



Universiteit  
Leiden  
The Netherlands

## Topics in the syntax of Sarikoli

Kim, D.; Kim D.

### Citation

Kim, D. (2017, September 20). *Topics in the syntax of Sarikoli*. LOT dissertation series. LOT, Utrecht. Retrieved from <https://hdl.handle.net/1887/55948>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/55948>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/55948> holds various files of this Leiden University dissertation

**Author:** Kim, Deborah

**Title:** Topics in the syntax of Sarikoli

**Date:** 2017-09-20

## 1

## Introduction

In far western China, to the north and northwest of the Himalayas, along the border with Tajikistan, Afghanistan, and Pakistan, the Sarikoli<sup>1</sup> (Uyghur: *sariqoli*) people live in the high valleys of the eastern Pamir mountains, which exceed 3000 meters in elevation. This group of people, numbering about forty thousand, speaks a language that is distinct from its Turkic neighbors.

Sarikoli [srh]<sup>2</sup> is an Eastern Iranian language of the Indo-European language family. It is easternmost of the extant Iranian languages, and the only Indo-European language spoken exclusively in China. Within the Iranian languages, it belongs to the Pamir sprachbund, which is spread across the Pamir Mountains in eastern Tajikistan, eastern Afghanistan, northern Pakistan, and western China. Due to its physical and political isolation from the other Pamir languages, Sarikoli is one of the most poorly described.

The present research describes the syntax of Sarikoli as it is spoken today. In the following sections of this chapter, the Sarikoli people are introduced in terms of their geographical, cultural, and historical situation (§1.1). This is followed by a linguistic overview of the Sarikoli language, including its classification, sociolinguistic situation, typological profile, and previous research (§1.2). Finally, the framework, data, and organization of the present study are presented (§1.3).

---

<sup>1</sup>Sarikoli is not a native designation; rather, it is a Western interpretation of the Uyghur word for the people group. Native speakers refer to themselves and their own language as *tudzik*, *sariquli*, or *sarikuj*. *tudzik* is the preferred endonym, as shown in examples (2.71), (2.118), (2.215), (3.73), (5.18), (6.51), (7.63), (10.7), (10.8), (10.42), (10.154), (10.194), (11.8), and (12.8), as well as in texts A.1, A.2, A.7 in Appendix A. When it is necessary to distinguish this group from the Tajik people of Tajikistan, the more specific ethnonyms *tcin tudzik* or *dzonggo tudzik* ‘China Tajik’ may be used.

<sup>2</sup>ISO 639-3 code (Lewis, Simons & Fennig 2016)

## 1.1 The Sarikoli people

### 1.1.1 Geographical and physical context

Sarikoli speakers primarily live among the mountains of Varshide (*varçide*), which is one of the westernmost counties in Xinjiang Uyghur Autonomous Region. This county, known in the Uyghur-based English name as Tashkorgan Tajik Autonomous County (塔什库尔干塔吉克自治县), is mostly settled by the Tajik ethnicity of China. The ethnonym “Tajik (塔吉克族)” in China covers Iranian peoples who speak three distinct native languages: Sarikoli, spoken by the majority, Wakhi (also Eastern Iranian), and Uyghur (Turkic). Even though Sarikoli and Wakhi are both Eastern Iranian languages, they are mutually unintelligible, and their speakers are culturally similar but ethnically distinct. Speakers of these three languages became an officially recognized ethnic minority of the PRC in 1954, the same year that their homeland officially became Tashkorgan Tajik Autonomous County. According to the Sixth National Population Census of the People’s Republic of China conducted in 2010, there were 51069 Tajiks in China. Since the majority of Chinese Tajiks speak Sarikoli, we estimate that there are about 40000 speakers of Sarikoli. The remainder of the Tajik ethnicity in China speaks Wakhi or Uyghur as their primary language. The Uyghur-speaking Tajiks speak neither Sarikoli nor Wakhi, but they identify with the Sarikoli and Wakhi speakers culturally and religiously.

Varshide County is officially composed of eleven *gungci* (公社), or communes<sup>3</sup>, which represent the main villages. The commune names are listed below in Table 1.1, followed by three other place names that have significant communities of Sarikoli residents and are frequently mentioned in conversations. The right-hand column shows how the village names are spelled in Neikramon Ibrukhim’s orthography. The central town and administrative county seat established by the Chinese government is also called Varshide, bearing the same name as the county itself. There are smaller villages which fall under the administration of each of the eleven main villages. Thavthor has the largest settlement of Wakhi speakers, although the Wakhi are thoroughly spread out among the other villages as well, partially due to intermarriage between the Sarikoli and Wakhi speakers. Kekyor is officially a Kyrgyz village which is primarily settled by the Kyrgyz, as it is the northernmost village and geographically closest to the Kizilsu Kyrgyz Autonomous Prefecture. Another

<sup>3</sup>Communes are a result of Maoist era Chinese government policy that dates only to the 1950s or later, not a Sarikoli cultural feature.

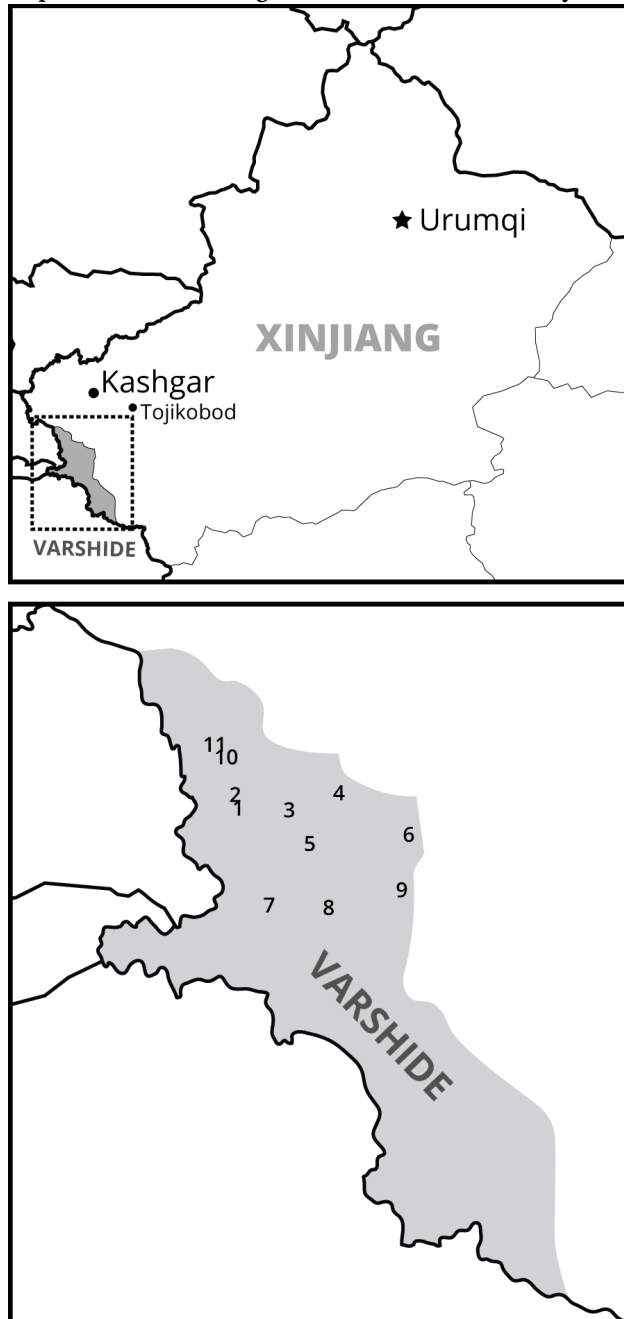
noteworthy village is a small village called Tor, located between Koghushluk and Teeng, but lying just outside the county border; it is the home of most Uyghur-speaking Tajiks.

