographical dialect continuum and acknowledge that any lines drawn on a map between dialects is often arbitrary (Chambers 2001). The model used here retains the

idea from wave theory that change has a core and spreads from that core but differs in that the spread may not have a neat geographic equivalent. Change may appear to 'jump over' certain areas, providing pockets where a change is more advanced while a neighboring area might appear to be in a less advanced state of change. As such, maps are a useful tool, but are not as useful as the implicational scales.

This dissertation will make primary use of implicational scales in analyzing the Sarikoli data along two lines. The first will be a comparison of the final results obtained in Chapter 3 with the other Pamiri mountain languages and with Avestan to obtain a historical perspective on the data. The second will be an analysis of the dialect variation found in Chapter 4. In both sections, the questions to be answered will be similar: Which of the languages/lects shows the most change? Which language/lect is the most progressive? What is the nature of the relationship between the languages/lects?

5.2 AN ANALYSIS OF THE PAMIRI MOUNTAIN LANGUAGES

Having extolled the virtues of including all aspects of variation in linguistic analysis, this section takes a temporary step back from including all lects of Sarikoli in an attempt to first analyze the historical development of the language. This is done merely as an expedient since all aspects of the language cannot be examined at once, but the individual dialect variations will be examined in Section 5.3. For the sake of simplicity, the forms given near the end of Chapter 3 will be analyzed in this section, limiting the variation analyzed here to that which can be seen among different languages and over time.

There is also, unfortunately, another limitation imposed here - that of lack of accessible data. Data on the Pamiri mountain languages is limited and hard to obtain. There is much information on Avestan, but it remains in the hands of true historical experts (like de Vaan 2003) to fully appreciate all the implications and patterns exhibited in that data. As such, not all of the words previously studied in this dissertation can be included for study in this section of the paper. Although this narrows the level of the claims that can be made from such a limited amount of data, it does not negate the fact that a first foray can be made into analyzing the data from a variationist and developmental perspective. The data included for analysis in this section is presented on page 179. The data that appear there contain data from four Pamiri mountain languages. The data for Sarikol Tajik is in accordance with the results obtained in Chapter 3. The data from the other four languages and Avestan is from a variety of sources, but Sokolova (1967) was the major contributor with Morgenstierne (1974) as the secondary contributor.

In following the implicational pattern scales as outlined above, the data has been analyzed using the following three rules:

Rule 1: Raising, Lowering

Rule 2: Diphthongizing or Shortening or Lengthening

Rule 3: Fronting, Backing

Data analyzed from a historical perspective

	Sarikoli	Shughnan	Roshani	Bartangi	Avestan	
(1)	[spɛːd]	[sipeːd]	[sipiːd]	[sipiːd]	[spaeːta]	'white'
(2)	[iːw]	[(j)i(w)]	[(j)i(w)]	[(j)i(w)]	[aeːva]	'one'
(3)	[toːw]	[tʊː]	[taː]	[taː]	[tava]	'you (sing.)'
(4)	[njɛw]	[noːw]	[naːw]	[naːw]	[nava]	'nine'
(5)	[ʁoːwl]	[ʁuːɣ]	[ʁoːw]	[ʁuː]	[gaoʃa]	'ear'
(6)	[tsaːz]	[tsa]	[tsa]	[tsa]	[tʃitt]	'what'
(7)	[poi]	[paːj]	[paːj]	[paːj]	[pajaːh]	'yoghurt'
(8)	[nai]	[naːj]	[naː]	[naː]	[na]	'no, not'
(9)	[waχen]	[wiχin]	[wiχin]	[waχin]	[wəχuni]	'blood'
(10)	[ðoːw]	[ðu]	[ðaw]	[ðaw]	[dva]	'two'

The first rule is said to apply if the vowel in question is either raised or lowered in relation to the Avestan vowel. Indeed, all the comparisons here assume the Avestan vowel as the basis for contrast since it was the historic predecessor of all of these languages. Since these two processes are mutually exclusive, there did not seem to be any reason to create two separate rules. Nor was the rule fine-tuned enough to categorize for the exact amount of raising or lowering. Since exactly how much to raise or lower the vowel could be affected by other factors in the language and since a more fine-tuned analysis was not revealing, the mere acknowledgement of raising or lowering

was deemed sufficient to reveal trends in the data here. It does deserve to be mentioned that raising seems to be a far more common pattern in the data than lowering overall. The second rule is said to apply if there is any operation that changes the length of the vowel. In this case, diphthongization was also put into this level since there is no actual data on the non-Sarikoli languages that differentiates diphthongs from vowel + semi-vowel sequences and because of the particular theoretical understanding of diphthongs taken here and reinforced by the results of Chapter 3. That is, shortening, lengthening and diphthongization are all seen as related processes (see Chapter 2 for a discussion of the relationship between long vowel, short vowels, and diphthongs). Also, the data itself seems to indicate that it is possible to put all of these into one category as there is rarely a language in the data that both lengthens and diphthongizes a vowel. Finally, the third rule is said to apply if the vowel in the Pamiri mountain language is either raised or lowered in relation to the vowel in the Avestan language. Once again, since a language cannot both raise and lower a vowel, it is possible to group both rules into one. In any event, the first, gross sorting of the trends of the languages is possible through such a set of rules. The outcome of the application of these rules is provided in Table 5.3. If none of the rules applied, i.e. the Avestan and the Pamiri language form were identical, then a '0' is indicated.

In the case of 'white', both Roshani and Bartangi raise the [e] of Avestan to [i], but no other changes are made. The Shughnan vowel is identical, and the Sarikoli vowel is lowered, but only slightly. In fact, the change is so slight that it is suspicious

(especially in an analysis that is temporarily leveling differences), but the difference is there nonetheless, so it is recorded. Nevertheless, it is clear that Roshani and Bartangi are following one path of development while Sarikoli and Shughnan seem to be following another - at least in relation to this one word.

Table 5.3 Results of the application of rule analysis

Sarikoli	Shughnan	Roshani	Bartangi	Avestan	
3	0	1	1	[spaeːta]	'white'
1,2	1	1	1	[tʃitt]	'what'
1,2	1	2	2	[dva]	'two'
1,2	2	2	2	[pajaːh]	'yoghurt'
1,2	1,2	2	2	[nava]	'nine'
1,2	1,2	2	2	[tava]	'you (sing.)'
2	2	2	2	[na]	'no, not'
1,2	1,2	1,2	1,2	[gaoʃa]	'ear'
1	1,2	1,2	1,2	[aeːva]	'one'
1,3	1,3	1,3	1,3	[wəχuni]	'blood'

In the case of 'what', Sarikoli both lowers and lengthens the vowel while the other three languages only lower the vowel.

'Two' seems to show only signs of diphthongization in both Bartangi and Roshani. In fact, this is likely because of the influence of the [v] in the original Avestan.

181

Although the [v] occurred before the vowel, it still impacts the process as a whole (Sokolova 1967). A similar process can be found in Sarikoli, but in this case there is both a semi-vowel and lengthening in addition to raising. Only the Shughnan seems to have not been influence by the Avestan [v] and contains simply a high short vowel. In this case, Sarikoli seems to occupy a place intermediate between Shughnan and the other two languages since it does raise the vowel in the same way that Shughnan does but goes one step further and also lengthens the vowel and retains a [w] as a reflex of the Avestan [v].

When examining 'yoghurt', it can be seen that Sarikoli raises the vowel and diphthongizes it while the other languages only lengthen the vowel.

'Nine' also shows a similar process happening in the two languages Roshani and Bartangi while Sarikoli and Shughnan stand slightly apart. Roshani and Bartangi simply lengthen the vowel and change the Avestan [v] to a [w]. Shughnan raises the vowel and lengthens it while Sarikol both raises the vowel and diphthongizes it.

The word 'you (singular)' also shows a similar pattern in which Bartangi and Roshani are similar in their path of development while Sarikol and Shughnan also show some similarities. Here, Bartangi and Roshani lengthen the vowel [a] likely in compensation for the loss of the Avestan [v]. Shughnan also lengthens the vowel but also raises it whereas Sarikoli both lengthens the vowel and raises it but also has a [w] as a reflexive of the Avestan [v].

'No' shows a path of development in which Bartangi and Roshani have both lengthened the vowel from Avestan. Shughnan has not only lengthened the vowel but also has a semi-vowel. Sarikoli has completely diphthongized the vowel.

'Ear' presents a particularly difficult word to analyze in part because the exact nature of the original sequence [ao] in [ɡaoʃa] is not made clear in the data. It is not clear if the two vowels are a diphthong in one syllable or a sequence of vowels in two syllables. It is clear, however, that the sequence has had some bearing on the developmental patterns of the languages under study here. Sarikoli, Shughnan, and Roshani all seem to retain a reflexive indicator of the sound while Bartangi has lost the sound altogether. In addition, all the of the Pamiri mountain languages contain long vowels in this position, which also seem to be due to the original vowel sequence (whether it was in one syllable or two). All of the languages have raised the original vowel but seem to have taken a slightly different approach as to how to handle the rest of the sequence.

In the word 'one', Sarikoli raises the vowel but retains the lengthened vowel from the original Avestan word. Also, it retains the [w] from the Avestan [v]. The other languages have raised the vowel to the same height as Sarikoli, but they have also shortened the vowel. Though this change seems to be in a state of flux - as indicated by the optional segments in parenthesis. This seems to be case in which the forms exhibit an unusually large amount of variation.

The vowel under consideration in 'blood' is the second one and not the first. This means that Shughnan, Roshani, and Bartangi are all similar in that they have raised and fronted the vowel to an [i]. Sarikoli is nearly identical in its pattern of development in that it has also raised and fronted the vowel, but in this case only as far as the [e].

