



January 2014

Subordination In Sarikoli

Deborah Kim

Follow this and additional works at: <https://commons.und.edu/theses>

Recommended Citation

Kim, Deborah, "Subordination In Sarikoli" (2014). *Theses and Dissertations*. 1672.
<https://commons.und.edu/theses/1672>

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Senior Projects at UND Scholarly Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of UND Scholarly Commons. For more information, please contact zeinebyousif@library.und.edu.

SUBORDINATION IN SARIKOLI

by

Deborah Kim
Bachelor of Arts, Trinity Western University, 2013

A Thesis
Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Arts

Grand Forks, North Dakota
August
2014

This thesis, submitted by Deborah Kim in partial fulfillment of the requirements for the Degree of Master of Arts from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

Robert Fried, Chair

Mark Karan

Adam Baker

This thesis meets the standards for appearance, conforms to the style and format requirements of the School of Graduate Studies of the University of North Dakota, and is hereby approved.

Wayne Swisher
Dean of the School of Graduate Studies

Date

PERMISSION

Title Subordination in Sarikoli
Department Linguistics
Degree Master of Arts

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in his absence, by the chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this thesis or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Deborah Kim

June 11, 2014

CONTENTS

Abbreviations	vii
Acknowledgements	viii
Abstract	ix
CHAPTER	
1. Introduction	1
1.1. The Pamir languages and Sarikoli	1
1.2. Previous linguistic work on Sarikoli	3
1.3. Overview of thesis	4
2. Preliminary notes on Sarikoli grammar	6
2.1. Basic constituent order	6
2.2. Grammatical relations	7
2.3. Verbal system	10
3. Relative Clauses	15
3.1. The common argument	15
3.2. The RC	20
3.2.1. Placement of the RC	20
3.2.2. Grammatical marking within the RC	21
3.2.3. Determiners, adjectives, and adpositions within the matrix clause	22
3.2.4. Non-core elements within the RC	24
3.2.5. Negation in RCs	25

3.3. Types of Sarikoli RCs	25
3.3.1. Finite RCs: the relativizer = <i>dzendz</i>	26
3.3.2. Infinitival RCs: the relativizer = <i>itfuz</i>	28
3.3.3. Headless RCs	31
3.3.4. The genitive marker <i>-an</i>	33
3.3.5. Unmarked infinitival RCs	33
3.3.6. Non-restrictive RCs	34
3.3.7. Adjectivizers: <i>-endz</i> , <i>-jendz</i> , and <i>-nendz</i>	35
3.4. Comparison with related languages	37
3.4.1. Persian & Tajik	38
3.4.2. Shughni	39
3.4.3. Rushani	40
3.4.4. Wakhi	40
3.4.5. Conclusion	41
4. Complement Clauses	43
4.1. Types of Sarkoli CCs	43
4.1.1. Nominalized complement	44
4.1.1.1. The <i>-i</i> suffix	44
4.1.1.2. Nominalized complement (with the <i>-i</i> suffix)	46
4.1.2. Infinitival complement	54
4.1.3. Unmarked finite complement (pre-verbal)	62
4.1.4. Marked finite complement (post-verbal)	63
4.2. Comparison with related languages	65
4.2.1. Persian & Tajik	65

4.2.2. Shughni	66
4.2.3. Rushani	66
4.2.4. Wakhi	66
4.2.5. Conclusion	67
5. Adverbial Clauses	68
5.1. Types of Sarikoli ACs	69
5.1.1. Time	69
5.1.2. Reason	71
5.1.3. Purpose	73
5.1.4. Condition	76
5.1.5. Concession	78
5.1.6. Means and Simultaneity	80
5.1.7. Substitution	82
5.1.8. Location	83
5.1.9. Manner	85
6. Conclusion	88
6.1. Summary	88
6.2. Areas for further research	89
References	91

ABBREVIATIONS

1	first person	NMLZ	nominalizer
2	second person	NPST	nonpast
3	third person	ORD	ordinal number
ACC	accusative	PL	plural
AC	adverbial clause	POSS	possessive
ADJ	adjectivizer	PRF	perfect
ADV	adverbial	PROH	prohibitive
BEN	benefactive	PST	past
CC	complement clause	Q	question particle
COMP	complementizer	RC	relative clause
COND	conditional	RECP	reciprocal
DAT	dative	REL	relativizer
F	feminine	RES	restrictive relative clause
GEN	genitive	SG	singular
INF	infinitive	SC	subordinating conjunction
NOM	nominative	TEMP	temporal
NEG	negation		

ACKNOWLEDGEMENTS

My dear committee members: thank you so much for helping this thesis come true. Dr. Robb Fried, you have been the best committee chair a linguistics student could ask for. Thank you for your genuine support and willingness to help in every way, and for your linguistic wisdom. Dr. Mark Karan, thank you for your encouragement and confidence in me. Dr. Adam Baker, thank you for bringing in your expertise in Iranian and Pamir languages.