Table 1.1 The eleven main villages of Varshide County, and other place names

	<b>IPA transcription</b>	<b>Orthographical spelling</b>
1	<i>varçide</i>	Varshide
2	<i>tuznef</i>	Teeznef
3	<i>baldir</i>	Baldir
4	<i>køuçluk</i>	Koghushluk
5	<i>watça</i>	Wacha
6	<i>tung</i>	Teeng
7	<i>ðavðor</i>	Thavthor
8	<i>marjong</i>	Maryong
9	<i>brumsol</i>	Brumsol
10	<i>taɣarmi</i>	Tagharmi
11	<i>kəkjor</i>	Kekyor
12	<i>toɖzikobod</i>	Tojikobod
13	<i>xwor</i>	Kashgar
14	<i>urumtçi</i>	Urumqi

The following map shows the locations of these eleven main villages. Their numbering is in the same order as they are listed in Table 1.1 above.

Figure 1.1 Map of the main villages of Varshide (created by Moss Doerksen)



There are a few Tajik resettlement towns outside of Varshide County, such as in Poskam County of Kashgar Prefecture (喀什地区泽普县), Akto County in Kizilsu Kyrgyz Autonomous Prefecture (克孜勒苏州阿克陶县), and Hotan Prefecture (和田地区), but the biggest and most prominent resettlement town is Tojikobod Town in Kashgar Prefecture (喀什地区塔吉克阿瓦提镇). These have started as Chinese government-initiated resettlements due to extensive flooding in Varshide, particularly in the villages of Teeng, Brumsol, and Koghushluk. However, Tojikobod Town now has many residents who have migrated from various villages of Varshide for reasons such as lower altitude, wider range of agricultural options, better educational opportunities, and proximity to the city of Kashgar. Tojikobod Town is inhabited by speakers of Sarikoli, Wakhi, and Uyghur, but residents in some of the other resettlement towns have virtually given up speaking Sarikoli and Wakhi in favor of Uyghur as they have lived in constant contact with Uyghur neighbors.

According to folk etymology, the ethnonym *sarikuj* derives from the Persian words *sar* ‘head’ and *kuh* ‘mountain’, which reveals the sense of pride and identity they take in living on the “Roof of the World” among some of the world’s highest mountains. The Pamir Mountains stretch from the Gorno-Badakhshan Autonomous Region of Tajikistan in the west, to Varshide in the east. The average elevation of Varshide County is 4000 meters, and the Muztagh Ata (7509 meters) and the Kongur Tagh (7649 meters) peaks are in the close vicinity of these people. The central town of Varshide is about 3000 meters.

The Sarikoli people are traditionally farmers and semi-nomadic herders. As farmers, they grow highland barley and green peas, and have begun to grow highland *maka* (*Lepidium Meyenii*) as a cash crop. As herders, they move to higher pasturelands in the summertime to graze their sheep, goats, and yaks for months at a time. Naturally, their diet primarily consists of meat, wheat, and dairy, although consumption of fruits and vegetables brought from Kashgar has been on the increase for those living in the central town of Varshide. Teeng and Koghushluk, the villages lower in elevation, also produce large quantities of fruits, especially apricots.

### 1.1.2 Religious and cultural context

Most Sarikoli people adhere to the Ismaili branch of Shi’a Islam, and claim continuity with Zoroastrian traditions as well. The three most significant festivals of the year are Sheawgeenbahor/Neawreez Eid (*çawgunbahor/nawruz ejd*), Qerbun Eid (*qurbun ejd*), and Pilik Eid (*pilik ejd*). Sheawgeenbahor Eid is the Iranian New Year and Zoroastrian festival, which begins on March 21 of the Western calendar. It is a three-day celebration during which everyone cleans their home (which is why it is also commonly called *tçed tçader ejd*

‘house cleaning festival’), wears new clothes, and visits all of the houses in their village to pass on good wishes and enjoy festival food. Qeerbun Eid, the ‘sacrifice festival’ of Islam, is celebrated on the tenth day of Dhu al-Hijjah in the Islamic calendar, in remembrance of Ibrahim’s willingness to obey God and sacrifice his son, Ishmael. At daybreak on the first day of this three- or four-day celebration, a ram is sacrificed on the rooftop, and its meat is shared with numerous guests who pay their visits throughout the day. Pilik Eid, the ‘wick festival’, is celebrated on the fourteenth and fifteenth days of Sha’ban in the Islamic calendar. The first day is called *tçed pilik* ‘home pilik’, on which they light a fire at home for the living family members. The second day is called *zurat pilik* ‘graveyard pilik’, as they go to the graveyards of their ancestors and light a small fire on the tomb of each deceased relative.

The Sarikoli people are patrilineal and patrilocal. Intermarriage with non-Tajik ethnicities is extremely rare, and currently all marriages are monogamous. Sarikoli people have preferred to marry within their extended family because travel to other villages has been difficult in the past, as well as the fact that relatives could assist each other financially and expected reasonable dowry and bride price. However, with improved road conditions and mobile communications, marrying a non-relative from another village has become possible and even commonplace. Before a wedding, the prospective groom’s male relatives first visit the prospective bride’s home to seek permission from her parents, taking some animals as gifts. Once permission has been granted, the engagement party (*rejmultarkol*, lit. ‘scarf to head’) is celebrated in both the prospective groom’s and bride’s homes. The wedding occurs a few months after this, and is celebrated for four to five days. The bride wears a red dress, adorns herself with jewelry and ornaments of silver and jade, and covers her face with a white veil (*tçumband*). The groom wears black, with a red and white cloth (*sala*) braided around the usual black wool hat (*tumox*) worn by men. Large celebrations take place at both the groom’s and the bride’s homes, each with crowds of guests, an abundance of food and sheep-slaughtering, and hours of dancing accompanied by loud music. Relatives and neighbors help with preparing and serving food, and guests enjoy themselves by dancing and watching others dance. On the third day or so, the groom, accompanied by a female relative (*rawots*) and two groomsmen (*xanitsamuug*), goes to the bride’s home to pick up the bride. After the bride and groom arrive at the groom’s home, they participate in the *niku*, a solemn religious ceremony performed by the *çalifa*, the religious leader. This is when they officially become bride and groom. On the final day of the wedding, the white veil covering the bride’s face is lifted, and the guests are able to see the bride’s face. Almost all weddings take place in the summertime or after harvest in the fall.



The eagle is the symbol for the Sarikoli people, as it is for the Pamir peoples in general. It represents freedom, strength, and beauty. The Sarikoli people mimic the eagle when they dance, play flutes (the *noj*) made of eagle wing bones, and claim that their noses resemble the eagle's beak.

The Sarikoli people's favorite leisurely activities include dancing, singing, and embroidery. Every major festival or significant event reserves special time and space for dancing and singing, but these activities may spring up at any gathering of Sarikoli people, often for no particular reason at all. The women are constantly embroidering pillows, home decorations, and ethnic hats whenever they have free time. When a woman gets married, she is expected to give a newly-embroidered ethnic cap (*cedoi* or *kulto*) to every female relative in the groom's extended family. As each cap generally takes at least a month to make, girls and their mothers are always busy embroidering caps when a wedding is imminent.

Colors, especially as shown on clothing, are significant for signalling social emotions. *χuŋi* 'happiness' is expressed by colors like red, orange, yellow, and pink. Since a wedding is a happy occasion, the bride is dressed in red from head to toe and the groom also wears red and white cloths around his hat. The bride is expected to wear red for at least one year after the wedding as well. Recently married women or women who are young and youthful generally wear traditional embroidered caps with happy colors. *χafagi* 'sadness', on the other hand, is expressed by colors like blue, green, and black. Everyone at a funeral wears these sad colors, often also with a blue or green cloth around their waist, and relatives and close friends of the deceased wear these colors for at least a year. They also refrain from activities that are perceived to be happy, such as dancing and singing or having a wedding within the family. Older women nearing death or women whose relatives have passed away recently will wear traditional caps containing more of the sad colors.