Keeping in mind that no actual language situation would be as clear as that in Table 5.2, it can still be argued that there are some patterns beginning to emerge in the implication chart. Firstly, it is clear that Roshani and Bartangi have more in common with each other than they do with either Sarikoli or Shughnan. In fact, in all of the words under consideration here, Roshani and Bartangi are undergoing a similar if not identical path of development. Overall, it can also be seen that these two languages have fewer changes from Avestan than do either Sarikoli or Shughnan. Where Sarikoli and Shughnan often have two numbers indicating two changes in their columns, Roshani and Bartangi have far fewer. A further examination of the implication matrix in Table 5.3 reveals that Sarikoli seems to be at the opposite end of the spectrum. In 7 of the 10 cases examined here, the vowel segment in Sarikoli has undergone two changes in comparison with the Avestan. (This counts the change in 'no' as two changes.) The next closest language in terms of actual number of changes in Shughnan. This indicates that Sarikoli has more in common with Shughnan in terms of development than it does with the other two languages. It also indicates that Shughnan is intermediate between the two, leaving the overall relationship pattern with Bartangi on end of the extreme and Sarikoli on the other end (Figure 5.1).

Sarikoli   Shughnan   Bartangi   Roshani
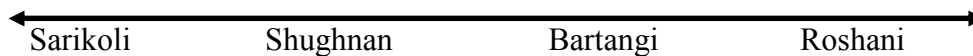
Figure 5.1 Relation on a continuum between the Pamiri mountain languages

Having indicated that Sarikoli seems to have undergone greater change than the other Pamiri mountain languages in this study, some further exploration of individual words is still necessary. In the case where Avestan already contained a long vowel, Sarikoli has preserved that long vowel. In this data set, that has occurred in two cases in 'white' and 'one'. This is important because inclusion of these two words in the data have obscured an otherwise significant trend. In addition, in one word 'blood' all the languages contain a simple long vowel undergoing a raising process. If these three cases are eliminated and the remaining cases are examined, a clear pattern emerges.

In Table 5.4, it can be seen that where Sarikoli contains a simple long vowel, that the other three languages all contain a short vowel. On the other hand, where Sarikoli contains a long vowel followed by a glide, the other three languages contain either a short vowel and a glide or a long vowel. In the case of 'two', the occurrence of the glide can be accounted for by the presence of [v] in the original Avestan influence from a surrounding segment. Where Sarikoli contains fully developed diphthongs, the other languages contain long vowels or long vowels plus glides. These data strongly suggest a pattern whereby Pamiri mountain languages develop diphthongs over time, passing through a process of simple monophthong → simple monophthong + glide → long monophthong → long monophthong + glide → diphthong. If this is the case, then

185

Sarikoli would be considered the most 'advanced' or 'progressive' of the languages and would be considered the origins of such a sound change assuming, as many would, that the language which demonstrates the greatest change is at the origin of the change--as articulated in Johannes Schmidt's Wavemodel (1872) and Hugo Schuchardt (1868). If a diphthong exists, then it indicates that a certain path of development has been followed.

Table 5.4 Patterns of vowels in Pamiri mountain languages

Bartangi	Roshani	Shughnan	Sarikoli	Avestan	English
S	S	S	L	[tʃitt]	'what'
S+G	S+G	S+G	L+G	[dva]	'two'
L	L	L	L+G	[tava]	'you (sing.)'
L	L+G	L	L+G	[gaoʃa]	'ear'
L+G	L+G	L+G	D	[pajaːh]	'yoghurt'
L+G	L+G	L+G	D	[nava]	'nine'
L	L	L+G	D	[na]	'no, not'

S = Short Vowel   G = Glide   L = Long Vowel   D = Diphthong

This pattern in the data combined with the fact that Sarikoli contains diphthongs where Avestan did but the other languages do not indicate that there is likely some other kind of abnatural change here as well. Sarikoli, for some reason not able to be identified here, has a tendency to either preserve or create diphthongs. It is not clear why this might be the case, but it is well known that connatural and abnatural changes sometimes

weave themselves together in a complex web that is hard to untangle. Although abnatural changes are just as real as the connatural language change, their explanation is fundamentally different. Speakers of a language may decide to make a change for social or prestige reasons and this change may run contrary to the unmarking principle outlined above. On the other hand, a disturbance of the natural pattern leads to the prediction that a connatural language will eventually either fix or delete the abnatural change to come into compliance with the system as a whole and linguistic universals (here Greenbergian type implication hierarchies).

The potential masking effects of change can be made clearer by providing an example of each. For example a higher order connatural change may make it difficult to decipher a lower order connatural change, but the change is still there. For example, Bailey (1981:48) maintains that the existence of [tk] in a language implies the existence of [kt]. Often, however, the [kt] cannot be found in a given language because it has been spirantized to [x] (at a lower level). This should not be seen as an exception to unmarking but as something that simply needs to be explained. The linguist need merely explain that the first change happened and then that another change operated later to make the subsequent change, masking the first. Likewise, markedness reversal, a type of abnatural change, is an exception to unmarking and may make it difficult to decipher the system as a whole. Even though markedness reversal is the opposite of the other types of change (i.e. will not lead to a move from more to less marked), it is nevertheless a 'natural' and normal change.

Bailey (1996:197) offers an example of markedness reversal from English diphthongs. In English, the <u+i> combination has been simplified (monophthongized) and <i> has been lost altogether, as in English *bruise, cruise, fruit*, etc. In fact English spelling still writes the diphthong, but it is no longer pronounced. This change in English is an example of unremarkable and normal connatural language change. In English, however, markedness reversal has taken place in two environments that did not follow the natural unmarking process of connatural change. The first was the situation in which the <u+i> sequence had a syllable boundary intervening between them. This can be seen in English words like *gooey, Buick,* and *fluid*. The diphthong is retained as a symbol of the syllabification. If it had not been retained, the syllabic information would have been lost in the pronunciation. As such, a process of markedness reversal was introduced to return these words to their pre-monophthongization state. The second environment in which the change was not seen was in foreign loan words brought into English, such as *bruin* and *Louis*. Loan words are usually impervious to connatural change until such a time as they have been perceived to be a part of the language. Since these words have not yet been totally incorporated into English, they have resisted the change. This is abnatural change for non-linguistic reasons motivated by the perceived foreignness of the words. However, it should be noted that connatural change is expected to work on these forms as they become more incorporated into the language and if there are no other abnatural changes to interrupt the change. This, indeed, can be seen in some varieties of Southern States English, where these sounds are now beginning to be

monophthongized by many speakers.   This example is also used to show how diphthongization (or monophthongization) can be used to point out (or marker) phonological development.

Returning to the case of the Pamiri mountain languages then, it should now be clear that the two words 'white' and 'one' are simply participating in some other overall pattern of development than the other words analyzed here. This pattern is likely due to the fact that these words initially had long vowels in the Avestan and therefore either participated in a different kind of connatural change in Sarikoli - one that simply preserves the long vowel because it existed in the original - or in some other kind of abnatural change. For example, given the overall and more generalized trend toward diphthongization, speakers may be trying to bring these two words into the overall pattern and alignment found elsewhere. In fact, this is what the theory predicts that speakers will do. Any words that are not initially part of a connatural change are predicted over time to be incorporated into the overall pattern.

Given the data here, then, is it possible to make a prediction about the future development of these languages? The answer must be yes (Bailey 1981, Bailey 1996). The data clearly show a pattern whereby the Pamiri mountain languages are tending toward diphthongization. The pattern was outlined earlier where languages in this group were argued to be moving through patterned and principled steps of diphthongization. Sarikoli seems to be most advanced into the process, which would account for the reason why so many previous linguists have debated and disagreed on where Sarikoli

stood in relation to the other languages. There is a pattern happening in which Sarikoli does contain more diphthongs than the other languages of the Pamiri mountains, but that pattern is still growing and developing, so the pattern is often not complete. In addition, the pattern has some counter-intuitive examples that initially seem to defy the overall development seen in other languages. As explained above, however, these counter examples (thus far) seem to be explainable as either changes at another level or as sociologically motivated changes. Therefore, all things being equal, these languages are expected to continue the pattern of development that Sarikoli has originated and will eventually also develop long vowels and eventually diphthongs in these environments in the same way that Sarikoli has. Sarikoli as well will continue to trend towards developing more diphthongs. Therefore, the implicational pattern will be that in Figure 5.2.

Sarikoli ⊃ Shughnan ⊃ Roshani ⊃ Bartangi

Figure 5.2 Implicational hierarchy of change among Pamiri mountain languages

The implicational hierarchy in Figure 5.2 indicates that a change that is seen in Sarikoli is the most progressive. If a lect has a particular feature, it will also have the features to the right on the implicational hierarchy but will not have those to the left. Here, it means that it if a language on the left has a diphthong, then it is predicted that the language to its right will also contain a diphthong. Therefore, it is not predicted by

this hierarchy that Roshani would have a diphthong (in a particular word) but Shughnan would not. To offer another example, if Shughnan only has a simple vowel, then Roshani and Bartangi should also have simple vowels. That is to say, that a lect on the right hand side of the hierarchy is expected to be at an earlier stage of the march towards diphthongization than a language on the left of the hierarchy.

At this point, there are several caveats that ought to be offered. Firstly, the results here only account for a small sub-set of the languages and only account for one small part of the phonetics/phonology and are based on a very small data sample. Therefore, saying that Sarikoli is the most 'advanced' down the path of this process does not in any way imply that it is also the most 'advanced' in any of the other types of development. Just as it is seems progressive in its development of diphthongs, it might just as easily be very conservative in other parts of the grammar. Secondly, it is difficult to imagine that this process will continue in the *ceterus paribus* (all things being equal) fashion assumed by the theory. Sarikol Tajiks are now by and large cut off from the other Pamiri mountain languages and now likely to begin to follow a different path of development. In addition, the Uighur language exerts a heavy influence -- not to mention the influence of Chinese and Wakhan Tajik upon Sarikoli. None of these three languages (Uighur, Wakhan, and Chinese) influences the other languages in the Pamiri mountain group to any great degree. Influence from Uighur is especially common among Sarikoli speakers and many Sarikolis seem to feel that it is a prestige language (Ibrahim 2006). This is a fact that can also be verified from personal observation. On

191

the other hand, this situation has already been in effect for 70+ years and so far Sarikoli does not seem to have departed from the path of development begun in the group as a whole. The end result remains to be seen.

The best way to gain access to the path of future development in Sarikoli is to examine how the individual dialects of the language are currently developing. This is done in the next section.