My dear Sarikuy friends: I could not have done anything without you. Thank you so much for your friendship, hospitality, and patience as I learned about your language. You will always have a very special place in my heart. Neikramon Ibrukhim, I cannot count how many questions I've asked you as I wrote this thesis; you have been so patient and helpful every single time. I admire your passion for your people and language. I feel very blessed to be your friend. Neeyuft Latfik, you are an amazing sister. Thank you for being my friend and offering to teach me your language. Thank you for introducing me to your wonderful family and opening up your home. Thank you for teaching me the Eagle Dance. I love your people.

My dear professors at CanIL, GIAL, and SIL-UND: thank you for inspiring me to fall in love with linguistics. Special thanks to Dr. Sean Allison, Dr. Doug Trick, Dr. Paul Kroeger, and Karl Anderbeck for encouraging me to pursue my interest in Sarikoli.

Dr. Pam Arlund, thank you for your generosity and hospitality. You have shared so much about the Sarikuy people and language. I, too, will share my knowledge and stories generously with others.

Moss Doerksen, thank you so much for making the Pamir map. You are very gifted.

And my dear family: thank you for loving me! Thank you for guiding me, supporting me, accompanying me on data collection trips, making me delicious food, and discussing fascinating things with me. I cannot imagine life without you.

ABSTRACT

Sarikoli [srh] is an Iranian language spoken in Tashkurgan Tajik Autonomous County in northwest China. This thesis describes three types of subordinate clauses in Sarikoli: 1) relative clauses, 2) complement clauses, and 3) adverbial clauses. The relative clause and complement clause structures are briefly compared with those found in related Iranian and Pamir languages (Persian, Tajik, Shughni, Rushani, and Wakhi).

Sarikoli relative clauses are placed before the head noun. Common nouns, proper nouns, demonstratives, and genetic terms may be relativized, but pronouns are generally not relativized. A wide range of syntactic functions are allowed for the common argument in both the relative clause and the matrix clause, including A, S, O, and oblique roles. The two main relativizers used for Sarikoli RCs are *=dʒendʒ* and *=itfuz*. The *=dʒendʒ* relativizer is used for finite RCs, while *=itfuz* is used for non-finite clauses (including future events with an infinitive verb). Other ways of forming RCs include headless RCs, unmarked RCs, and using the genitive marker *-an*.

Sarikoli has at least two types of finite complement clauses and two types of non-finite complementation strategies. In the nominalized complement, the nominalizer *-i* attaches to the infinitive stem of the verb. The infinitival complement also contains the infinitive stem of the verb, but is unmarked. The pre-verbal finite complement clause is unmarked and contains a finite verb stem and a subject-verb agreement clitic. The post-verbal finite complement clause is placed after the matrix clause verb; it is introduced by the subordinating conjunction *iko* and contains a finite stem of the verb and a subject-verb agreement clitic.

Adverbial clauses are marked by various subordinating morphemes, including *tsa* 'if', *qati* 'with', *alo* 'when', *avon* 'for', *az* 'from', and the dative marker *=ir*, which generally occur at the end of the adverbial clause. Most types of adverbial clauses are non-finite, containing verbs in their infinitive stem and lacking subject-verb agreement clitics. Like

regular adverbs, adverbial clauses usually precede the entire matrix clause or immediately follow the subject. Sarikoli has structurally-distinctive adverbial clause constructions for expressing time, reason, purpose, condition, concession, means and simultaneity, and substitution.

CHAPTER 1

Introduction

Sarikoli [srh] is an Iranian language spoken in the Xinjiang Uyghur Autonomous Region of northwest China. It has received little attention in Iranian linguistics, and very few researchers have published anything about Sarikoli grammar. This thesis provides a grammatical description of relative clauses, complement clauses, and adverbial clauses in Sarikoli. Some notable differences between subordinate clause structures in Sarikoli and those in other Iranian and Pamir languages are also highlighted.

1.1 The Pamir languages and Sarikoli

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 language family. The Pamir languages are located in the far eastern edge of the area where Iranian languages are distributed.

Pamir languages are spread across the Pamir Mountains in four Central Asian countries: Tajikistan, Afghanistan, Pakistan, and China. Shughni, Rushani, and Yazgulyam are mainly spoken in Tajikistan; Sanglechi, Ishkashimi, and Munji are mainly spoken in Afghanistan; Yidgha is mainly spoken in Pakistan; Sarikoli is only spoken in China; and Wakhi is spoken in all four countries (Payne 1989; Lewis 2013). Figure 1 indicates where these languages are located.



Figure 1: Map of Pamir languages (created by Moss Doerksen; data from multi-tree.org)

Some linguists divide the Pamir language family into northern and southern subgroups (Edelman & Dodykhudoeva 2009a). According to Dodykhudoeva, Shughni, Rushani, Yazgulyam, and Sarikoli belong to the Shughni-Rushani subgroup, and are the only genetically closely related languages among the Pamir languages (2004:149). Geographically, these languages are also located closely together in the northern part of the Pamir Mountains, although Sarikoli is somewhat isolated. The other Pamir languages—Wakhi, Ishkashimi, Sanglechi, Munji, and Yidgha—belong to the South Pamir subgroup and are not closely related genetically (2004:149).