### 1.1.3 Historical context

Sarikoli lacks a native account of origins and history. The people themselves often claim to have been living in the Pamir Mountains since the beginning of time, and that they are the oldest Iranian civilization speaking the original or most ancient variety of Persian. Given the harsh conditions on the eastern Pamir plateau, Sarikoli people reason that no one would choose to ascend the mountain; instead, they conveniently descended from their mountain dwellings.

Shughni and Rushani, the most closely-related languages to Sarikoli, are spoken in eastern Tajikistan and Afghanistan. According to Dodykhudoeva, the

Sarikoli people migrated several centuries ago from the Upper Bartang of the Gorno-Badakhshan Autonomous Region of Tajikistan. More populations fled from Upper Bartang in 1911, when the massive Sarez-Pamir earthquake triggered landslides and destroyed their villages (2004:2).

## 1.2 The Sarikoli language

### 1.2.1 Classification: The place of Sarikoli in Iranian languages

The Iranian languages are a branch of the Indo-European language family, and are subdivided into eastern and western groups. The Western Iranian languages include Kurdish, Balochi, and Persian languages. The Eastern Iranian language family includes the Pamir languages, as well as Pashto, Ormuri, Parachi, Yaghnobi, and Ossetian. The Pamir languages, which are spread across the Pamir Mountains in Tajikistan, Afghanistan, Pakistan, and China, are located on the far eastern edge of the area where Iranian languages are spoken today.

There is general agreement that the Pamir languages constitute a common Pamir sprachbund, or areal grouping, rather than a genetic grouping (Morgenstierne 1938; Sokolova 1967; Paxalina 1969 & 1983; Payne 1989; Edelman & Dodykhudoeva 2009a; Wendtland 2009). Within the Pamir sprachbund, etymological evidence suggests that Sarikoli, Shughni, Rushani, and possibly Yazgulyam comprise a genetically-related subgroup, whereas the others—such as Wakhi, Ishkashimi, Munji, and Yidgha—are not closely related genetically (Sokolova 1967; Payne 1989; Edelman & Dodykhudoeva 2009a).

### 1.2.2 Sociolinguistic situation

Sarikoli is surrounded by unrelated languages. The political border between China and the Central Asian countries limits Sarikoli speakers' contact with speakers of other Pamir languages to the west, while increasing their relative contact with speakers of Turkic languages. Xinjiang is the homeland of tens of millions of speakers of Turkic languages, including Uyghur, Kyrgyz, Kazakh, Uzbek, and Tatar. Mandarin Chinese is also increasing in prominence due to education policies and socioeconomic pressures.

Besides the Tajik ethnicity, the three largest ethnic groups living in the county of Varshide are Han (the Chinese majority), Uyghur, and Kyrgyz, but they constitute an extremely small portion of the overall population of the county. The Hans and Uyghurs come to Varshide to run small businesses, a trade which

the Tajiks rarely get involved in. The Uyghurs come from various places in Xinjiang, especially Kashgar, the nearest city in China which is 300 kilometers northeast of Varshide. The Hans come from much more distant places all over China. The Kyrgyz are generally farmers and herders, just like the Tajiks, and they are close to their homeland because they belong to the Kizilsu Kyrgyz Autonomous Prefecture and Kekyor, the Kyrgyz village in Varshide.

Currently, each of the 10 main villages besides the county seat has a small elementary school, and the county seat has a very large elementary school with thousands of students, which provides room and board for students from other villages. Elementary school education is six years, followed by three years of middle school and three years of high school. The only middle school in the entire county is located in the county seat, and is also a boarding school with thousands of students. There are no high schools in Varshide, so students must leave Varshide and go to cities such as Kashgar, Urumqi, or other cities in Xinjiang or Innerland China to pursue higher education. Rather than the national-level Law on Nine-Year Compulsory Education, Varshide complies with southern Xinjiang's Law on Twelve-Year Compulsory Education, so all Tajik children must leave their hometown and spend at least three years in a generally Han- or Uyghur-speaking city. The majority of students attend the No. 6 High School and No. 2 High School in Kashgar, but the top students are granted the privilege of receiving their high school education in a city in eastern China on a government scholarship. Tajik students who attend high school in Innerland China (outside of Xinjiang) are obligated to also attend college in Innerland China, and these students usually become more comfortable with Mandarin than their native language.

Sarikoli is not taught in schools, neither as the language of instruction nor as a separate language subject. Up until a few years ago, the languages of instruction at the schools in Varshide were Mandarin and Uyghur. When being enrolled in first grade, students and their parents were to choose either the Mandarin track or the Uyghur track, a decision which lasted until the end of their education career. Initially, most chose Uyghur, which is why many people from the middle-aged generation now are more comfortable with Uyghur than Mandarin. However, around 2010, the Uyghur track has been abolished in the first grade, leaving Mandarin as the only option for the entire class. As the Mandarin-only classes move up each year, Tajik children are increasingly speaking more Mandarin. Mandarin is the only language that is permitted in school, both in class and outside of class, and children are forbidden to communicate with each other in Sarikoli or other languages.

Television and radio are available in Uyghur and Mandarin only. Families watch Uyghur television together after the evening meal, as Uyghur continues

to be the language understood by the older and younger generations alike. However, based on current trends, Mandarin seems likely to take over as the dominant second language in the future.

As the Language of Wider Communication and one of the official government languages of the province, Uyghur is naturally viewed as having higher prestige than Sarikoli. It also has a rich literary tradition and has been a language of instruction in schools, which have not been opportunities for Sarikoli. Apart from these official domains, Uyghur is also ubiquitous in popular media, both on television and radio. It is the language spoken by an ethnic group with a much larger population and greater political power than the Sarikoli people. It provides far greater socio-economic opportunities.

The Sarikoli people retain a positive attitude toward their own language. They have a strong sense of identity as the only Iranian-speaking group in China, and take great pride in their language and culture. Language use is vigorous, and speakers of all generations are fluent in their language, unless they have spent most of their lives studying in Innerland China. As Varshide is isolated from other Han- or Uyghur-majority cities, Sarikoli speakers still use their native language for most interactions with people in their daily lives. In addition, they show great enthusiasm and passion for cultural artifacts in Sarikoli, such as songs, poetry, and proverbs.

Within the Tajik ethnicity of China, Sarikoli has a higher prestige than Wakhi because it is spoken by the majority. Most of the Wakhi people also learn to speak Sarikoli fluently in order to communicate with other Tajiks, but some communicate with them through Uyghur. Sarikoli speakers rarely learn to speak Wakhi fluently; if they do, it is usually because they were raised by Wakhi-speaking family members. Intermarriage between the Sarikoli and Wakhi groups is common. However, the Uyghur-speaking Tor Tajiks (*tor tudzik*) tend to take more pride in their unique identity and are less likely to intermarry with Sarikoli or Wakhi Tajiks.

Because speakers are spread out throughout the mountains and valleys across 52400 square kilometers of land, Sarikoli is not homogenous. Paxalina (1966:3) noted dialectical differences among three general regions: central (including the county seat of Varshide, Teeznef, Cheekhmon, and parts of Baldir), near eastern (including Wacha, Maryong, and parts of Baldir), and far eastern (including Teeng and Brumsol). Differences among these variants are mostly phonetic, with some lexical variation as well.

### 1.2.3 Typological overview

Sarikoli is a moderately agglutinating language with SOV basic word order. Peripheral arguments and adverbial modifiers are typically placed between the subject and the object. Head-final morphosyntactic behavior is shown through the ordering of constituents: objects precede the verb, nominal modifiers precede the head noun, and degree words precede the adjective. Both prepositions and postpositions are used, some of which are coded for relative elevation. Suffixes are more prevalent than prefixes. Interrogative words occur *in situ* in content questions, and the question enclitic which marks polar questions occurs sentence-finally. Grammatical relations are signaled through case and function marking on nouns and pronouns, constituent order, and pronominal subject-verb agreement clitics. Verbs can be analyzed in five different stems, and aspect is indicated through a combination of the choice of verb stem, aspectual clitics and suffixes, and the form and placement of pronominal clitics.