5.3 AN ANALYSIS OF SARIKOLI DIALECTS

Analysis of data in Chapter 4 revealed that there appear to be (at least) three dialects of Sarikol Tajik clustered in Tashkorgan, Vacha, and Burungsal. The Tashkorgan lect contains generally higher and more front vowels than Vacha. In turn, the vowels in Vacha are higher and more front than those in Burungsal. This reveals a geographic trend moving from east to west in which vowels become gradually lower and more back the farther one moves to the east.

In Section 5.2., analysis showed that the trend operating across the various Pamiri mountain languages is a trend towards diphthongization in these languages. Given this situation, a look at the various dialects of Sarikoli must necessarily examine which of the lects has moved most towards having a stable diphthong in these words. By a stable diphthong, it seems reasonable to look for an iconic diphthong, i.e. a diphthong that is common both cross-linguistically and within this language family. Likewise, it seems reasonable to take into account the data examined in Chapter 3 in which it was revealed that Sarikoli currently contains only three diphthongs overall.

Although, it is possible that new diphthongs will arise in the language, it seems more reasonable to assume that changes in the individual dialects are currently changes towards one of the diphthongs already in existence.

For example, the Sarikoli word [nai] 'no, not' has already been revealed to contain a diphthong, while the other languages in the language family do not yet contain a diphthong, but show a trend in that direction. Of the four languages analyzed in this dissertation, Shughnan appears to be the farthest along in the trend and has already developed a long vowel followed by a semi-vowel while the other two languages, Bartangi and Roshani contain merely simple long vowels. The variation among the languages in the family show that there is a trend towards diphthongization. However, just as each of the four languages in the family are at various stages in the process of diphthongization, the three dialects of Sarikoli are also at various stages of diphthongal development.

Since the overall trend of the historical data shows a trend towards diphthongs in the development of the language family, it is possible to argue that individual dialects are also following the same trend. In this case, it would be the Burungsal dialect that would be considered to be farthest along in achieving this 'goal' since the diphthong in that language has gone farthest towards achieving a target vowel position. By containing the lowest and farthest back vowel of the three lects analyzed here, its vowel comes closest to the iconic /ai/ diphthong position. Therefore, it will be seen as being

the origin of the change and being farthest along in the change from a simple vowel to a diphthong.

Although, it is not necessary for sound changes to follow geographic patterns, it does help to bolster the notion of a trend when they do (Bailey 1996, Chambers 2001). In many cases, dialect change is seen as being cumulative, in which 'the further we get from our starting point, the larger the differences will become' (Chambers 2001:5). Such changes reflect a situation in which adjacent lects are mutually intelligible but the larger the separation in geographical space, the more difficult it is for speakers of any two given lects to communicate. This seems to be reflected anecdotally in speakers from the Pamiri plateau. It is not uncommon for speakers from China to be able to communicate from village to village but also with people from one of the other countries in which Pamiri mountain speakers reside. However, the greater the geographical distance of the other speaker from one's home town, the more difficult this communication becomes. This notion of geographic distance displays itself in the data analyzed here. As lects become more eastern geographically, the change becomes more progressive. This change is not only seen in individual dialects of Sarikoli but also across the other languages of the family as a whole. This helps to underscore the apparent arbitrariness of language division that is often prevalent in dialect studies. In fact, it is difficult to separate mountain Tajik languages on purely (non socio-) linguistic grounds and a great deal of mutual intelligibility is reported among them.

This trend among the Pamiri mountain lects toward diphthongization can be seen in other words as well, but a few examples should suffice to establish the overall trend. For instance, this trend can be seen in the Sarikoli word 'what'. This word contains a long vowel which exhibits change across the three lects of Sarikoli. Following the general pattern, the Burungsal dialect contains a vowel relatively lower and farther back than the vowel in the other dialects. If this information is placed in light of the overall trend identified across all Pamiri mountain languages, the pattern of change is clearer. In this case, Roshani, Bartangi, and Shughnan all contain simple vowels while Sarikoli is unique in containing a long vowel. The overall trend of the data leads to the prediction that each of these languages will move along progressively towards overall diphthongization. This being the case, then the Burungsal dialect of Sarikoli is the most advanced in the movement in that direction, already showing signs of lowering the vowel in preparation for a more iconic diphthong.

Similarly, 'nine' shows a trend towards diphthongization. In this case, however, Sarikoli has moved farther down this path of development than it has in 'what'. In 'nine', Sarikoli already contains a diphthong, while Shughnan contains a long vowel + glide, and Bartangi and Roshani contain long vowels. All things being equal, this path of development is expected to continue. This being the case, then Burungsal is the most advanced along this path of development.

Such an analysis by necessity makes underlying assumptions that ought to be made explicit here. The first is that, all things being equal, the languages here will not

only form diphthongs, but are also trending towards more ideal or prototypical diphthongs. A prototypical diphthong is one in which the sonority differences between the onset and offset positions are maximal (Miret 1998). This indicates that diphthongs composed of all mid-central vowels are disfavored. Diphthongs composed entirely of middle vowels are rare cross-linguistically and languages that contain them tend to increase the sonority of one or the other part of the diphthong in actual realization. Such a trend can be seen in an example from Spanish offered by Miret. In Spanish *aire* 'air' originally contained a sequence of [a.e], in which the period indicates a syllable boundary. Because of the syllable boundary, there was originally no violation of the sonority hierarchy by having two vowels of such similar sonority contiguous in the word. Later, the same segment became [ae] (i.e. the syllable boundary was lost) and eventually became [ai]. The last change is seen to be motivated by a change in status that occurred when both vowels were moved to the same syllable. When the change to [ae] occurred and a diphthong was formed, a conflict began in which it became difficult to identify the nucleus of the syllable due to the similar sonority of the two contiguous segments. To rectify the violation of sonority and to create a clear peak at the nucleus of the syllable, the second vowel was raised and fronted to [i], lowering its sonority and also allowing the [a] to be clearly identified as the nucleus of the syllable.

Trends like those in Spanish and the Sonority Hierarchy Principle make it reasonable to predict that Sarikoli long vowels that begin their path towards diphthongization will also move to either raise or lower one of the contiguous elements.

Once the two elements begin to be interpreted as occurring in the same syllable, then a conflict occurs if both of the elements are mid-central or in any other way too near to each other on the sonority hierarchy. In this case, Sarikoli is leading the change towards diphthongs and the Burungsal dialect specifically is spearheading that change. The change is being smoothed and made to fit in with general linguistic principles by also beginning a process of lowering the first segment of the emerging diphthong in an attempt to increase the difference in the overall sonority difference between the onset and offset components. Cross-linguistically, so-called 'falling' diphthongs in which the first vowel is sonorous followed by a less sonorous segment are the prototypical norm. In fact, they are so common that traditionally only these diphthongs were considered to be diphthongs at all (Miret 1998). All of which points to the high likelihood that this is the reason that the Burungsal dialect has a tendency to have lower and more back vowels than other Sarikoli dialects - it is a trend to create a more prototypical diphthong situation in preparation for the actual appearance of the diphthong.

The second assumption which needs to be made clear here is that this process of sound change is not seen as being teleologically motivated. That is to say, there is no concept here of purposefulness behind any of these changes either on the part of the individual speakers or on the part of the language itself. In the past few years there has been revived interest on the part of linguists in the biological metaphor in language change, which was at least partially motivated by the popularity that variation and change has received in recent linguistic work. This interest in change has led some

197

linguists to explore the evolutionary metaphor in a great deal of detail (McMahon 1999:Chapter 12). This exploration of linguistic evolution has in turn led some linguists to argue that change can be teleological, but that is not the position taken there. It is true that the words employed in this dissertation do often come from a biological metaphor, words like 'development', for example. The desire here is for ease of communication rather than any stated or implied acceptance of the use of the biological metaphor (i.e. the evolutionary metaphor) in linguistic endeavors. In fact, the position taken here is that such a discussion is moot (cf. Ohala 2005). The prediction for the future made here is based strictly on data and seeks to identify trends based on linguistic drift. As stated earlier, in the frame of reference adopted here, the motivation for connatural change is not an issue. It is known that languages drift and change over time, and the goal here is to try to identify that change throughout history - not to answer the question about why languages change.

In fact, the idea that languages change but that such change need not be motivated by any particular factor (sociolinguistic or otherwise) seems to be gaining momentum in both linguistics and biology both. It has been said that 'perceived directionality is accepted in current evolutionary theory as resulting from random variation and natural selection, which combine to produce order with no external direction' (McMahon 1999:337). It seems that languages are 'ebbing and flowing like the tide, but neither progressing nor decaying…' (Aitchison 2004:253). This change is epitomized by that kind of change that has been called here connatural language change.

In fact, it was made quite clear from the beginning that the developmental approach to language accepts that change is non-teleologic in many circumstances claiming, 'It would be a waste of time to reiterate that developmentalists do not attribute to speakers a knowledge of the past history of their language' (Bailey 1996:31). In fact, it has likely been the overwhelming popularity of Labov that has helped influence the notion that variation and change need (nearly) always be sociolinguistically motivated. Labov, himself, is certainly concerned primarily with such changes, but such sociologically motivated sound changes need not be considered the only reason for sound change or even the primary reason for sound change when examined in light of all of history. Actually, in his seminal work *Principles of Linguistic Change* (2001b) Labov dedicates one volume to the examination of internal changes and one to the examination of external changes. Certainly, historical linguists have not occupied themselves with why such changes occurred but merely documented that they did. This research takes a similar approach to the sound change here. The change is not considered to be motivated either by any desire to reach a particular goal in the language or by any particular sociological motivation.

Thus, the pattern of sound change in the case of the Pamiri mountain languages can be seen as a bundle of features that have as their point of origin the Burungsal dialect of Sarikoli Tajik. This relationship is summarized in Figure 5.3. In this case, the Sarikoli language has been replaced by the names of the three dialects. The breaks between the three lects are artificial since there is a smooth and continuous change

among the languages that flows from the source: the Burungsal dialect of Sarikol Tajik. It is predicted here that if there were more data available on the specific dialects of individual Shughnan, Bartangi, and Roshani communities that the change could be tracked through the individual villages, creating a dialect continuum that can be tracked geographically from east to west.