Sarikoli [srh] is the easternmost of the extant Iranian languages (Payne 1989:147), and is only spoken in China. The Sarikoli-speaking community is located on the westernmost edge of the Xinjiang Uyghur Autonomous Region in northwest China. The majority of Sarikoli speakers live in Tashkurgan Tajik Autonomous County, which borders Tajikistan, Afghanistan, and Pakistan. According to Dodykhudoeva, Sarikoli speakers first migrated to Xinjiang several centuries ago and again in 1911, following the Sarez-Pamir earthquake that triggered numerous landslides and destroyed villages in eastern Tajikistan (2007:69).

The exact number of Sarikoli speakers is unknown, because Sarikoli is not officially a distinct ethnic minority group in China. China has lumped together its two Pamir people groups, Sarikoli and Wakhi, as one of its 55 official minority groups, under the ethnonym “塔吉克族” (tajikeyu), which means, “Tajik people”. According to the Sixth National Population Census of the People’s Republic of China that was conducted by The National Bureau of Statistics of China (2010), there are 51,069 Tajiks, of which the majority are Sarikoli speakers. Sarikoli speakers live in all 12 of the main villages of Tashkurgan Tajik Autonomous County, whereas the Wakhi speakers are mostly concentrated in one village called Dafdar. Because Sarikoli and Wakhi are mutually unintelligible, Sarikoli is used as the language of wider communication among the “Tajiks” of China (Arlund 2006:4). The Tajiks of China are surrounded not by other Iranian languages, but by Turkic languages and Mandarin Chinese. Uyghur, a Turkic language, is the lingua franca among the minority peoples of Xinjiang, and Mandarin is the national language.

Sarikoli has not had a writing system until very recently, so it is primarily used as a language of oral communication. See below for more information on the orthography. The data presented in this thesis are transcribed in the International Phonetic Alphabet.

1.2 Previous linguistic work on Sarikoli

Information about Sarikoli has been included in some materials about the Pamir languages as a whole, but very little has been written specifically about Sarikoli. According to Arlund, Sarikoli has been “the most isolated and most understudied” of the Pamir languages because it is restricted to a remote area in the western edge of China (2006:6). Very few linguists have analyzed Sarikoli based on their own data. Researchers who have actually collected and studied their own Sarikoli data include T. N. Pakhalina, Gao Erqiang, and Pam Arlund. While other authors sometimes mention Sarikoli in their works, they rely on the data collected by these researchers as their primary sources (2006:9).

The first English account of Sarikoli was made by an English linguist named Robert B. Shaw in 1876 (Arlund 2006). Nearly a century later, T. N. Pakhalina, a Russian scholar, carried out an extensive study of Sarikoli and other Pamir languages, under the leadership of Ivan Ivanovich Zarubin of the Soviet Academy of the Social Sciences. Although Zarubin

did not conduct in-depth research on Sarikoli, he was one of the first Pamir researchers who identified speakers of some Shughni dialects with Sarikoli speakers who migrated to Tashkurgan, Xinjiang (Dodykhudoeva 2007:69). Pakhalina did most of her fieldwork in the 1950s and published her work in the 1960s and 1970s (Arlund 2006). She wrote descriptions of Sarikoli in Russian and published some texts transcribed using the Russian Iranologist transcription system (Pakhalina 1960 & 1966).

In 1985, the Chinese scholar Gao Erqiang published *Tajikeyu Jianzhi* [*Concise grammar of Tajik*], a “Tajik” volume to the Chinese series of books that describe minority languages, in which he describes Sarikoli and Wakhi, the two Pamir languages spoken in China. In the 1990s and the 2000s, Pam Arlund conducted in-depth research on Sarikoli. She wrote her dissertation on Sarikoli diphthongs (Arlund 2006), as well as doing grammatical analysis. Together with Neikramon Ibrukhim, she wrote *A Chinese Tajik reader: An introduction to Sarikoy (Sarikol) Tajik* (2013) for people learning Sarikoli as a foreign language. Edelman and Dodykhudoeva wrote a brief typological overview of the Pamir languages (Edelman & Dodykhudoeva 2009a), which includes some information about Sarikoli. Most recently, Neikramon Ibrukhim, a native speaker of Sarikoli, has created the first Sarikoli writing system. He also wrote the first Sarikoli primer, *Chinese Tajik Alphabet* (Ibrukhim 2012), in which he introduces the Sarikoli writing system along with some short texts.

Among the previous linguistics works, there has been no description of subordination in Sarikoli in English. The contribution of this thesis is to describe the relative clauses, complement clauses, and adverbial clauses found in Sarikoli.

1.3 Overview of thesis

Sarikoli is rich in subordinate clauses; it often makes use of the possibilities for subordinating one proposition to another. In this thesis, I describe relative clauses (Chapter 3), complement clauses (chapter 4), and adverbial clauses (chapter 5) in Sarikoli. For relative clauses and complement clauses, at the end of the respective chapters, the Sarikoli constructions are compared with those of some related Iranian and Pamir languages. Chapter 2 introduces some basic aspects of Sarikoli grammar, as a foundation for discussing the

subordinate clauses in the chapters that follow. Chapter 6 includes a summary of my findings and directions for further research.