### 1.2.4 Previous research

Sarikoli is an underdescribed and poorly documented language. Arlund describes it as “the most isolated and understudied of the [Pamir] languages” as a result of its confinement to a remote border area of China, presenting great challenges to linguists in terms of geographical remoteness, requirement of Mandarin proficiency, and the red tape and surveillance of the Chinese government (Arlund 2006:6). Paxalina speculates that Sarikoli has kept many words and forms lost in other Pamir languages due to its geographical and political isolation from other Pamir languages (Paxalina 1966:4).

Few linguists have produced descriptions of Sarikoli based on data from their own fieldwork, and they will be introduced in this section. Although Sarikoli has also been mentioned in several general works on Pamir languages or the Shughni-Rushani subgroup (Lentz 1933; Sköld 1936; Morgenstierne 1938 & 1974; Payne 1989; Skjærvø 1989; Edelman & Dodykhudoeva 2009a; Wendtland 2009), those works are based on materials published by those who did original research in the 1870s and 1950s: Shaw (1876) and Paxalina (1966).

The first English mention of Sarikoli appeared in 1875, when Britain sent an official mission to Eastern Turkestan (present-day Xinjiang) led by diplomat Thomas Douglas Forsyth in 1873, during the closing decades of the Great Game, the struggle between Victorian Britain and Tsarist Russia for geopolitical power in Central Asia. Two of the participants of this expedition, medical Dr. Henry Walter Bellew and Colonel John Biddulph, collected substantial wordlists and twenty phrases of Sarikoli (to which they refer as *Sarigh Culi*

and *Sirikolee*, respectively). These data are in chapter 15 of Forsyth's report on this mission, which also includes rich historical, geographical and ethnographical information on western Xinjiang (Forsyth 1875). Bellew and Biddulph's wordlists can be useful for historical-comparative work.

The first English description of Sarikoli was written by Robert B. Shaw, a British political agent who was on special duty at Kashgar (Shaw 1876). In 1868, he was "the first Englishman who ever went to Yarkund" (Forsyth 1871), a county off the northeast border of Varshide, just a short distance away from the village of Teeng. In 1872, when he returned to England, he was awarded the patron's gold medal by the Royal Geographical Society for his service in exploring Eastern Turkestan (Lee 1897). He also published several linguistic descriptions of the languages of Xinjiang and the Pamir Mountains, including: *On the Ghalchah languages (Wakhi and Sarikoli)* (1876), *On the Shigni (Ghalchah) dialect* (1877), *A Sketch of the Turki Language as spoken in Eastern Turkestan* (1878a), and *On the Hill Canton of Salar: the most easterly settlement of the Turk race* (1878b).

In *On the Ghalchah languages (Wakhi and Sarikoli)* (1876), Shaw provides a brief sketch of Wakhi and Sarikoli grammar, followed by several narrative texts in each language, accompanied by literal English translations. He also includes a lengthy lexicon of Sarikoli and Wakhi. This work is a resource for a diachronic study of Sarikoli, with texts and lexicon from the 1870s. It is useful for investigating how the language has changed and developed since then, and which elements have remained constant. Paxalina (1966) evaluates Shaw's work as beneficial, even though there are mistakes and inaccuracies because he was not a trained linguist.

About eight decades later, in the 1950s, a Russian linguist named Tatiana N. Paxalina came to research Sarikoli and related Pamir languages, including Shughni, Rushani, Ishkashimi, and Wakhi. She collected Sarikoli data in 1956.<sup>4</sup> In 1966, she produced a sketch of Sarikoli grammar which also includes narrative texts with literal translations into Russian, and later in 1971 published a Sarikoli-Russian dictionary. Because of the amount of detail she provides in her description of Sarikoli and her extensive experience researching Pamir languages, her work is considered the most reliable and in-depth grammatical analysis of the Sarikoli language to date.

In the 1960s, a Chinese linguist, Gao Erqiang, conducted research on Sarikoli and Wakhi, the two Iranian languages spoken by the "Tajik" ethnicity of China

<sup>4</sup>My field research has put me in contact with someone who remembers Paxalina conducting research in Varshide when he was a child.

(Gao 1963). This was part of the Chinese initiative to produce brief descriptions of each of the minority languages of China. In 塔吉克语简志 (Outline of the Tajik language), which appeared in 1985, he presents an overview of the phonetics, lexicon, morphology, and syntax of Sarikoli, and also includes a description of Wakhi, referring to it as a “dialect” of Sarikoli. In 1996, he published a Tajik-Mandarin dictionary.

In the 1990s, Pamela Arlund, an American linguist, began researching Sarikoli. Her PhD dissertation, an acoustic analysis of Sarikoli diphthongs, appeared in 2006. A few years later, she co-authored an English-language primer intended for non-linguist learners of Sarikoli, in cooperation with Neikramon Ibrukhim, a native speaker of Sarikoli. This primer does not contain original native texts, but has grammatically acceptable translations of Uyghur texts, as well as word lists with English, Mandarin Chinese, and Uyghur glosses.

Neikramon Ibrukhim is a Sarikoli scholar who is passionate about promoting and developing his own language. He is a professor of Russian at Xinjiang University and also works at the Foreign Affairs Office of the university. In addition to co-authoring the English-language primer with Arlund, he has developed a Roman-script-based orthography of Sarikoli and published a primer introducing his alphabet (Ibrukhim 2012; see Appendix B for correlations with the IPA). Although his alphabet is still far from being widely used within the Sarikoli community, he transcribes stories, poems, song lyrics, and news articles with his orthography and disseminates it on social media. By doing so, he hopes to pass on the language to younger generations and maintain its vitality.

Publications based on the most recent original linguistic research conducted on Sarikoli include works by Kim (2014, 2015) and Palmer (2016). These are not comprehensive grammars, but descriptions of specific phenomena of Sarikoli syntax and morphology.

## 1.3 The present study

### 1.3.1 Scope and descriptive theoretical framework

This dissertation presents an analysis of selected topics in the syntax of Sarikoli. It was originally intended to form half of a joint dissertation, but the demands of life, family, and education have made it more prudent to write separate dissertations covering different topics. Upon completion of Timothy Palmer’s dissertation on topics including Sarikoli phonology, morphology, and the verb

and verb phrase, the two dissertations will be joined together as a single comprehensive grammar of Sarikoli. While this description focuses on syntactic topics, the verb phrase is not covered in detail here, because analysis of verbs and verb phrases include much analysis of morphology and especially of aspect, which my partner is better prepared to address. In order to help the reader understand the discussions in this dissertation, a brief phonological and morphological sketch including verbal morphology is provided in §1.4.

The present study is a synchronic description of the syntactic structure of Sarikoli. As such, I do not theorize about the place of Sarikoli within Iranian languages through diachronic analysis, though the data and description provided here may be useful for historical-comparative work in future studies. Because this is a descriptive grammar, I have chosen to use a descriptive theoretical framework, Basic Linguistic Theory (Dixon 1997, Dryer 2006), rather than an explanatory theory, to analyze and present my data. The descriptive focus of this grammar assumes a minimal amount of theoretical knowledge on the part of the reader, and I use terms that are generally familiar to all linguists. In exceptional cases where it is necessary to use terms specific to Sarikoli, they are explained as they are introduced.

### 1.3.2 Fieldwork and data

This section summarizes the process of fieldwork and the scope of data on which this grammar is based.