Burungsal   Vacha   Tashkorgan   Shughnan   Bartangi   Roshani
        (Dialects of Sarikoli)

Figure 5.3 Relation on a continuum between the Pamiri mountain languages

5.4 CONCLUSION

This chapter has examined the developmental path of four Pamiri mountain languages. They have been examined in terms of their changes in vowels and diphthongs over time, using Avestan, the ancient predecessor to these languages as a basis of comparison. Data organized into implicational hierarchies reveal that there is a change happening among the languages in which simple vowels develop over time into long vowels and eventually into diphthongs. Sarikoli seems to be the origin of this change and specifically the Burungsal dialect of Sarikoli seems to be the most advanced in this change (and therefore the source of the change).

At least in regards to diphthongs, this seems to undermine the generally held opinion that Sarikoli is the most conservative of the languages in this language group.

In fact, quite the opposite seems to be the case. Sarikoli is leading the way in the change to diphthongs in the language. Certainly, it could be true that when such results are placed in an overall framework that takes into account all facets of the language that Sarikoli will truly be considered the language that is least likely to change. While such an assessment is the object of a different study for a different day, the changes and results identified here should certainly be taken into consideration and generally held pre-conceptions should be re-examined in light of more concrete evidence and a more rigorous and systematic approach to the study of the data. Generally speaking, when linguists have made such broad-reaching claims about Sarikoli being conservative, no proper documentation has followed. In light of the evidence revealed here, such statements can no longer be accepted without concrete arguments to bolster such claims. Chapter 6 will examine the theoretical implications of the results of this analysis.

CHAPTER 6

THEORETICAL IMPLICATIONS

Even for those not particularly interested in Sarikol Tajik or in Indo-Iranian languages, this dissertation still offers interesting and potentially useful implications for how languages change over time, how diphthongs manifest themselves phonetically, and how variation can be used to enhance linguistic theory. This chapter begins with an examination of the Pamiri mountain languages, which should provide insight into that language family. For those less interested in Pamiri mountain languages than in theoretical implications, Section 6.2 should be of particular interest.

6.1 PAMIRI MOUNTAIN LANGUAGES

The exact relationship among Pamiri mountain languages is unclear. For many of these languages, data is scarce and often outdated. Furthermore, due to its relative geographic and political isolation, even less is known about Sarikoli Tajik than the other languages in the group. The most exhaustive studies made to date were made either before the computer revolution or were made from a very traditional perspective. In this light, the data offered and analyzed here is a large step forward in general knowledge and understanding of Sarikoli Tajik and its dialects.

Sarikoli dialects have been a controversial issue for a long time. Gao (1985) claims that Sarikoli dialects differ on a north to south axis while Pakhalina (1971)

claims that their differences lie primarily in an east - west direction. Gao gives no further reasoning to his claim, but merely states what he has concluded. Pakhalina, on the other hand, claims that the primary difference between the dialects seems to be in the pronunciation of vocalic segments. This study has proven Pakhalina's theory to be right in that it has confirmed that the dialect variation is primarily east to west and that the languages do differ primarily in vocalic pronunciation.

Similarly, there has been a long-running controversy among those few who have studied Sarikoli as to why it is so different than the other languages in the Pamiri mountain language family. For example, it was argued that Sarikoli could potentially have up to 12 diphthongs, which was unusual in the world and certainly unusual within the family. Nevertheless, the argument persisted since no one previously had the ability to do a spectrographic analysis of the data and definitively solve the problem. This issue has now been resolved, and Sarikoli appears to be much more 'normal' than previously supposed. In fact, Sarikoli does not contain so many diphthongs and the few that it does contain are generally accepted and common diphthongs in the languages of the world.

Likewise, there was great controversy surrounding why Sarikoli was the only language in the language group to not contain a short vowel/long vowel distinction. This was also argued to be a way in which Sarikoli was unusual in its language family. All the other Pamiri mountain languages contain a phonemic short and long vowel distinction, but no such distinction was previously found or even suspected in Sarikoli. This assumption has now been challenged and proven to be false. Sarikoli does indeed

have the same short and long vowel distinction as the other languages in the family, but many of the long vowels are leading the way towards diphthongizing those long vowels. When analyzed spectrographically, it became clear that Sarikoli long vowels show considerable drift in formants (but do not show prototypical diphthong movements), which is why they were previously thought to be diphthongs and not long vowels. Once again, Sarikoli turned out to be more 'normal' than had previously been supposed. Thus, this study has been a major unmasking of thinking that placed Sarikoli as an unusual language and has instead placed it squarely within the parameters of the other languages in this language family.
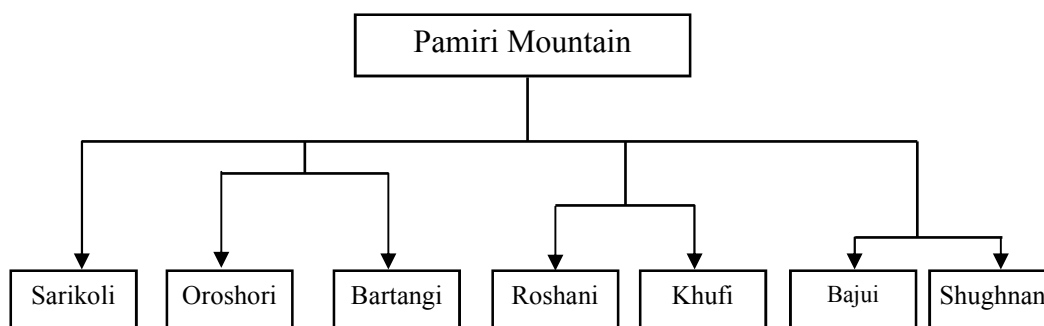
Figure 6.1 Sokolova's representation of the relationship between
the Pamiri mountain languages

Neither Pakhalina nor anyone else, however, has ever before tried to incorporate knowledge of Sarikoli into an overall dialect continuum with the other Pamiri mountain languages. In that regard, this study is unique. In fact, the Soviet scholars seemed to

accept the established languages borders and boundaries *a priori* and never challenged the notion of a discrete separation of the languages, never arguing that actually one dialect transitioned smoothly into another. In Figure 1.2, a representation was given in which Sarikoli is considered to be a dialect or kind of Roshani rather than a separate language. However, in Figure 6.1, Roshani is considered a separate language relatively distantly related to Sarikoli. The point here is that no systematic examination has to date been made on the exact nature of the relationship between these lects. The work on diphthongs here does seem to indicate that they form a dialect continuum and that one lect does gradually and systematically turn into another. Although making such a claim based on only the study of diphthongs would also be overstating the case, the study conducted here certainly leads to that initial conclusion and will serve as a first hypothesis for future studies.

There have also been several hypotheses regarding the genetic or historical relationship among the Pamiri mountain languages. One such representation was given in Figure 1.2. In that representation, Sarikoli was listed as being intermediate between Shughnan and Roshani. Payne says that the languages 'form a closely related subgroup,' but no basis for such a grouping is offered and it appears likely that he is basing his analysis on Sokolova's work (1980:153). In her analysis, Sokolova (1967) portrays Sarikoli as having more in common with Bartangi and Roshani than Shughnan (Figure 6.1). It is true that such an analysis was made based on many decisions made in many parts of the phonology, lexicon, and syntax of the languages. However, this initial

analysis of the diphthongs seems to indicate a different relationship - at least in terms of one part of the overall grammar. Instead of Sarikoli being intermediate between the other languages in its development of diphthongs, Sarikoli is on the edge leading the development. Therefore, this analysis provides new insight into the overall relationships between these languages and warrants a new analysis of overall genetic relationships. It is possible that the previous analysis will stand, for Sarikoli does seem to have some other indications that it is unusual within the language family, such as lack of gender distinctions in the noun system that is present in the other languages and possessing an absolute-oblique case marking distinction for nouns that is not present in the other languages (Payne 1980). Nevertheless, a re-examination of the relationship between Sarikoli and the other languages in the Pamiri mountain language family is warranted in light of the conclusions reached here.

6.2 APPROACHES TO DIPHTHONGS

Definitions of diphthongs are numerous and often contradictory. For example, a brief review of some of the more scholarly definitions of diphthongs on the internet offers descriptions as diverse as the following:

- 'A diphthong is a phonetic sequence, consisting of a vowel and a glide, that is interpreted as a single vowel' (Loos 1997).

- 'Diphthongs are those sounds that consist of a movement or glide from one vowel to another' (Luscombe 1996).

- 'A *diphthong* is a speech sound in which the articulatory mechanism moves continuously from an initial vowel position to a final vowel position' (Kaiser 1997).

These definitions serve to highlight some of the major controversies surrounding the understanding of diphthongs in linguistics. For example, must a diphthong be interpreted as a single vowel or not? While this is key to the first definition, the others make no reference to it. Likewise, the first definition mentions a vowel and a glide but does not mention another vowel. Does this definition then imply that the diphthong need not reach its second target segment at all? On the other hand, the second two definitions seem much more open in what would be an acceptable diphthong. In those two instances, mere movement from one position to another is important. There is no mention about interpretation or syllable boundaries in those two definitions, however. Is it mere oversight, something they left out of their definition, due to their assumption that diphthongs must be in one syllable or not?

Thus, despite numerous definitions and attempts at carefully defining diphthongs (cf. Chapter 2), the definition of a diphthong remains as elusive as ever. This is partially because many linguists carry with them underlying assumptions about diphthongs that they do not include in their definitions. For example, most linguists seem to assume that diphthongs are segments that occur in one syllable, but not all take the time to explicitly state this. Another reason for the lack of a cohesive definition of diphthongs is in the data itself. Different research taking different approaches has

yielded different results. At the end of the day, linguists are left debating what a diphthong is.

This dissertation makes three noticeable contributions to the unraveling of this conflict concerning diphthongs and upon linguistic theory in general. The first is the notion that variation is normal and should be accepted in language studies. The second is the inclusion of time into the study of the phenomenon. The third is the notion that it is possible to understand diphthongs as a fluid concept, recognizing that there are more and less prototypical diphthongs rather than adopting one single definition. Each of these is examined in turn in the following sections.