I rely on “basic linguistic theory” (Dixon 1997, Dryer 2006) as the theoretical framework for this thesis, rather than using terms from any one explanatory framework. My aim is to describe Sarikoli on its own terms, so that this thesis may provide an adequate description of Sarikoli “in all of its complexity and idiosyncrasy” (Dryer 2006:10).

The analysis presented in this thesis is predominantly based on data from my own research in Tashkurgan Tajik Autonomous Region and Urumqi, Xinjiang. They include oral texts that were recorded and transcribed, elicited sentences, and lessons learned through informal conversations I had with Sarikoli speakers. My fieldwork totals up to seven weeks in Tashkurgan Tajik Autonomous Region and one month in Urumqi. In addition, questions were asked and answered through correspondence after I had left Xinjiang. The examples in this thesis are drawn from approximately 400 elicited sentences, as well as 300 lines from 13 oral texts that have been interlinearized (Kim 2013).

I have also relied on data from another source, which is *A Chinese Tajik reader: An introduction to Sarikoy (Sarikol) Tajik* (Arlund & Ibrukhim 2013), a textbook for English speakers learning Sarikoli. It contains numerous short Sarikoli texts and vocabulary words.

Example sentences from other sources are cited in parentheses. Examples transcribed in the Russian Iranologist Transcription were re-written in IPA, for greater accessibility. Data from my own research are in broad phonetic transcription.

CHAPTER 2

Preliminary notes on Sarikoli grammar

This chapter aims to provide a brief overview of Sarikoli grammar before the more in-depth study of Sarikoli subordinate clauses in the remaining chapters. It covers the basic constituent order, how grammatical relations are marked, and how various verbs stems and tenses work.

2.1 Basic constituent order

Like most Iranian languages, Sarikoli is a verb-final language with the SOV basic constituent order:

- (1) *zulfia poj furd*
Zulfia yogurt slurp.PST
'Zulfia ate yogurt.'

Sarikoli is also a null-subject language, so an explicit subject is not required for an independent clause. As shown in the following sentences, the subject can be omitted if it can be inferred from the context.

- (2) *(wođ) xipik ka = jin*
3PL.NOM flatbread do.NPST = 3SG.NPST
'They are making flatbread.'
- (3) *waz = am noj χid*
1SG.NOM = 1SG.PST eagle.flute play.PST
'I played the eagle flute.'
- (4) *noj = am χid*
eagle.flute = 1SG.PST play.PST
'I played the eagle flute.'

In sentences with non-verbal predicates, copulae are not required. The subject may simply be placed next to the non-verbal predicate:

- (5) *jad m = oto*
 this 1SG.POSS = father
 'This is my father.'
- (6) *tiznejf mu dijur*
 Tiznap 1SG.POSS hometown
 'Tiznap is my hometown.'

2.2 Grammatical relations

Sarikoli is a nominative-accusative language. In Sarikoli, grammatical relations are expressed through word order, case marking, and person marking.

As mentioned above, the SOV word order is one way in which grammatical relations are indicated. In (1), 'Zulfia' is the subject and 'yogurt' is the object because 'Zulfia' precedes 'yogurt'.

Case marking is also used to some extent. Personal pronouns operate on a case system based on person and number. The following table shows the set of personal pronouns used in Sarikoli (Payne 1989:432, Arlund & Ibrukhim 2013:15).¹

Table 1. Personal Pronouns

	Singular	Plural
1.N	waz	
1.A	mu	maʃ
2.N	təw	
2.A	ta	tamaʃ
3.N	ju	woð
3.A	wi	wəf

¹ Note that the possessive pronouns are identical to the accusative pronouns below. If the possessor is the same referent as the subject of the sentence, *χu* 'own' is used as the possessive pronoun. Both common-noun possessors and possessive-pronoun possessors may optionally take the genitive marker, *-(j)an*. Whether the possessor and possessed item are simply juxtaposed or linked together with the genitive marker, the possessor always precedes the head noun, the possessed item.

For common nouns, not all cases are marked. The nominative case is unmarked unless the subject is plural. The accusative case is marked with the *a =* proclitic, but Sarikoli has differential object marking: definite objects are obligatorily marked with the accusative case marker, and indefinite objects may be marked or unmarked. As shown in (9), the differential object marker attaches to accusative pronouns as well.

- (7) *waz barqo vrej = am*
 1SG.NOM lamb find.NPST = 1SG.NPST
 'I will find a lamb.'
- (8) *waz a = barqo vrej = am*
 1SG.NOM ACC = lamb find.NPST = 1SG.NPST
 'I will find the lamb.'
- (9) *waz a = ta tfardz wejn = am*
 1SG.NOM ACC = 2SG.ACC good see.NPST = 1SG.NPST
 'I like you.'

The dative case is always marked with the *=ir/=ri* enclitic. The form of this case marker is phonologically conditioned by the final segment of the word it attaches to. Consonant-final words take *=ir* and vowel-final words take *=ri*. If the indirect object is a pronoun, the dative case marker attaches to the accusative form of the relevant pronoun.