We conducted the fieldwork for this dissertation between September 2014 and December 2016. The three principal locations of field research for this grammar are: 1) various villages in Varshide County (Varshide, Teeznef, Tagharmi, Rabut, Teeng, Wacha); 2) Tojikobod Town, the Sarikoli resettlement town in Kashgar Prefecture; and 3) Urumqi, the provincial capital of Xinjiang Uyghur Autonomous Region. Data collection was carried out in Sarikoli, as it is the language that I use to communicate with the Sarikoli people. Most of the writing of this grammar was done on-site during fieldwork, and all of the examples have been checked by native speakers.

We recorded 15 folktales (142 minutes), 25 cultural, traditional, and historical texts (239 minutes), 10 personal experience texts (32 minutes), 15 conversation texts (35 minutes), 20 procedural texts (16 minutes), 4 poems (6 minutes), 3 hortative texts (3 minutes), 7 traditional songs (22 minutes), and a collection of proverbs (29 minutes). Texts were transcribed, analyzed and glossed in FieldWorks Language Explorer (FLEx), and translated into English. This dissertation is based on these texts and conversations. Some sample texts representing various genres and topics are provided in Appendix A. In addition



to sentences taken from this corpus of recorded natural data, many examples are taken from utterances that occurred in natural conversations, which were transcribed on the spot.

Twenty-nine Sarikoli speakers of a variety of ages, occupations, village origins, and genders contributed oral texts for this study. Of these, twenty-eight were born and raised in Varshide County and one was born and raised in Tojikobod Town. A large number of native speakers also assisted by providing and translating data. Neikramon Ibrukhim, who is originally from the Varshide county seat, provided great help by introducing us to Sarikoli speakers in various villages who were willing to share oral texts. Gawar Deyqun, a native of Wacha, has accompanied us when collecting some of the oral texts and has worked with us for countless hours on transcription and translation. He has provided much insight into his language and culture.

In cases where there are differences among the dialects, we describe the majority or most pervasive form.

### **1.3.3 Transcription**

In this grammar I use a phonemic IPA representation of Sarikoli, as this is a dissertation written in English primarily geared towards an international linguistic audience. Sarikoli does not have an officially implemented orthography yet, and different members of the Sarikoli community wish to use different types of script for their orthography, so a phonemic IPA representation appears to be the most appropriate and politically neutral choice for the purposes of this grammar.

In the free translations of examples, proper nouns (mainly names of people, places, and festivals) and names of cultural items or concepts that are unique to Sarikoli are given in the orthography developed by Neikramon Ibrukhim, which is based on the Roman script. This orthography and correlations with the IPA are presented in Appendix B.

## **1.4 Phonological and morphological sketch**

This section gives a brief overview of the phonology and morphology of Sarikoli in order to provide the reader a basis for understanding the discussions on syntax in the following chapters. The present study does not include an in-depth examination of phonology, morphology, and verbs (including aspect and transitivity) beyond what is discussed in this short section.

### 1.4.1 Phonology

#### 1.4.1.1 Consonant and vowel phonemes

Sarikoli distinguishes thirty consonant phonemes and eight vowel phonemes, as listed in Table 1.2 and Table 1.3 below.

Table 1.2 Sarikoli consonant phonemes

	Labial	Dental	Alveolar	Alveolo-palatal	Velar	Uvular	Glottal
<b>Stop</b>	p b		t d		k g	q	
<b>Affricate</b>			ts dz	tʃ dʒ			
<b>Fricative</b>	f v	θ ð	s z	ʃ ʒ	x ɣ	χ ʁ	h
<b>Nasal</b>	m		n				
<b>Trill</b>			r				
<b>Lateral</b>			l				
<b>Glide</b>	w			j			

Table 1.3 Sarikoli vowel phonemes

	Front	Central	Back
<b>Close</b>	i		u
<b>Close-mid</b>	e		o
<b>Mid</b>		ə	
<b>Open-mid</b>	ɛ		
<b>Open</b>		a	

#### 1.4.1.2 Stress

In general, primary stress falls on the final syllable of nouns, adjectives, and adverbial modifiers. Verb stress is more variable and sometimes falls on the first syllable, as shown in (1.1). In a compound verb, stress usually falls

on the final syllable of the nominal element, and not on the inflected verb that follows, as in (1.2). Most grammatical morphemes, such as pronominal agreement clitics, function-marking clitics and adpositions, aspectual morphemes, conjunctions, and modal particles like the conditional *tsa* and the ability marker *tçi*, are not stressed, as in (1.3). The nominalizer *-i*, diminutive suffix *-ik*, and negators (*na*, *nist*, *mo*, *naj*) are exceptions, as they do receive stress, as in (1.4). In the following examples, stress is indicated in the second line.

- (1.1) *askar-xejl = af*                      *a = bejroq naymuug*  
*as'kar-xejl = af*                      *a = bej'roq 'naymuug*  
 soldier-PL.NOM = 3PL.PFV    ACC = flag    hide.PFV  
 ‘The soldiers hid the flag.’
- (1.2) *niso pa maktab xtsuvd usul xumand sut*  
*ni'so pa mak'tab x'tsuvd u'sul xu'mand sut*  
 Niso LOC school eagle dance learn    become.PFV  
 ‘Niso learned the eagle dance at school.’
- (1.3) *ɛadar*                      *tɛdz*    *tsa*    *pa*    *puiz*    *dejd*    *tçi*  
*ɛa'dar*                      *'tɛdz*    *tsa*    *pa*    *pu'iz*    *'dejd*    *tçi*  
 three.days.hence go.IPFV COND LOC train enter.INF CAP  
  
*ka*  
*'ka*  
 do.IPFV  
 ‘If you go three days from today, you can get on the train.’
- (1.4) *xalisa az turik-i*            *xudz na*    *đord*  
*xali'sa az turik-'i*            *xudz 'na*    *đord*  
 Halisa ABL dark-NMLZ fear NEG fear.3SG.IPFV  
 ‘Halisa is not afraid of the dark.’

#### 1.4.1.3 Glide epenthesis

The glide [j] is epenthesized between two adjacent vowels as a hiatus resolution strategy. In the following examples, the first line represents the bare lexical forms and the second line accounts for morphophonemic epenthesis.

- (1.5) *watça at baldir*  
*watça jat baldir*  
 Wacha CONJ Baldir  
 ‘Wacha and Baldir’
- (1.6) *na=am vuusond*  
*na=jam vuusond*  
 NEG = 1SG.PFV show.PFV  
 ‘I did not show it.’
- (1.7) *samu=at tçəwg=o*  
*samu=jat tçəwg=o*  
 walk = 2SG.PFV do.PFV = Q  
 ‘Did you take a walk?’
- (1.8) *a=di mu=ri hat ka=o*  
*a=di mu=ri hat ka=jo*  
 ACC = 3SG.NNOM.PROX 1SG.NNOM = DAT open do.IPFV = Q  
 ‘Will you open this for me?’
- (1.9) *ar ujnak agar m=k=dos tçost tsa*  
*ar ujnak agar m=k=dos tçost tsa*  
 LOC glass if CATA = ANA = manner look.3SG.IPFV COND
- u*  
*ju*  
 COND  
 ‘If he looks into the mirror like this...’

## 1.4.2 Morphology

### 1.4.2.1 Verb stems

Each Sarikoli verb can be analyzed as having an infinitive stem, as well as four finite stems: imperfective, third-person singular imperfective, perfective, and perfect. The formation of these aspectual stems is somewhat predictable for some verbs; in these regular verbs, the perfective stem is usually formed by adding a /t/ or /d/ ending to the imperfective stem (depending on the voice of the segment it attaches to), and the perfect stem is formed by changing those endings to /tç/ or /dç/ (Payne 1989:436). Sometimes the infinitive stem is identical to the perfective stem. The third-person singular imperfective stem

is identical to the past stem or the infinitive stem, or sometimes unique. Some regular verbs and their stems are presented in Table 1.4.