6.2.1 VARIATION THEORY

This study underlines the importance of allowing variation in data. Perhaps this is not unique or surprising given that dialectology has never forsaken the need to allow variation in data and this study is in many ways a classic study of dialects. However, in the modern paradigm of linguistics overall, this is significant and worth mentioning. While the predominant school of thought focuses on the grammar of the ideal speaker-hearer in a homogeneous community, variation theorists have continued to argue that such a situation is too much of an abstraction from reality. While many researchers continue the classic distinction between speech and language or competence and performance, others argue that such a distinction simply fails to account for the data. Chambers is typical in arguing 'The most elusive questions about language variation can only be answered by confronting linguistic variability in all its profusion - as

ordinary human beings confront it daily as they go about their mundane tasks' (2001:148). And, accounting for the data spoken by ordinary human beings should be the first and basic requirement of any theory of language.

This study is a study of language in change, which is one kind of variation (Chambers 2004a). In this study, it was shown that the various lects of Pamiri mountain languages are in a state of change. It is precisely this change which is interesting and telling. It is the change itself which accounts for the data and also allows a prediction to be made about the future of the language. If the change is ignored and one particular lect is chosen as the 'competence' lect (as opposed to 'performance') or the 'language' lect (as opposed to 'speech'), then it is impossible to decide which lect to choose as the proper object of study. It is, in fact, very likely that the linguist confronted with such a choice among Sarikoli lects would choose to study the Tashkorgan lect. This is the lect that has the greatest potential to have any social prestige at all, but then again, the choice is being made for what the generative school would call 'non-linguistic' reasons. It becomes clear that making such a choice puts the linguist in a position of ignoring some data in favor of other data, but such a choice need not be thrust upon the linguist, indeed must not be thrust upon the linguist if linguistics is to maintain any modicum of being a science.

It appears that Chomsky and Halle were aware of the need to provide optionality in phonology rules and provided two mechanisms with which to do so: optional rules and allomorphy (Anttila 2004). These optional rules have recently been

given a great deal of attention within the generative tradition itself and have given rise to the development of Optimality Theory, first proposed in the early 1990s. Since its first introduction, the theory has become ever more popular. However, Optimality Theory by and large only addresses internal factors of the grammar such as phonology and syntax. It does not address variation that might be conditioned by external factors such as gender, age, socio-economic class. It certainly does not address the kinds of variation studied here - the kind that occurs in language drift. Its continuing lack of ability to address such issues means that it still fails to address many issues that make linguistics both interesting and revealing. In fact, it still continues to focus on only 'language/competence' and continues to place 'speech/performance' outside of the realm of investigation.

Recent work in phonetics has realized the need to allow for variation in the theory to reflect the variation in the data in reality. Rather than continuing to pursue the invariant definitions of phonetic phenomenon, the focus of phonetics is now shifting:

The lack of invariant physical correlates of speech features and segments has questioned the premise that inspired them, i.e., that mental representations are invariant, categorical and abstract. Thus in the last ten years the focus of the phonetics-phonology interface has moved away from discrete and invariant internal representations, stripped off the phonetic details and has turned to prototype, e.g., usage-based and

experience-based probabilistic models, which assume that variation is

part of the internal representation. (Sole 2003:1)

Switching to a model which allows for variation through prototype theory (cf. Section 6.2.3 below) allows the model to truly account for the data, i.e. to account for language in use. It is more than a debate about what should be the proper course of study in linguistics. If linguistics continues to follow formal models that seek to establish all things as invariant models and gradient possibilities and variation are ignored, then it become impossible to explain or analyze key linguistic problems - those like the one analyzed here. Indeed quantitative approaches to speech have been argued to be better able to explain phenomenon in first and second language acquisition, production and perception problems, the neutralization of phoneme classes, and allophonic splits as well as sound changes - the particular use such an approach is put to in this study (Sole 2003). If data is only analyzed in such a way that variability is not taken seriously or investigated, then by definition no sound changes will ever be studied in linguistics. Thus, variability is a key cornerstone in the study of dialects and historical change.

If diphthongs in Sarikoli (and perhaps in any language) are to be truly understood, then they will need to be understood in their real world variation rather than as an invariant model. Actually, many previous linguists who have studied the phenomenon have understood that diphthongs vary across dialects, across speech registers, and among individual speakers at different times (cf. Chapter 2). It is because linguists have approached the task of identifying diphthongs with the attitude that

diphthongs may never vary that such controversy has arisen. This study underlines the elegant simplicity of allowing the variation to exist and then accounting for the data. As this study shows, it is possible to elegantly account for the data, provide rational explanations, and predictions for the future while also allowing variation to be a part of legitimate linguistic inquiry.

6.2.2 TIME THEORY

If diphthongs are only studied synchronically, they will never be fully understood. This is because they rise and fall over time, form and reform, appear and disappear (cf. Chapter 2). Thus, if a particular language is only studied at one point in time, then it is highly likely that some diphthongs will be at one stage of development while other diphthongs are at another stage of development. That is to say, it is reasonable, given the relationships previously discussed between diphthongs and monophthongs (both short and long) to find many varieties of diphthongs in a language at any given moment in time. Some will be farther along the path of development towards a prototypical diphthong while others may still be more monophthong-like. This is certainly the case in the Sarikoli data examined here.

In modern linguistics, diachronic studies have fallen into disfavor. Following the Chomskyan revolution, linguists devoted themselves by and large to synchronic studies of individual speakers. The notion of time and change was relegated to the back-burner of historical linguistics. Indeed, historical linguistics carries with it the notion that time is key and that languages are always changing. Even a basic introduction to

historical linguistics makes this clear: 'A basic assumption in historical linguistics is that languages are constantly changing. Rather than assuming that languages are static, non-changing 'things,' we need to think about them as one of the most dynamic areas of culture' (Effland 1995). While the notions of time, change, and variation became relegated to the realm of historical linguistics, core areas of linguistics like syntax, phonetics, and phonology conducted within the Chomskyan framework no longer accepted time explanations as legitimate explanations. This essentially removed the historical perspective as a legitimate path of exploration and explanation from the realm of theoretically acceptable (Janda 2005).

However, time is what can explain the apparent differences in the different lects of Pamiri mountain languages. In Sarikoli Tajik and in the Pamiri mountain languages in general, some lects contain monophthongs while others contain diphthongs. In addition, Sarikoli, being the most advanced in terms of the rise of diphthongs in the language has many long monophthongs that sound deceptively like diphthongs when only examined aurally. As was shown in this study, some of the segments identified as diphthongs sounded like a diphthong when listened to, but showed distinct patterns that were separate from the sharp and rapid movement in formants expected of diphthongs. Thus, it is possible to debate from different perspectives that a particular segment is or is not a diphthong. When examined aurally, one might conclude that Sarikoli contains up to twelve diphthongs. However, when examined spectrographically, it becomes clear that such segments are not at all the same.

To account for this difference in any kind of logical way is impossible if the element of time is removed from the analysis. In his analysis of diphthongization, Anderson pointed out more than thirty years ago that '[a]n essential fact about evolutive [internally motivated change] is its gradual character' (1972:12). Therefore, if the data are only examined synchronically, one can argue that those who posit that Sarikoli has twelve diphthongs are wrong, but such an explanation begs the question. How can one explain that there are segments in the language that are almost a diphthong, but not quite if reference is only made to one state of the language at one point in time? In short, no explanation is possible without examining the data over time. Anderson further argues that sound change in general and the rise of diphthongs in particular can be manifest across time in three ways: (1) Speakers of different ages may show the change; (2) individual speaker register variation; and (3) speakers of the same age from adjacent communities may show different stages of the change (1972). The third situation he describes is exactly the situation found in the case of Sarikoli. Thus, time tracked across speech communities is a key facet in understanding the phenomenon of diphthongization.

When data is only examined synchronically, linguists are doomed to a repeating cycle of debating on and off glides and whether or not formant movement is actually rapid enough to be considered a diphthong or not. However, if time is introduced back into the equation, then one can point out that diphthongs and monophthongs come and go over time and so variation in the data is to be expected. Languages are always

214

changing, and time can go a long way towards explaining many of the apparent inconsistencies in data if time is allowed as a legitimate method of linguistic explanation.

The case of Sarikol Tajik has proven to be a good example of the importance of diachronic study. Although linguists have debated whether or not Sarikoli has diphthong for many years, no previous linguist had examined the problem from a historical viewpoint. Once the historical viewpoint was explored, it became clear why the debate had raged for so long. In fact, there is variation in the data. In fact, different speakers from different places have different pronunciations. If one remains open to the fact that variation is normal and that time is a legitimate avenue of explanation, then the debate falls by the wayside. In fact, some diphthongs are exactly what linguistic definitions would predict, but others are not. Often, the reason for this is their point in development, which means that time must be allowed as a legitimate and viable method of explanation in linguistic theory.

6.2.3 PROTOTYPE THEORY

In this study, it was shown firstly that linguists cannot agree on the definition of a diphthong. Then, it was shown in Chapter 2 that even when such a definition was either decided or agreed upon, actual realizations of that definition were difficult to find. This led most linguists who had approached the theory of diphthongs to go back to the drawing board, concluding that their definition of diphthong must have been wrong if they could not find actual instantiations of it.

Going back to the drawing board is not necessary every time a non-perfect example of a diphthong is encountered. Such action is only necessary if linguistic theory sets for itself the false dichotomy that something either is 100% a diphthong or it 100% is not. Such all or nothing approaches to linguistics are not new and became even more popular after Chomsky and Halle's *Sound Patterns of English* (1968) introduced the concept of binary features to the world of linguistics. In this model, features are seen as being either turned on or turned off; they are either present or they are not. There is no room for something to be partially one thing and partially another. A segment must either fully possess the quality of the feature or fully not possess it. This model has influenced the thinking of all linguistics and also influenced the way linguists write definitions. Therefore, it has been argued that a segment is either a diphthong or it is not there is no room to be more or less 'diphthong-like' in this way of thinking.

There is, however, an alternative in prototype theory. This theory builds on the work of Eleanor Rosch (1976) and later that of George Lakoff (1987). Taylor (1989)and Langacker (1987) both argued that prototype theory can be successfully applied to phonetic/phonemic questions. In fact, there have been other phonological proposals that also argue that phonemes ought to be seen as places along a gradient rather than as absolute black and white membership (Taylor 1989). Prototype theory similarly contends that membership in a category is not absolute. Instead, membership in a category is more or less, depending on how far a particular instantiation varies from the prototype for that category. Langacker argues that such an approach is preferable to the

216

binary approach by saying, '…[W]e do not forsake the possibility of either precise description or strong empirical claims. The choice is not an a priori matter of preference or scientific rigor, but a factual one pertaining to the organization and complexity of the data' (1987:17). In other words, the complexity of the data not only do not fit the previous binary model, but demand a theory more able to account for complex data.