- (10) *alima kud = ir tamoq ðud*
 Alima dog = DAT food give.PST
 'Alima gave food to the dog.'
- (11) *alima batfo = ri tamoq ðud*
 Alima child = DAT food give.PST
 'Alima gave food to the child.'
- (12) *mu = ri jurdam ka*
 1SG.ACC = DAT help do.NPST
 'Help me.'

The personal interrogative pronouns distinguish case as well. The nominative form is *tfoj* ‘who’ and the accusative form is *tçi* ‘whom’. The accusative form with the dative marker =*ri* is used for the dative interrogative pronoun.

- (13) *tfoj* *a=ta* *ðud*
 who.NOM ACC=2SG.ACC hit.PST
 ‘Who hit you?’
- (14) *təw=at* *a=tçi* *ðud*
 2SG.NOM=2SG.PST ACC=who.ACC hit.PST
 ‘Whom did you hit?’
- (15) *təw=at* *tçi=ri* *ðud*
 2SG.NOM=2SG.PST who.ACC=DAT give.PST
 ‘Whom did you give it to?’

Case marking also interacts with the encoding of plurality. Sarikoli also has two different plural markers for nouns, corresponding to the nominative and accusative cases. The suffix *-xejl* is used for pluralizing nominative arguments, while *-(j)ef* is used for pluralizing accusative arguments:

- (16) *batfo-xejl=af* *jot*
 child-NOM.PL=3PL.PST come.PST
 ‘Children came.’
- (17) *a=batfo-jef* *jɛt* *laka=jit*
 ACC=child-ACC.PL come.INF let.NPST=2PL.NPST
 ‘Let the children come. (speaking to you-plural)’

Finally, grammatical relations are also expressed through obligatory person marking: the pronominal clitics show agreement between the subject and the verb. The form of the pronominal clitics agrees with the person and number of the subject; its form and placement agrees with the tense of the verb. These pronominal clitics are presented in the following table (Payne 1989:437). The [j] indicated in parentheses is inserted when the clitic attaches to a vowel-final stem.

Table 2. Pronominal Agreement Clitics

	Non-past singular	Non-past plural	Past Singular	Past Plural
1	= (j)am	= (j)an	= (j)am	= (j)an
2	= ∅	= (j)it	= (j)at	= (j)af
3	= t/d	= (j)in	= i/ = ∅	= (j)af

Arlund points out that the non-past third person singular verb stem is often irregular and needs to be memorized separately, even though they usually end in [t] or [d] (Arlund & Ibrukhim 2013:14). For this reason, the non-past third person singular verb stem will be glossed as its own stem, with the *-t/d* agreement clitic merged into it.

2.3 Verbal system

In Sarikoli, each verb has three finite stems, non-past, past, and perfect, as well as an infinitive stem. There are some regular verbs in which the formation of these verb stems is somewhat predictable, although the infinitive is not predictable. In these regular verbs, the past stem is formed by adding a *-t* or *-d* ending to the present stem (depending on the voice of the segment it attaches to), and the perfect stem is sometimes formed by changing those endings to *-tʃ* or *-dʒ* (Payne 1989:436). Sometimes the infinitive stem is identical to the past stem, but this is not always the case.

Table 3. Examples of Regular Verbs (verb paradigms from Arlund & Ibrukhim 2013; table compiled by me)

English gloss	NON-PAST	PAST	PERFECT	INFINITIVE
'say'	lev	levd	levdʒ	levd
'gather'	wix	wixt	wixtʃ	wixt
'use'	rafon	rafond	rafondʒ	rafond
'celebrate'	narzamb	narzambd	narzambdʒ	narzambd
'give'	ðo	ðud	ðudʒ	ðod
'sit'	niθ	nalust	nalustʃ	nalist
'get up'	indiz	indəwd	indəwdʒ	indejd

However, there are many irregular verbs in which the 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 three stems. Some irregular verbs are listed below.

Table 4. Examples of Irregular Verbs (verb paradigms from Arlund & Ibrukhim 2013; table compiled by me)

English gloss	NON-PAST	PAST	PERFECT	INFINITIVE
'do'	ka	tʃəwg	tʃəwɣ	tʃejg
'become'	so	sut	səðdʒ	sjet
'come'	joð	jot	iθ	jet
'play' (instrument)	χej	χid	χεð	χid
'turn around'	ʋirs	ʋejrd	ʋejr	ʋejrd
'sweep'	zdor	zdug	zduɣ	zdig
'sew'	insov	insuɪvd	insivd	insivd
'stand'	warofs	waruɪvd	waruɪvd	warvid

The following six examples illustrate how the four verb stems of 'sit' are used. (18) uses the non-past stem, *niθ*; (19) uses the past stem, *nalust*; (20) uses the perfect stem, *nalustf*; (21), (22), and (23) use the infinitive stem, *nalist*. Notice that the non-past, past, and perfect stems may be used as the main verb in a matrix clause, whereas the infinitive stem only occurs as an embedded verb within a relative clause (as in (21)), a complement clause (as in (22)), or an adverbial clause (as in (23)).

(18) *waz* *fitf* *niθ = am*
 1SG.NOM now sit.NPST = 1SG.NPST
 'I will sit now.'