Table 1.4 Examples of regular verbs

	IPFV	3SG.IPFV	PFV	PRF	INF
‘say’	<i>lev</i>	<i>levd</i>	<i>levd</i>	<i>levdʒ</i>	<i>levd</i>
‘gather’	<i>wix</i>	<i>wixt</i>	<i>wixt</i>	<i>wixtʃ</i>	<i>wixt</i>
‘ask’	<i>pars</i>	<i>parst</i>	<i>parst</i>	<i>parstʃ</i>	<i>parst</i>
‘dig’	<i>kəw</i>	<i>kəwd</i>	<i>kəwd</i>	<i>kəwdʒ</i>	<i>kəwd</i>
‘write’	<i>naviç</i>	<i>naviçt</i>	<i>naviçt</i>	<i>naviçtʃ</i>	<i>naviçt</i>
‘use’	<i>rafon</i>	<i>rafond</i>	<i>rafond</i>	<i>rafondʒ</i>	<i>rafond</i>
‘know’	<i>wazon</i>	<i>wazond</i>	<i>wazond</i>	<i>wazondʒ</i>	<i>wazond</i>

However, there are a number of more morphologically variable verbs whose stems cannot be predicted. The stem modification in these irregular verbs involves vowel and consonant alternation, but the first segment of the verb usually remains the same in all five stems. Table 1.5 lists some irregular verbs and their stems. The first is a morphologically suppletive paradigm.

Table 1.5 Examples of irregular verbs

	IPFV	3SG.IPFV	PFV	PRF	INF
‘do’	<i>ka(n)</i>	<i>kaxt</i>	<i>tʃəwɔŋ</i>	<i>tʃəwɔydz</i>	<i>tʃejɔŋ</i>
‘become’	<i>so</i>	<i>səwd</i>	<i>sut</i>	<i>seðdʒ</i>	<i>set</i>
‘eat’	<i>χor</i>	<i>χird</i>	<i>χuɔŋ</i>	<i>χuɔydz</i>	<i>χig</i>
‘come’	<i>joð</i>	<i>joðd</i>	<i>jot</i>	<i>iθtʃ</i>	<i>jɛt</i>
‘bring’	<i>vor</i>	<i>virɔd</i>	<i>vəwɔŋ</i>	<i>vəwɔydz</i>	<i>vejɔŋ</i>
‘grind’	<i>jon</i>	<i>jigɔd</i>	<i>juɔŋ</i>	<i>juɔydz</i>	<i>jig</i>
‘disappear’	<i>bis</i>	<i>bast</i>	<i>bejd</i>	<i>beðdʒ</i>	<i>bejd</i>

Sentences are formed by combining a verb stem with the appropriate subject-verb agreement clitic, based on the person (1/2/3) and number (singular/plural) of the subject. This pronominal agreement clitic attaches to the verb in the imperfective aspect and to a preverbal element in the perfective and perfect aspects. The forms of these agreement clitics are given in §3.2. The infinitive stem is only used for subordinate clauses, so it generally does not occur with pronominal agreement clitics. Table 1.6 below shows the conjugations of the verb *χig* ‘eat’.

Table 1.6 Conjugations of *χig* ‘eat’

	IPFV	PFV	PRF
<b>Clitic:</b>	on verb	preverbal	preverbal
<b>1SG</b>	<i>waz χor = am</i> ‘I (will) eat.’	<i>waz = am χuug</i> ‘I ate.’	<i>waz = am χuuydz</i> ‘I have eaten.’
<b>2SG</b>	<i>təw χor = ∅</i> ‘You (will) eat.’	<i>təw = at χuug</i> ‘You ate.’	<i>təw = at χuuydz</i> ‘You have eaten.’
<b>3SG</b>	<i>ju χird</i> ‘S/he (will) eat.’	<i>ju = ∅ χuug</i> ‘S/he ate.’	<i>ju = ∅ χuuydz</i> ‘S/he has eaten.’
<b>1PL</b>	<i>maε χor = an</i> ‘We (will) eat.’	<i>maε = an χuug</i> ‘We ate.’	<i>maε = an χuuydz</i> ‘We have eaten.’
<b>2PL</b>	<i>tamaε χor = it</i> ‘You(pl) (will) eat.’	<i>tamaε = af χuug</i> ‘You(pl) ate.’	<i>tamaε = af χuuydz</i> ‘You(pl) have eaten.’
<b>3PL</b>	<i>wod χor = in</i> ‘They (will) eat.’	<i>wod = af χuug</i> ‘They ate.’	<i>wod = af χuuydz</i> ‘They have eaten.’

Examples (1.10) - (1.19) illustrate how the five verb stems of *χig* ‘eat’ are combined with pronominal agreement clitics to form sentences. In the imperfective aspect, the imperfective stem, *χor*, has an imperfective clitic attached to it. (1.10) has the first person singular imperfective clitic, *= am*, and (1.11) has the second person plural imperfective clitic, *= it*.

(1.10) *waz xipik χor = am*  
1SG.NOM flatbread eat.IPFV = 1SG.IPFV  
‘I (will) eat flatbread.’

(1.11) *tamaε xipik χor = it*  
2PL.NOM flatbread eat.IPFV = 2PL.IPFV  
‘You(pl) (will) eat flatbread.’

(1.12) & (1.13) have the third-person singular imperfective verb stem, *χird*, and no overt agreement clitic, which is a feature of the imperfective aspect with a third person singular subject.

(1.12) *mu jaχ xipik χird*  
1SG.NNOM sister flatbread eat.3SG.IPFV  
‘My sister eats/will eat flatbread.’

- (1.13) *juu xipik χird*  
 3SG.NOM.DIST flatbread eat.3SG.IPFV  
 ‘He eats/will eat flatbread.’

The perfective aspect is formed with the perfective stem, *χug*, with the perfective clitic attached to a preverbal element. (1.14) has the first person plural perfective clitic, =*an*, and (1.15) has the third person plural perfective clitic, =*af*.

- (1.14) *maç = an ingum xipik χug*  
 1PL.NOM = 1PL.PFV just.now flatbread eat.PFV  
 ‘We ate flatbread just now.’

- (1.15) *doð = af ingum xipik χug*  
 3PL.NOM.PROX = 3PL.PFV just.now flatbread eat.PFV  
 ‘These people ate flatbread just now.’

The perfect aspect contains the perfect stem, *χuydz*, as well as the perfective clitic attached to a preverbal element. (1.16) has the second person singular perfective clitic, =*at*, and (1.17) has the third person plural perfective clitic, =*af*.

- (1.16) *təw = at xipik tagəw na χuydz*  
 2SG.NOM = 2SG.PFV flatbread at.all NEG eat.PRF  
 ‘You have not eaten any flatbread at all. (Evidential/New information)’

- (1.17) *woð = af xipik tagəw na χuydz*  
 3PL.NOM.DIST = 3PL.PFV flatbread at.all NEG eat.PRF  
 ‘They have not eaten any flatbread at all. (Evidential/New information)’

Finally, (1.18) & (1.19) use the infinitive stem, *χig*, which does not occur with an agreement clitic because it is within a subordinate clause.

- (1.18) *jad pugan χig = itçuz xipik*  
 3SG.NOM.PROX tomorrow eat.INF = REL flatbread  
 ‘This is flatbread that will be eaten tomorrow.’

- (1.19) *mu dil xipik χig*  
 1SG.NNOM heart flatbread eat.INF  
 ‘I want to eat flatbread.’

Causative verbs are formed through stem modification. They cannot be formed for all verbs, although many verbs do have a causative counterpart. Causative forms are not completely predictable because the vowel and/or consonant from the final syllable of the non-causative form is sometimes altered, but they are often recognizable as causatives because they generally end with /ond/. Causatives typically have the same form for infinitive, perfective, and third person singular imperfective stems, as they all end with /ond/. The imperfective stem does not have a final /d/ and the perfect stem always ends in /dz/. Table 1.7 presents some causatives that are commonly used, along with the corresponding non-causative verb.