This need to be able to account for complexity certainly seems to be true of the situation regarding diphthongs. Diphthongs are a complex and varied phenomenon and, as such, the theory that accounts for them also needs to be able to handle complexity. In this study, it was shown that Pamiri mountain language vowels are in the process of becoming diphthongs. In this process, they go through a number of stages. This means that, particularly near the end of that process, there are segments that are very diphthong-like, though they do not strictly meet the definition of a diphthong just yet. By appealing to prototype theory, the data can be accounted for without causing any theoretical gymnastics. It is not necessary to modify the definition of a diphthong each and every time a diphthong does not truly reach its off target. Nor is it necessary to revamp the theory of diphthongs when such speakers utter a long monophthong that contains such long drift as to begin to sound like a diphthong to the speakers hearing it. Instead, one can merely classify those instances as being less prototypical instantiations of a diphthong. On the other hand, without appealing to prototype theory, the linguist is mired in a world of data of infinite variety but forced to call each instance either 'diphthong' or 'not diphthong'. It seems that such expectations are part of what has

217

contributed to the overwhelming amount of vagueness and lack of definiteness in the literature regarding diphthongs. If it is true that language is changing and diphthongs are one of the segments most likely to change, then the theory should account for that and quit trying to force a phenomenon characterized by variety into a theory that does not account for that variety.

Early in the rise of prototype theory Jaeger and Ohala (1984) found evidence that prototype theory is relevant to phonetics and phonology. They trained informants to categorize sounds of English words based on phonological criterion. It was found that the informants defined parameters like voicing along a continuum rather than defining two discrete values. In other words, some segments were found to be a better example of voicing than others. This leads to the conclusion that the informants had formed an idea of a prototypical voiced segment that they were measuring input against. Taylor (1989) also argues that the very notion of the sonority hierarchy indicates that a prototype is in effect. Some segments are known to be much more sonorous than others and known to play a key role in many phonological processes in many languages around the world, particularly in regard to syllable shape. Given the clearly established and widely accepted role of the sonority hierarchy in phonology (cf. Kenstowicz 1994) and the results of Jaeger and O'hala's study, it seems that prototypes might be able to more accurately account for phonetic and phonological phenomenon than the discrete category model.

Indeed, prototype theory appears to already be gaining wide acceptance in the field of phonetics and phonology. Child language acquisition studies have argued that children form a prototype of phonemes and store them even before they actually begin to use them consciously in reading and writing (Rechziegel 1998). In addition, both children as young as six months old and adults were more easily able to identify instantiations of /i/ phonemes when comparisons were made against a prototype than when they were compared to non-prototypical [i] values (Kuhl 1991). Further research into the prototype of [i] revealed that subjects who were asked to choose the best example of [i] typically chose a very high and front vowel, though such a vowel actually rarely occurs in real speech. This led researchers to conclude that the minds of speakers store prototypes of vowels even if instantiations of those prototypes rarely appear in actual speech (Boersma 2005). Finally, it has been argued that 'the prototype appears to function as a ''magnet'' for other stimuli in the category, in a way similar to that shown for prototypes of other cognitive categories. Moreover, the perceptual magnet effect depends on exposure to a specific language' (Kuhl 2000:11850). Note that this research indicates that each language might have its own prototypes, which would have an immense impact on the notion of phonemes in general and diphthongs in particular.

In addition, the problem of diversity in the speech signal has long been known to phoneticians. Given such a wide array of input that varies in many ways, the question has always been how listeners manage to properly decode the speech stream. Actual

speech in performance is affected by issues as widely divergent as registers, ambient noise, mood, physical problems, socio-linguistic factors and a whole host of others. There have been two main solutions proposed to the problem of phone identification. In the exemplar model, speakers are said to 'store multiple categorized instances in memory. Categorization proceeds by matching incoming instances to the set of exemplars for each category' (Scarborough 2005). The other model is prototype theory in which speakers form prototypes. The difference is that in the prototype model, the actual instantiations of each phone are not stored in memory - only the prototype is stored. Recent evidence seems to argue in favor of the prototype theory. This is partially because of the results of some recent experiments. It has been found that informants who were confronted with a task of choosing the best form of [i], consistently chose the prototype of the [i] even over a recording of their actual own pronunciation (Boersma 2005). In other words, they chose the prototype as the better form of [i] even over the forms they actually produced themselves. Although the debate continues between the two models, such evidence lends support to the saliency of prototypes in the formation of phones.

6.2.4 DIPHTHONG THEORY

At the end of such an investigation and in light of variation theory, diachronic analysis, and prototype theory it is finally possible to identify diphthongs in a potentially meaningful and revealing way. In a way that accounts for the data, makes predictions for the future, and is elegantly simple. In light of the previous research and

the Sarikoli situation, diphthongs can be prototypically defined as *a vowel interpreted as one segment by native speakers in which formant movements show a transition from an onset position to an offset position and contain both on onset steady state and on offset steady state.*

Such a definition is only a beginning. It is already widely known that diphthongs develop from and into monophthongs, so it will not be surprising to find less prototypical instantiations of diphthongs in actual data. Such data will then need to be examined diachronically so as to make an evaluation of the path of development of such non-prototypical monophthongs. Or, alternatively, they can be examined for any sociologically motivated processes or conditions such as rapid or slow speech that might also affect the prototypicality of the diphthong. If linguists know that diphthongs are affected by all of these measures, then it makes no sense to adopt a theory that cannot incorporate such knowledge into itself. In fact, it seems that diphthongs are less controversial than they originally appeared to be if they are seen in this light. It is the theory that makes diphthongs unnecessarily controversial by not properly providing a place for linguists to incorporate all that is known about them. A good theory ought to be able to both explain data and predict phenomenon. The theory offered here can accomplish both of these goals. Approaching diphthongs with a mentality that variation will exist and interesting phenomenon can only be observed over time, supplemented with a concrete understanding of what a prototypical diphthong is, seems the best hope for ever really explaining diphthongs at all.

6.3 DIRECTIONS FOR FUTURE RESEARCH

First and foremost, this study would be greatly benefited by a true Indo-Iranian historical linguist specialist to examine the data. The reaches of historical linguistics are deep and hard to reach for the non-specialist. Although this study was able to make some predictions for future development, such an analysis would undoubtedly greatly benefit from the insights of someone who is an expert in the path of historical development of the language as a whole. Likewise, such a person would also benefit from being able to fill in gaps in previous knowledge in that field of specialty due to lack of Sarikoli data.

Next, it will have to be left to another round of data gathering to amass more information, but this would be a major next step in the expansion of the understanding of this phenomenon. Due to the constraints that exist in China, it was not possible at the time of this writing to gather more data from wider sources, but this would certainly be desirable. Especially data from people of different age groups and genders would be of benefit as they might reveal more about the trend towards diphthongization in progress. For example, is the trend the same among young people or not? Likewise, having speakers of different ages would make it more possible to conduct a so-called 'apparent time' analysis in which the pronunciations of the younger respondents would give a good indicator of the more recent trends in the language. Also, it would be beneficial to have data from a number of villages to see if the dialect continuum can be mapped more accurately. This would be true of Pamiri Tajik speaking villages on the China side of

the border, but the study should also incorporate villages all along the Gorno-Badakhshan valley. This would help to confirm or deny the current language boundaries and lect borders that are hypothesized to exist.

Even using the data already in hand, there is much further study that can be done. No acoustic or laboratory analysis has ever before been made on any dialect of Sarikoli. Thus, there are plenty of phonetic and phonological questions to be explored. Simply establishing some further phonetic basics of the language may seem mundane, but is a worthy and necessary process. For example, the data here show some signs of secondary palatalization by some speakers but not others. Currently, it is not at all clear how often or where or why such palatalization occurs. Furthermore, some speakers seem to have a much heavier onset aspiration than others. Stops initially appear to have roughly the same distinction as English stops. That is, onset stops not in clusters are heavily aspirated. Stops that occur as the second segment in a cluster (such as [sp]) seem to not be aspirated and the onset sibilant seems to become fortis. And, stops in final position seem to be unrealeased. However, no systematic analysis has ever been conducted to establish this as fact. These are just two examples of the large number of phonetic topics that could be explored with the data already in hand.

In terms of establishing the genetic relationship of Sarikoli to the other lects in the region, a wider and more broad-based analysis ought to be conducted. Rather than the trees that are offered in the current literature, a systematic study based on current phonetic and grammatical theory ought to be conducted. Such a study should rigorously

document on what basis such relationships are established and how such decisions were reached. Currently, the family tree diagrams that exist are based on data 50-60 years old and are rarely, if ever, systematically formed. Instead, they seem to be based on the overall intuitions of the linguists who studied the languages at the time. While such an intuition may ultimately be proven correct, still scientific processes ought to be observed.

Finally, the overall understanding of diphthongs ought to be re-examined in light of the usefulness of the approach taken here. Many of the languages examined in Chapter 2 were never studied in terms of newer approaches to linguistics and certainly have never been examined diachronically. It could be that some of the mysteries and headaches concerning diphthongs by other researchers could be solved if the data is re-examined using variation theory, prototype theory, and a diachronic approach.

REFERENCES

ACP. July/August 2001. Primer on Statistical Significance and P Values. Effective
Clinical Practice.

AGUILAR, LOURDES. 1999. Hiatus and diphthong: Acoustic cues and speech situation
differences. Speech Communication, 28.57-74.

AITCHISON, JEAN. 2004. Language change: Progress or decay: Cambridge approaches to
linguistics. Cambridge: Cambridge University Press.

ANDERSON, HENNING. 1972. Diphthongization. Language, 48.11-50.

ANTTILA, ARTO. 2004. Variation and phonological theory. The handbook of language
variation and change, ed. by Peter Trudgill J.K. Chambers, et al., 206-43.
Malden, MA: Blackwell.