(19) *waz = am* *ingum* *nalust*
 1SG.NOM = 1SG.PST just.now sit.PST
 'I sat just now.'

(20) *tʃardʒ = at* *nalustf = o*
 good = 2SG.PST sit.PRF = Q
 'Have you been sitting well? (for: Have you been doing well?)'

- (21) *jad [mejmun-χejl nalist=itfuz dzu]*
 this [guest-NOM.PL sit.INF=REL place]
 ‘This is the place [where guests sit].’
- (22) *waz <ajdzjamol-an wi kudzur nalist-i>*
 1SG.NOM <Aijiamol-GEN 3SG.POSS where sit.INF-NMLZ >

wazon = am
 know.NPST = 1SG.NPST
 ‘I know <where Aijiamol will sit>.’
- (23) *{taw nalist alo} waz mas niθ = am*
 {2SG.NOM sit.INF TEMP} 1SG.NOM also sit.NPST = 1SG.NPST
 ‘{When you sit}, I will also sit.’

Verbs are not an open lexical class, in the sense that new verbs with all four stems are not regularly added to the lexicon. More commonly, Sarikoli uses a large amount of nouns and adjectives that express verbal meanings with the help of the verb ‘do’. In the following table, *at* is ‘open’ (adj), *tej* is ‘wedding’ (n), *gap* is ‘word’ (n), *afu* is ‘forgiveness’ (n), and so on.

Table 5. Examples of Nouns and Adjectives with ‘Do’

English gloss	NON-PAST	PAST	PERFECT	INFINITIVE
‘open’	at ka	at tʃəwɣ	at tʃəwɣ	at tʃejɣ
‘marry’	tej ka	tej tʃəwɣ	tej tʃəwɣ	tej tʃejɣ
‘say’	gap ka	gap tʃəwɣ	gap tʃəwɣ	gap tʃejɣ
‘forgive’	afu ka	afu tʃəwɣ	afu tʃəwɣ	afu tʃejɣ
‘finish’	adu ka	adu tʃəwɣ	adu tʃəwɣ	adu tʃejɣ
‘work’	tɕer ka	tɕer tʃəwɣ	tɕer tʃəwɣ	tɕer tʃejɣ
‘play’	skit ka	skit tʃəwɣ	skit tʃəwɣ	skit tʃejɣ
‘influence’	tasir ka	tasir tʃəwɣ	tasir tʃəwɣ	tasir tʃejɣ
‘travel’	sajoat ka	sajoat tʃəwɣ	sajoat tʃəwɣ	sajoat tʃejɣ
‘help’	jurdam ka	jurdam tʃəwɣ	jurdam tʃəwɣ	jurdam tʃejɣ
‘do homework’	topɕurruq ka	topɕurruq tʃəwɣ	topɕurruq tʃəwɣ	topɕurruq tʃejɣ

Tense is formed by the non-past, past, and perfect verb stems, together with the pronominal agreement clitics. The non-past tense, which expresses present and future events, is formed by attaching the appropriate non-past pronominal agreement clitic to

the non-past verb stem. The non-past tense ambiguously expresses the present and future together; or it clearly express one of those through the context, such as by using time words (as in (26)):

- (24) *maf lejligul tfardz wejn = an*
 1PL.NOM Leiligeel good see.NPST = 1PL.NPST
 ‘We love Leiligeel.’
- (25) *wođ kutub xuj = in*
 3PL.NOM book read.NPST = 3PL.NPST
 ‘They are reading books/They will read books.’
- (26) *waz pugan topfuruq ka = m*
 1SG.NOM tomorrow homework do.NPST = 1SG.NPST
 ‘I will do homework tomorrow.’

The past tense is formed by using the appropriate past pronominal clitic with the past stem of the verb. But unlike the non-past tense, the pronominal clitic moves around and does not attach to the verb. It usually attaches to the first major constituent of the sentence. This different placement of the pronominal clitic, together with the use of the past tense verb stem, creates the past tense:

- (27) *wođ = af kutub xojd*
 3PL.NOM = 3PL.PST book read.PST
 ‘They read books.’
- (28) *waz = am topfuruq tfəwg*
 1SG.NOM = 1SG.PST homework do.PST
 ‘I did my homework.’
- (29) *pa dijur = an tujd*
 to hometown = 1PL.PST go.PST
 ‘We went up to our hometown.’

Any usage of the infinitive stem of the verb will involve subordination, since infinitive stems cannot function as the main verb of the matrix clause. If an infinitival clause functions as a modifier of a head noun, as in (30) below, it is an unmarked relative clause

(see 3.3.5 for more discussion). If an infinitival clause functions as an argument of the matrix verb, it is an infinitival complement clause, as in (31), (32), and (33) below (see 4.1.2 for more discussion).

(30) *nur [dam zoxt] maθ*
 today [rest get.INF] day
 'Today is a day [on which one gets rest].' (Arlund & Ibrukhim 2013:136)

(31) *[xipik tfejg] qilo nist*
 [flatbread do.INF] difficult NEG
 '[Making flatbread] is not difficult.'