Table 1.7 Examples of causative verbs

	<b>IPFV</b>	<b>3SG.IPFV</b>	<b>PFV</b>	<b>PRF</b>	<b>INF</b>
‘lie’	<i>alos</i>	<i>alost</i>	<i>alud</i>	<i>aludz</i>	<i>alid</i>
‘lie.CAUS’	<i>alazon</i>	<i>alazond</i>	<i>alazond</i>	<i>alazondz</i>	<i>alazond</i>
‘reach’	<i>frops</i>	<i>fropst</i>	<i>fript</i>	<i>friptɕ</i>	<i>fript</i>
‘reach.CAUS’	<i>frapon</i>	<i>frapond</i>	<i>frapond</i>	<i>frapondz</i>	<i>frapond</i>
‘sleep’	<i>xufs</i>	<i>xufst</i>	<i>xuvd</i>	<i>xuvdz</i>	<i>xovd</i>
‘sleep.CAUS’	<i>xafson</i>	<i>xafsond</i>	<i>xafsond</i>	<i>xafsondz</i>	<i>xafsond</i>
‘read’	<i>xuj</i>	<i>xujd</i>	<i>xojd</i>	<i>xojdz</i>	<i>xojd</i>
‘read.CAUS’	<i>xajon</i>	<i>xajond</i>	<i>xajond</i>	<i>xajondz</i>	<i>xajond</i>
‘eat’	<i>ɣor</i>	<i>ɣird</i>	<i>ɣug</i>	<i>ɣuydz</i>	<i>ɣig</i>
‘eat.CAUS’	<i>ɣuron</i>	<i>ɣurond</i>	<i>ɣurond</i>	<i>ɣurondz</i>	<i>ɣurond</i>
‘cry’	<i>nəw</i>	<i>nəwd</i>	<i>niwd</i>	<i>niwdz</i>	<i>niwd</i>
‘cry.CAUS’	<i>nawon</i>	<i>nawond</i>	<i>nawond</i>	<i>nawondz</i>	<i>nawond</i>
‘burn’	<i>θəw</i>	<i>θəwd</i>	<i>θud</i>	<i>θɛðdz</i>	<i>θid</i>
‘burn.CAUS’	<i>θawon</i>	<i>θawond</i>	<i>θawond</i>	<i>θawondz</i>	<i>θawond</i>
‘move’	<i>dzumb</i>	<i>dzumbd</i>	<i>dzumbd</i>	<i>dzumbdz</i>	<i>dzumbd</i>
‘move.CAUS’	<i>dzumbon</i>	<i>dzumbond</i>	<i>dzumbond</i>	<i>dzumbondz</i>	<i>dzumbond</i>

The following pairs of sentences contrast how causatives and non-causatives are used. The subject of a non-causative verb becomes the direct object (as in (1.20b) & (1.21b)) or indirect object (as in (1.22b)) of a causative verb, and the causative verb takes an additional argument as its subject:

- (1.20) a. *jad = ik*                                 *uz*     *nəwd*  
           3SG.NOM.PROX = DUR     again     cry.3SG.IPFV  
           ‘This one is crying again.’



- b. *tɕoj*      *a=wi*                      *nawond*  
 who.NOM ACC=3SG.NNOM.PROX cry.CAUS.PFV  
 ‘Who caused her to cry?’
- (1.21) a. *tamaç*    *dɕald* *χu*              *leq*      *pamedz=it*,  
 2PL.NOM fast REFL.NNOM clothing wear.IPFV=2PL.IPFV  
  
*tamoq* *χor=it*  
 food eat.IPFV=2PL.IPFV  
 ‘Put your(pl) clothes on quickly and eat.’
- b. *waz=am*                      *a=tamaç*              *ðes sul*  
 1SG.NOM=1SG.PFV ACC=2PL.NNOM ten year  
  
*χurond*              *pamedzond*  
 eat.CAUS.PFV wear.CAUS.PFV  
 ‘I have fed you and clothed you for ten years.’
- (1.22) a. *mu*              *bob*              *xats*    *bruxt*  
 1SG.NNOM grandfather water drink.PFV  
 ‘My grandfather drank water.’
- b. *waz=am*                      *χu*              *bob=ir*              *xats*  
 1SG.NOM=1SG.PFV REFL.NNOM grandfather=DAT water  
  
*brazond*  
 drink.CAUS.PFV  
 ‘I fed my grandfather water.’

#### 1.4.2.2 Compound verbs

Verbs are not an open lexical class in that new verb stem paradigms are not regularly added to the lexicon. Instead, Sarikoli uses a large number nouns and adjectives in combination with other existing verbs to express verbal meanings. *tɕejg* ‘do’, *set* ‘become’, *ðod* ‘give’, and *χig* ‘eat’ are among the most common verbs to be used in compound verbs. Table 1.8 lists examples of frequently-used compound verbs.

Table 1.8 Examples of compound verbs

Compound verb	Components	Meaning
<i>iç tcejg</i>	cold + do	‘be cold’
<i>tej tcejg</i>	wedding + do	‘marry’
<i>hat tcejg</i>	open + do	‘open’
<i>gap tcejg</i>	word + do	‘talk’
<i>jordam tcejg</i>	help + do	‘help’
<i>ub set</i>	melt + become	‘melt’
<i>ago set</i>	awake + become	‘wake up’
<i>aluk set</i>	tired + become	‘get tired’
<i>çafø set</i>	upset + become	‘get upset’
<i>azmud set</i>	born + become	‘be born’
<i>mut ðod</i>	fist + give	‘punch’
<i>lutç ðod</i>	kick + give	‘kick’
<i>para ðod</i>	sell + give	‘sell’
<i>fand ðod</i>	false + give	‘lie’
<i>dzεq ðod</i>	squat + give	‘squat’
<i>leç çig</i>	bump + eat	‘bump into’
<i>çam çig</i>	bend + eat	‘bend’
<i>ditçur çig</i>	encounter + eat	‘encounter’
<i>bazun çig</i>	wither + eat	‘wither’
<i>rawudž çig</i>	thriving + eat	‘thrive’

The nominal (noun or adjective) element of a compound verb does not function as the direct object of the verb, as it is part of the verb. This is exemplified in (1.23) – (1.25), in which compound verbs occur with accusative arguments. Other compound verbs, as shown in (1.26) – (1.28), are used intransitively and do not take accusative arguments. Morphologically, the nominal elements of compound verbs are distinct from both verbs and NP arguments. Unlike verbs, they do not occur in five different stems and do not host pronominal agreement clitics in the imperfective aspect. Whereas NP arguments are usually marked with function-marking clitics or adpositions, the nominal element of a compound verb is not. It is part of the compound verb but does not take inflections that are limited to verbs or nouns. But it is a separate word which can anchor enclitics, as in (1.24).

- (1.23) *farzana a=sandεq hat tçəwg*  
 Farzana ACC=box open do.PFV  
 ‘Farzana opened the box.’

- (1.24) *a=maç=at* *fand ðudz*  
 ACC = 1PL.NNOM = 2SG.PFV false give.PRF  
 ‘You have lied to us. (Evidential/New information)’
- (1.25) *χu radzen tsava para ðo=am*  
 REFL.NNOM daughter how sell give.IPFV = 1SG.IPFV  
 ‘How could I sell my own daughter?’
- (1.26) *mur=af wayti ago sut*  
 today = 3PL.PFV early awake become.PFV  
 ‘They woke up early today.’
- (1.27) *kalo-χejl=af mas iç tçəwydz*  
 sheep-PL.NOM = 3PL.PFV also cold do.PRF  
 ‘The sheep also got cold.’
- (1.28) *waz=am i suat dzeq ðud*  
 1SG.NOM = 1SG.PFV one hour squat give.PFV  
 ‘I have squatted for one hour.’