ATWELL, ERIC, BAYAN ABU SHAWAR, ET EL. 2006. Unsupervised Learning of Linguistic
Significance. Paper presented at The 25th Leeds Annual Statistical Research
Workshop, Leeds.

BAILEY, CHARLES-JAMES N. 1981. Theory, description, and differences among linguists
(or, what keeps linguistics from becoming a science). Language &
communication, 1.39-66.

—. 1996. Essays on time-based linguistic analysis. Oxford: Clarendon Press.

BASHIRI, IRAJ. 1997. The languages of Tajikistan in perspective.
<http://www.angelfire.com/mb/bashiri/tajling/tajling.html>.

BEEKES, ROBERT S.P. 1995. Comparative Indo-European linguistics. Amsterdam: John
Benjamins.

BLADON, ANTHONY. 1985. Diphthongs: a case study of dynamic auditory processing.
Speech Communication, 4.145-54.

BOERSMA, PAUL. 2005. Prototypicality judgments as inverted perception.
<http://www.fon.hum.uva.nl/paul/papers/Prototypicality.pdf>.

BOND, Z.S. 1978. The effects of varying glide durations on diphthong identification.
Language and Speech, 21.253-63.

—. 1982. Experiments with synthetic diphthongs. Journal of Phonetics, 10.259-64.

BORZONE DE MANRIQUE, ANA M. 1976. Acoustic study of /i,u/ in the Spanish
diphthong. Language and Speech, 19.121-28.

BUREAU, XINJIANG STATISTICS. 2004. Xinjiang statistics bureau yearbook. Urumqi:
Xinjiang Provincial Government.

BYBEE, JOAN L. forthcoming. Word frequency and context of use in the lexical diffusion
of phonetically conditioned sound change. Language Variation and Change.

CHAMBERS, J.K. 2004a. Patterns of variation including change. The handbook of
language variation and change, ed. by Peter Trudgill J.K. Chambers, et al., 349-
72. Malden, MA: Blackwell.

CHAMBERS, J.K. AND PETER TRUDGILL. 2001. Dialectology. Beijing: Peking University
Press and Cambridge University Press.

CHAMBERS, J.K. AND PETER TRUDGILL, ET AL (ed.) 2004b. The handbook of language
variation and change. Malden, MA: Blackwell.

CHOMSKY, NOAM AND MORRIS HALLE. 1968. The sound pattern of English. New York: Harper and Row.

COHEN, A. 1971. Diphthongs, mainly in Dutch. Form and substance. Phonetic and linguistics papers presented to Eli Fischer-Jorgensen, 11th February 1971., ed. by L.L. Hammerich and Roman Jakobson and Eberhard Zwerner, 277-89. Copenhagen: Akademisk Forlag.

COLEMAN, LINDA AND PAUL KAY. 1981. Protoype semantics: the English word Lie. Language, 57.26-44.

COLLIER, RENÉ AND FREDERICKA BELL-BERTI, ET AL. 1982. Some acoustic and physiological observations on diphthongs. Language and Speech, 25.305-23.

DAVIS, LAWRENCE M. 1990. Statistics in Dialectology. Tuscaloosa and London: The University of Alabama Press.

DE VAAN, MICHIEL. 2003. The Avestan Vowels.vol. 12: Leiden Studies in Indo-European. Amsterdam: Rodopi.

DOLAN, WILLIAM B. AND YOKO MIMORI. 1986. Rate-dependent variability in English and Japanese complex vowel transitions F2 transitions. UCLA working papers in phonetics, 63.125-53.

EDELMAN, D.I. 1980. History of the consonant systems of the North Pamir Languages. Indo-Iranian Journal, 22.287-310.

EDMONDSON, JEROLD. 2006. Personal Communication, ed. by Pam Arlund. Dallas: Texas

EFFLAND, RICHARD, ET AL. 1995. Historical Linguistics. June 28, 2006.

      <http://www.mc.maricopa.edu/~reffland/anthropology/language/history1.html>.

ESPY-WILSON, CAROL Y. 1992. Acoustic Measures for linguistic features distinguishing

      the semivowels /w j r l/ in American English. Journal of the Acoustical Society

      of America, 92.736-56.

FASOLD, RALPH W. AND ROGER W. SHUY (ed.) 1975. Analyzing Variation in Language.

      Washington, D.C.: Georgetown University School of Languages and

      Linguistics.

GAO, ERQIANG. 1963. Tajikeyu Gaikuang. Zhongguo Yuwen, 123.161-75.

—. 1985. Tajikeyu jianzhi. Beijing: Nationalities Publishing.

GAY, THOMAS. 1968. Effects of speaking rate on diphthong formant movements.

      Journal of the Acoustical Society of America, 44.1550-73.

GEORGETOWN_UNIVERSITY. 2005. Research methods and statistics resources. October

      1, 2006.

      <http://www.georgetown.edu/departments/psychology/researchmethods/index.ht

      m>.

GORDON, MATTHEW, PAUL BARTHMAIER, AND KATHY SANDS. 2002. A cross-linguistic

      acoustic study of fricatives. Journal of the International Phonetic Association,

      32.141-74.

GOTTFRIED, MICHAEL AND JAMES D. MILLER, ET AL. 1993. Three approaches to the

      classification of American English diphthongs. Journal of Phonetics, 21.205-29.

GRIMES, BARBARA (ed.) 1996. The ethnologue:

http://www.sil.org/ethnologue/countries/taji.html.

GRIMSON, A.C. 1970. An introduction to the pronunciation of English. London: Arnold.

GUSSENHOVEN, CARLOS AND FLOR AARTS. 1999. The dialect of Maastricht. Journal of

the International Association, 29.155-66.

HARRINGTON, JONATHAN AND STEPHEN CASSIDY. 1994. Dynamic target theories of

vowel classification: Evidence from monophthongs and diphthongs in

Australian English. Language and Speech, 4.357-73.

HATTAWAY, PAUL. 2000. Operation China. United Kingdom: Piquant.

HATZIVASSILOGLOU, VASILEIOS. 1996. Do We Need Linguistics When We Have

Statistics? A Comparative Analysis of the Contributions of Linguistic Cues to a

Statistical Word Grouping System. The Balancing Act: Combining Symbolic

and Statistical Approaches to Language, ed. by Judith L. Klavans and Philip S.

Resnik, 67–94. Cambridge, Massachusetts: MIT Press.

HIRASAKA, FUMIO AND SEIZABURO KAMATA. 1981. English and Japanese Dipthongs: an

acoustic approach. Sophia Linguistica, 8-9.183-95.

IBRAHIM, NEIKRAHMON. 2006. Personal Communication, ed. by Pam Arlund

JAEGER, R. AND J.J. OHALA. 1984. On the structure of phonetic categories. Proceedings

of the Berkeley Linguistics Society, 10.15-26.

JANDA, RICHARD D. AND BRIAN D. JOSEPH. 2005. On language, change, and language

change - or, of history, linguistics, and historical linguistics. The handbook of

historical linguistics, ed. by Brian D. Joseph and Richard D. Janda, 3-180. Malden, MA: Blackwell.

JHA, SUNIL KUMAR. 1985. Acoustic analysis of the Maithili diphthongs [1985]. Journal of Phonetics, 13.107-15.

KAISER, ED. 1997. (March 15, 1997) Diphthongs as moving vowels. June 28, 2006. <http://speech.bme.ogi.edu/tutordemos/SpectrogramReading/cse551html/cse551 /node28.html>.

KEISER, STEVE HARTMAN. 2000. Sound changes across speech islands: the diphthong /aI/ in two midwestern Pennsylvania German communities. Working papers in Linguistics, 54.143-70.

KENSTOWICZ, MICHAEL. 1994. Phonology in generative grammar. Cambridge, MA: Blackwell.

KENT, RAY D. AND CHARLES READ. 2002. Acoustic analysis of speech. Albany: Singular.

KOPP, GEORGE A. KOPP AND HARRIET GREEN. 1966. Visible speech. New York: Dover.

KU'ERBAN, XIREN. 1994. Zhongguo Tajike. Urumqi: Xinjiang University Publishing.

KUHL, P.K. 1991. Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not. Perceptual Psychophysiology, 50.93-107.

KUHL, PATRICIA K. 2000. A new view of language acquisition. Publication of the National Academy of Sciences, 97.11850-57.

LABOV, WILLIAM. 2001a. Some principles of linguistic methodology. Studies in sociolinguistics: Selected papers by William Labov, ed. by William Labov, 441-76. Beijing: Beijing Languages and Literature University Press.

—. 2001b. Principles of linguistic change.vol. 2. Cambridge, MA: Blackwell.

LADEFOGED, PETER. 1996. Elements of acoustic phonetics. Chicago: The University of Chicago Press.

—. 2001. Vowels and Consonants. Malden, MA: Blackwell.

LAKOFF, GEORGE. 1987. Women, fire, and dangerous things: What categories reveal about the mind. Chicago: University of Chicago Press.

LANE, DAVID. 2006. HyperStat Online Statistics Textbook. September 28, 2006.

LANGACKER, RONALD W. 1987. Foundations of cognitive grammar.vol. 1. Stanford: Stanford University Press.

LASS, ROGER. 1984. Vowel system universals and typology: Prologue to theory. Phonology Yearbook, ed. by Colin J. and John M. Anderson Ewen, 75-112. Cambridge: Cambridge University Press.

LEHISTE, ILSE. 1967. Diphthongs versus vowel sequences in Estonian. Proceedings of the sixth international congress of phonetic sciences held at Prague 7-13 September 1967, ed. by Bohuslav and Milan Romportl, 539-44. Prague: Academic Publishing House of the Czechoslovakian Academy of Sciences.

LEHMANN, WINFRED P. 1993. Theoretical bases of Indo-European linguistics. London: Routledge.

LINDAU, MONA AND KJELL NORLIN, ET AL. 1990a. Some cross-linguistic differences in diphthongs. Journal of the International Phonetic Association, 20.10-14.

—. 1990b. Some cross-linguistic differences in diphthongs. UCLA working papers in phonetics, 61.40-44.

LINDAU-WEBB, MONA. 1985. Hausa vowels and diphthongs. UCLA working papers in phonetics, 60.40-54.