### 1.4.2.3 Clitics

In this grammar, clitics are defined as grammatically separate morphemes that are phonologically dependent on another word (Dixon & Aikhenvald 2003). A clitic is attached to its host after phonological rules have been applied, so it usually does not receive primary stress even if it is the final syllable of a phonological word. Unlike affixes, which are more restricted in their choice of host, clitics can attach to words belonging to multiple lexical classes, or to entire clauses. There are eight categories of clitics in Sarikoli, as presented in Table 1.9:

Table 1.9 Categories of clitics

Category	Function	Members
Agreement	Shows person and number of the subject; indicates aspect through form (perfective vs. imperfective form) and placement (attaching to the verb vs. preverbal element), in combination with the type of verb stem	Introduced in §3.2

<b>Category</b>	<b>Function</b>	<b>Members</b>
<b>Function-marking</b>	Marks the clausal function of an NP	<i>a</i> = (ACC), = <i>ir</i> / = <i>ri</i> (DAT)
<b>Aspectual</b>	Indicates lesser aspects in combination with verb stem and pronominal agreement clitic	= <i>ik</i> (DUR)
<b>Subordinating</b>	Forms nominalized complement clauses and reason adverbial clauses	= <i>i</i> (SC)
<b>Relativizer</b>	Forms relative clauses	= <i>endz</i> (REL), = <i>itɕuz</i> (REL)
<b>Demonstrative</b>	Marks anaphora, cataphora, and physical distance	<i>k(i)</i> = (ANA), <i>m(i)</i> = (CATA)
<b>Emphatic</b>	Attaches to an emphasized constituent	= <i>aθ</i> (EMP)
<b>Interrogative</b>	Attaches to a sentence or particular constituent and forms polar questions	= <i>o</i> (Q)

#### 1.4.2.4 Aspect

Major aspects—perfective, imperfective, and perfect—are indicated through a combination of the type of verb stem and the form and placement of pronominal agreement clitics, as shown in examples (1.10) - (1.18). Besides the major aspects, lesser aspects are formed by adding the durative enclitic *=ik* or cessative suffix *-it*.

*=ik*, which Palmer analyzes as a durative marker (2016:106), is used with situations that are ongoing, occurring, or coming about. It may attach to the verb but more commonly attaches to a preverbal element that is not an adposition or adnominal modifier. It is a key element in a number of different constructions, as shown in the following examples.

It is used with the imperfective stem for present continuous aspect:

- (1.29)    *m-ono = ik*                      *tamoq kaxt*  
           1SG.NNOM-mother = DUR    food    do.3SG.IPFV  
           ‘My mother is making food.’

- (1.30) *malum-χejl = ik a = tamaç tços = in*  
 teacher-PL.NOM = DUR ACC = 2PL.NNOM watch.IPFV = 3PL.IPFV  
 ‘The teachers are waiting for you(pl).’

It is used with the perfective stem for past habitual aspect, which involve iterative events that have occurred in the past:

- (1.31) *paləw = am = ik χuug*  
 pilaf = 1SG.PFV = DUR eat.PFV  
 ‘I have eaten pilaf (multiple times).’
- (1.32) *malum pa tçed = am = ik dejd*  
 teacher LOC house = 1SG.PFV = DUR enter.PFV  
 ‘I have gone to the teacher’s house (multiple times).’

It is used with the perfect stem and cessative suffix *-it* in counterfactual adverbial clauses:

- (1.33) *tamaç = af uz i maθ = ik tsa nahuçtç-it*  
 2PL.NOM = 2PL.PFV again one day = DUR COND sit.PRF-CESS
- maç = an = ik tup amad ar tej*  
 1PL.NOM = 1PL.PFV = DUR group Amad LOC wedding
- σεδδç-it*  
 become.PRF-CESS  
 ‘If you(pl) had stayed one more day, we would have all gone to Amad’s wedding together.’
- (1.34) *mu-an hansu ziv kasp vid tçi dzuj*  
 1SG.NNOM-GEN Han tongue major be.INF LOC place
- ingles ziv kasp = ik tsa veδδç-it*  
 English tongue major = DUR COND be.PRF-CESS
- waz = am = ik az ta ingles ziv*  
 1SG.NOM = 1SG.PFV = DUR ABL 2SG.NNOM English tongue
- χumand σεδδç-it*  
 learn become.PRF-CESS  
 ‘If my major had been English instead of Mandarin, I would have learned English from you.’

It is used with the perfective stem in temporal adverbial clauses:

- (1.35) *az dars = am = ik                      xovd                      tu = ri*  
 ABL lesson = 1SG.PFV = DUR go.down.PFV 2SG.NNOM = DAT

*lev = am*

say.IPFV = 1SG.IPFV

‘I will tell you when I have gotten out of class.’

- (1.36) *jad                      kinu = ik                      adu sut                      pa buzur*  
 3SG.NOM.PROX movie = DUR finish become.PFV LOC bazaar

*so = an*

become.IPFV = 1PL.IPFV

‘We will go to the bazaar once this movie is finished.’

Finally, it is used with the imperfective stem for reporting direct speech:

- (1.37) *na səwd = ik                      levd*  
 NEG become.3SG.IPFV = DUR say.3SG.IPFV

‘He is saying, “It is not okay”.’

- (1.38) *ta                      dil = ik                      lev = in*  
 2SG.NNOM heart = DUR say.IPFV = 3PL.IPFV

‘They are saying, “It is up to you”.’

The cessative suffix *-it* attaches to the perfect stem of verbs to form the pluperfect aspect, which is used for situations which “have been completed at a past time reference” and whose resultant state is also in the past (Palmer 2016:103). It is also used in counterfactual adverbial clauses, as in (1.33) & (1.34). The following examples contain sentences in the pluperfect aspect:

- (1.39) *i maθ = am                      a = wi                      wandz-it*  
 one day = 1SG.PFV ACC = 3SG.NNOM.DIST see.PRF-CESS

‘I saw(/had seen) her the other day.’

- (1.40) *wođ = af                      parus                      i                      xots mu = ri*  
 3PL.NOM.DIST = 3PL.PFV last.year one girl 1SG.NNOM = DAT

*buxtç-it*

send.PRF-CESS

‘They sent me a girl last year.’

- (1.41) *waz nardzɛd alo taw=at mu tçi*  
 1SG.NOM pass.INF TEMP 2SG.NOM=2SG.PFV 1SG.NNOM LOC  
*kol çindz-it*  
 head laugh.PRF-CESS  
 ‘When I passed by, you laughed at me.’
- (1.42) *nur kampir a=mu pa tçed levdz-it*  
 today old.lady ACC=1SG.NNOM LOC house say.PRF-CESS  
*tçoj broxt=ir*  
 tea drink.INF=DAT  
 ‘Today the old lady invited me to her house for tea.’
- (1.43) *waz=am utç tur sɛðdz-it, pa*  
 1SG.NOM=1SG.PFV very thirsty become.PRF-CESS LOC  
*tçed=am dejt, tazo xats=am bruxt*  
 house=1SG.PFV enter.PFV very water=1SG.PFV drink.PFV  
 ‘I got very thirsty, went into the house, and drank a lot of water.’
- (1.44) *eej ju tçurik mu=ri levdz-it iko*  
 INTJ 3SG.NOM.DIST man 1SG.NNOM=DAT say.PRF-CESS SC  
*ditçur çuydz=endz a=rusq tagəw mo*  
 encounter eat.PRF=REL ACC=portion ever PROH  
*patəw*  
 throw.IPFV  
 ‘Oh yeah, that man told me, “Never throw away an offered portion that you come across”.’
- (1.45) *ha ðod=ir=ik vəw buðon~mabuðon qati*  
 INTJ give.INF=DAT=DUR be.IPFV saddle~RDP COM  
*ðo, ingum=at mu pa gap na*  
 give.IPFV just.now=2SG.PFV 1SG.NNOM LOC word NEG  
*tçimbdz-it*  
 be.willing.PRF-CESS  
 ‘Ah, if you are going to give it to me, give me the saddle as well, since you were unwilling just a moment ago.’