LOOS, EUGENE E. 1997. (Sep. 18, 1997) What is a diphthong? June 28, 2006. <http://www.sil.org/LINGUISTICS/GlossaryOfLinguisticTerms/WhatIsADipht hong.htm>.

LUSCOMBE, STEPHEN. 1996. (June 20, 1996) On-Line phonology course. June 28, 2006. <http://www.celt.stir.ac.uk/staff/HIGDOX/STEPHEN/PHONO/VOWEL/DIPH. HTM>.

MACLAURY, ROBERT E. 1991. Prototypes revisited. Annual review of anthropology, 20.55-74.

MAUDER, ELIZABETH AND VINCENT J. VAN HEUVEN. 1996. On the rise and fall of the Spanish diphthongs. Linguistics in the Netherlands 1996, ed. by Crit Cremers and Marcel den Dikken, 171-82. Amsterdam: John Benjamins.

MCMAHON, APRIL, M.S. 1999. Understanding language change. Cambridge: Cambridge University Press.

MINKOVA, DONKA AND ROBERT STOCKWELL. 1998. Are diphthongs neglected? Publication of the American Dialect Society.34-49.

MIRET, FERNANDO SANCHEZ. 1998. Some reflections on the notion of diphthong. Papers and studies in contrastive linguistics 34, ed. by Adam Mickiewicz, 27-51.

MOOSMÜLLER, SYLVIA. 1998. The process of monophthongization in Austria (Reading material and spontaneous speech). Papers and studies in contrastive linguistics, 34.9-25.

MORGENSTIERNE, GEORG. 1938. Indo-Iranian frontier languages.vol. 1 & 2. Cambridge, MA: Harvard University Press.

—. 1974. Etymological vocabulary of the Shughni group. Wiesbaden: Dr. Ludwig Reichert Verlag.

NEAREY, T.M. AND ASSMANN, P.F. 1986. Modeling the role of inherent spectral change in vowel identification. Journal of the Acoustical Society of America, 80.1297-308.

NORLIN, KJELL. 1984. Acoustic analysis of vowels and diphthongs in Cairo Arabic. Working papers (Lund, Sweden), 27.185-208.

OHALA, JOHN J. 2005. Phonetics and historical phonology. The handbook of historical linguistics, ed. by Brian D. Joseph and Richard D. Janda, 669- 86. Malden, MA: Blackwell.

PAKHALINA, T.N. 1960. Sarykolskii dialekt i ego otnoshenie k drugim dialektum shugnano-rushanskoi iazykovoi. Moscow: Izd-vo vostochnoi lit-ry.

—. 1966. Sarikolsko - Russkii slovar. Moscow: Institute of Linguistics, Soviet Academy of Sciences.

—. 1969. Pamirskie iazyki. Moscow: Academy of Social Sciences.

—. 1971. *Sarikolskii yazik*. Moscow: Institute of Linguistics, Soviet Academy of Sciences.

—. 1983. *Issledovanie po sravnitelyno-istoricheskoi fonetike pamirskih yazikob*. Moscow: Institute of Linguistics, Soviet Academy of Sciences.

PAYNE, J.R. 1980. The decay of ergativity in Pamir languages. *Lingua*, 61.147-86.

PEETERS, W.J.M. 1987. Vowels, diphthongs, and vowel clusters. A quantitative dynamic approach through synthesis by rule. Paper presented at New methods in dialectology: Proceedings of a workshop held at the Free University, Amsterdam, December 7-10, 1987, Amsterdam.

—. 1991. *Diphthong dynamics: a cross-linguistic perceptual analysis of Dutch, English, and German*, Utrecht University: Doctoral dissertation.

PETERS, ERIC AND ELLEN PETERS. 1996. *Engish - Sarikoli Lexicon*. Urumqi

PETERSON, G.E. AND H.E. BARNEY. 1952. Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, 24.175-84.

PHILLIPS, B.S. 1999. The mental lexicon: evidence from lexical diffusion. *Brain and Language*, 68.104-09.

PICKETT, J.M. 1998. *The acoustics of speech communication: fundamentals, speech, perception theory and technology*. Needham Heights, MD: Allyn & Bacon.

PONGWENI, ALEC J.C. 1983. Does Shona have diphthongs? *Journal of the International Phonetic Association*, 13.24-31.

RECHZIEGEL, ADELA. 1998. On the notions of target and target undershoot in the

    phonetics of L1. Paper presented at Institute of Phonetic Sciences, University of

    Amsterdam.

REZAKHANI, KHODADAD. 2001. The Iranian language family. May 12, 2006.

    <http://www.iranologie.com/history/ilf.html>.

RICKFORD, JOHN R. 2004. Implicational scales. The handbook of language variation and

    change, ed. by Peter Trudgill J.K. Chambers, et al., 142-67. Malden, MA:

    Blackwell.

ROSCH, ELEANOR. 1976. Basic Objects in Natural Categories. Cognitive Psychology,

    8.382-439.

SALZA, PIER LUIGI. 1988. Durations of Italian diphthongs and vowel clusters. Language

    and Speech, 31.97-113.

SANDS, KATHY. 2004. Patternings of vocalic sequences in the world's languages,

    Linguistics, University of California Santa Barbara.

SCARBOROUGH, REBECCA. 2005. Handout 14. June 30, 2006.

    <http://www.stanford.edu/class/linguist205/index_files/Handout%2014%20-

    %20Speech%20Perception.pdf>.

—. 2006. Lexically-conditioned phonetics: Frequency and neighborhood effects on

    coarticulation. September 20, 2006,

    <http://www.stanford.edu/~rscar/index_files/LSA%20handout.pdf>.

SCHMIDT, JOHANNES. 1872. Die Verwandtschaftsverhältnisse der indogermanischen

    Sprachen. Weimar: Hermann Böhlau.

SCHOUTEN, M.E.H. AND W.J.M. PEETERS. 2000. Searching for an explanation for diphthong perception: dynamic tones and dynamic spectral profiles. Phonetica, 57.17-39.

SCHUCHARDT, HUGO. 1868. Der Vokalismus des Vulgarlateins. Leipzig: Teubner.

SCHWARZ, HENRY G. 1984. The minorities of northern China: A survey. Bellingham, WA: Western Washington Press.

SIMS-WILLIAMS, NICHOLAS. 1996. Eastern Iranian languages. Encyclopædia Iranica, ed. by Ehsan Yarshater. New York: Encyclopedia Iranica Foundation.

SOKOLOVA, V.S. 1953. Ocherki po fonetike iranskikh iazykov. Moscow: Institute of Linguistics, Soviet Academy of Sciences.

—. 1967. Geneticheskie Otnosheniya Iazgulamskogo yazika i shignanskoy iazykovoy gruppy. Moscow: Institute of Linguistics, Soviet Academy of Sciences.

SOLE, MARIA-JOSEP. 2003. Is variation encoded in phonology? <http://seneca.uab.es/pilarprieto/Articles_Eines/Sol%E9%202003-Variation%20in%20phonology.pdf>.

SOLOMON, I. AND S.J. SARA. 1984. English diphthongs, [ai, oi, ou]. Proceedings of the 10th International Congress of Phonetic Sciences, ed. by M.P.R. Van De Broecke and A. Cohen, 411-14. Dordrecht: Foris.

STAMPE, DAVID. 1972. On the natural history of diphthongs. Paper presented at Papers from the Eighth Regional Meeting Chicago Linguistic Society, April 14-16, 1972, Chicago.

STEVENS, KENNETH N. 1998. Acoustic Phonetics. Cambridge, MA: MIT Press.

STRANGE, W. 1989. Evolving theories of vowel perception. Journal of the Acoustical Society of America, 85.2081-87.

SZEMERÉNYI, OSWALD, J.L. 1999. Introduction to Indo-European Linguistics. Oxford: Oxford University Press.

TAYLOR, JOHN R. 1989. Linguistic Categorization. Oxford: Clarendon Press.

THOMAS, ERIK R. 2001. An acoustic analysis of vowel variation in New World English.vol. 85: American Dialect Society.

VOLLMANN, RALF. 1996. Phonetics of informal speech: The Viennese monophthongization. Studia Phonetica Posnamiensia, 5.87-100.

WARDEN, MICHAEL. 1979. The phonetic realization of diphthongs. Toronto English: Studies in phonetics to honour C.D. Rouillard, ed. by Pierre R. Leon and Philippe J. Martin et al, 35-47. Ottawa: Didier.

WEIL, KAREN S. AND JAMES L. FITCH ET AL. 2000. Diphthong changes in style shifting from Southern English to Standard American English. Journal of Communication Disorders, 33.151-63.

WEISMER, GARY. 1999. Motor Speech Disorders. The Handbook of Phonetic Sciences, ed. by William J. Hardcastle. Cambridge, Mass.: Blackwell.

WOODS, ANTHONY AND PAUL FLETCHER, ARTHUR HUGHES. 1986. Statistics in Language Studies: Cambridge Textbooks in Linguistics. New York: Cambridge University Press.

WOUTERS, JOHAN AND MICHAEL W. MACON. 2002. Effects of prosodic factors on

    spectral dynamics. I. Analysis. Journal of the Acoustical Society of America,

    111.417-27.

ZONNEVELD, WIM AND MIKE TROMMELEN. 1980. Egg, onion, ouch! On the

    representation of Dutch diphthongs. Studies in Dutch Phonology, 265-92.

BIOGRAPHICAL INFORMATION

Pam Arlund received her BA *cum laude* in International Affairs with an East Asia Concentration from The George Washington University in 1993. In 1996, she received her MA in Linguistics from the University of Texas at Arlington. Since 1997, she has been living, working, and studying in China. Most recently, she has served as a Foreign Expert of English and Linguistics at Xinjiang University in Urumqi, Xinjiang, China. Her study of the Uighur language led her to translate (from English to Uighur) a book for the study of the Uighur language and also to assist Uighur speakers in learning English. The book was published by Xinjiang University Press in 2006. Currently, a book for the study of the Sarikol Tajik language is in press at Xinjiang University. She has made numerous television appearances and appeared before many government officials in China, where she hopes to continue to do research in Sarikoli, Uighur, and issues affecting second language learners in the future.