(32) *waz = am [tu = ri tilfon tfejg] ranuxtf*
 1SG.NOM = 1SG.PST [2SG.ACC = DAT phone do.INF] forget.PRF
 'I forgot [to call you].'

(33) *awal [usul tfejg] xumand so-m*
 first [dance do.INF] learn become.NPST-1SG.NPST
 'First, I will learn [to dance].'

CHAPTER 3

Relative Clauses

Relativization in Sarikoli involves two clauses, the relative clause (RC) and the matrix clause, that are linked together by a common argument. The RC modifies the common argument within the matrix clause (Dixon 2010). Throughout this thesis, RCs will be enclosed in square brackets.

3.1 The common argument

In Sarikoli, the common argument that is shared by the RC and the matrix clause can be realized in two ways. First, the fullest statement of the common argument may be in the matrix clause, and gapped in the RC. This is an externally-headed RC. The common argument in (34), *ejrplon* ‘airplane’, is fully stated in the matrix clause and not stated at all in the RC. Sarikoli does not use resumptive pronouns, which state the common argument within the RC, as shown in the ungrammatical example (35).

(34) *[εη buland rowixt=itfuz] ejrplon utup kaxt*
[most high fly.INF=REL] airplane win do.3SG.NPST
‘The airplane [that flies the highest] wins.’

(35) **[ju εη buland rowixt=itfuz] ejrplon utup kaxt*
[3SG.NOM most high fly.INF=REL] airplane win do.3SG.NPST
‘The airplane [that it flies the highest] wins.’

Second, the common argument may be stated in neither clause, creating a headless RC. If the common argument can reasonably be understood from the situational context, it may simply be omitted. In (36), an example of a headless RC, the common argument is not stated at all but the RC still modifies some object that is understood.

- (36) [*εγ buɫand rowixt = itfuz*] *utup kaxt*
 [most high fly.INF = REL] win do.3SG.NPST
 ‘The one [that flies the highest] wins/will win.’

The common argument in the matrix clause is the head of the RC. In Sarikoli, certain types of NP head are more likely to function as the common argument (i.e. be relativized). Common nouns are the most commonly relativized, as shown in many of the examples in this chapter. Proper nouns are also allowed, though they are less common. Relativized proper nouns get non-restrictive RCs, even though there is no structural difference between restrictive and non-restrictive RCs in Sarikoli:

- (37) [*usul χuf vεð = dzɛndz*] *gulnuz puɫan ar tej usul tfejg mejdz*
 [dance happy be.PRF = REL] Gulnuz tomorrow at wedding dance do.INF plan
 ‘Gulnuz, [who likes to dance], plans to dance at the wedding tomorrow.’

In general, pronouns are not relativized, and the two propositions are simply expressed as two separate sentences. For example, rather than the ungrammatical sentence in (38), Sarikoli speakers say, ‘She likes to dance. She plans to dance at the wedding tomorrow.’ And instead of (39), they say, ‘He knows everything. I will ask him’.

- (38) **[usul χuf vεð = dzɛndz] ju puɫan ar tej usul tfejg*
 [dance happy be.PRF = REL] 3SG.NOM tomorrow at wedding dance do.INF
mejdz
 plan
 ‘She, [who likes to dance], plans to dance at the wedding tomorrow.’

- (39) **waz az [a = dzam tsiz wazond = itfuz] az wi adi*
 1SG.NOM from [ACC-all thing know.INF = REL] from 3SG.ACC this
pars = am
 ask.NPST = 1SG.NPST
 ‘I will ask him, [who knows everything].’

Demonstratives may be relativized, if the context allows the hearer to understand what they are referring to. For example, they may be used when the speaker is pointing at the relativized object(s):

- (40) *[m-oto xjɛb zuxt=tʃɛndʒ] ujeʃ mo-χor*
 [1SG.POSS-dad yesterday buy.PRF=REL] those PROH-eat.NPST
 ‘Don’t eat those [that dad bought yesterday].’

- (41) *maf [olim ɖud=dʒɛndʒ] madi rafon=an*
 1PL.NOM [Olim give.PST=REL] this use.NPST=1PL.NPST
 ‘Let’s use this [which Olim gave].’

Finally, some generic terms¹, such as *dʒuj* ‘place’, *waxt* ‘time’, and *rang* ‘manner’, may be relativized. When these generic terms function as RC heads, they form locative, temporal, and manner adverbial clauses, respectively. These will be further discussed in chapter 5. Apart from these words, generic terms are not commonly relativized; instead, headless RCs are often used when the RC is modifying something general, as in the following example:

- (42) *a=bilat=an [bilat paraɖod=itʃuz]=ir ɖud*
 ACC=ticket=1PL.PST [ticket sell.INF=REL]=DAT give.PST
 ‘We gave our tickets to the one [who sells tickets].’ (Arlund & Ibrukhim 2013:82)

Sarikoli allows a wide range of syntactic functions for the common argument in the matrix clause and the RC. In the matrix clause, the common argument can be in the core argument (A, S, and O²) positions:

- (43) *[zord puqzo, χalg-ɛʃ a=muunoswat tʃardʒ ramud=itʃuz] χalg*
 [heart pure person-ACC.PL ACC=relationship good cause.INF=REL] person

barakat vrejɖ
 blessing find.3SG.NPST
 ‘People [whose heart is pure, and cause good relationships among people] will find blessing.’ (matrix: A, RC: S & A)

- (44) *[ɛʃ buɫand rowixt=itʃuz] ejrplon utuʋ kaxt*
 [most high fly.INF=REL] airplane win do.3SG.NPST
 ‘The airplane [that flies the highest] wins/will win.’ (matrix: S, RC: S)

¹ The term "generic term" is used in Dixon (2010) in the section discussing possible heads of relative clauses.
² A is the most agent-like argument of a transitive clause; S is the single argument of an intransitive clause; O is the most patient-like argument of a transitive clause (Dixon 1972:128).

(45) *waz = am* *[maj tci mofin zjer wjeð = dzendz]* *a-vaðo ðud*
 1SG.NOM = 1SG.PST [1PL.POSS at car rock throw.PST = REL] ACC-boy hit.PST
 ‘I hit the boy [who threw a rock at our car].’ (matrix: O, RC: A)

(46) *[maktab muu = ri ðud = dzendz]* *pul mu xojd rasqut,*
 [school 1SG.ACC = DAT give.PST = REL] money 1SG.POSS study.INF fee

jatoq rasqut, tamoq rasqut = ir fropst
 dorm fee food fee = DAT reach.3SG.NPST
 ‘The money [that the school gave me] is enough for tuition, room, and board fees.’
 (matrix: S, RC: O)

The common argument may also be an indirect object, possessor, or possessed item in the matrix clause, as in the following examples. However, it should be noted that it is rare for the common argument to function as a possessed item in the matrix clause, as in (49).

(47) *malum a = mukofot bujum [dzam sawul-ef = ir toʁri dzawub*
 teacher ACC = prize object [all question-ACC.PL = DAT correct answer

ðud = dzendz] *oquʁutci = ri ðud*
 give.PST = REL] student = DAT give.PST
 ‘The teacher gave the prize to the student [who gave the correct answer to all the questions].’ (matrix: indirect object, RC: A)

(48) *[mudz varcidje iθ = tfendz]* *χalg-an wi kol ðizd*
 [new Tashkurgan come.PRF = REL] person-GEN 3SG.POSS head hurt.3SG.NPST
 ‘People [who come to Tashkurgan for the first time]’s heads hurt.’ (matrix: possessor, RC: S)

(49) *kuud [χu sujib-an a = dver at tfəwy = dzendz]* *awudz xud*
 dog [own owner-GEN ACC = door open do.PRF = REL] sound hear.PST
 ‘The dog heard its owner’s sound [that the door was opened].’ (matrix: possessed item; RC: none?)³

Finally, the common argument may also be the object of comparison, as in (50), or play oblique roles of time or location, as in (51) & (56):

³ The bracketed clause in this example is not the best example of a RC because it does not meet the criteria of having a common argument. The head noun ('sound') is not a gapped argument that plays a role within the RC. However, I still use it here as an example of a RC because it uses the relativier = *dzendz* and modifies the head noun 'sound' in some way. Alternatively, this could be analyzed as a noun phrase and a complement clause in apposition (Dixon 2006:11).

- (50) [pa tʃɛd tʃəwɣ=dʒɛndʒ] xipik [pa buzur tʃəwɣ=dʒɛndʒ] az⁴ xipik
 [at home do.PRF=REL] flatbread [at market do.PRF=REL] from flatbread

χɛg
 sweet

‘Flatbread [that is made at home] is more delicious than flatbread [that is made at the market]. (matrix: object of comparison, RC: O)

- (51) [dars boʻlamif tʃəwɣ=dʒɛndʒ] maθ, malum-χejl dʒam aldroy
 [class start do.PRF=REL] day teacher-NOM.PL all busy
 ‘On the day that classes begin, all the teachers are busy.’ (matrix: oblique; RC: oblique)

- (52) ju [waz parus tuijd=dʒɛndʒ] ar maktab tuijd
 3SG.NOM [1SG.NOM last.year go.PST=REL] to school go.PST
 ‘He went to the school [that I went to last year].’ (matrix: oblique; RC: oblique)

Within the RC, the common argument may also be in a wide range of syntactic functions. The examples above (some of which are repeated below) show that common arguments may be in the A, S, O, and oblique functions:

- (53) waz=am [maf tʃi moʻfin zjer wjɛð=dʒɛndʒ] a-vaðo ðud
 1SG.NOM=1SG.PST [1PL.POSS at car rock throw.PST=REL] ACC-boy hit.PST
 ‘I hit the boy [who threw a rock at our car].’ (matrix: O, RC: A)

- (54) [ɛŋ buʻland rowixt=itʃuz] ejrplon utuup kaxt
 [most high fly.INF=REL] airplane win do.3SG.NPST
 ‘The airplane [that flies the highest] wins/will win.’ (matrix: S, RC: S)

- (55) [maktab muu=ri ðud=dʒɛndʒ] pul mu xojd rasqut,
 [school 1SG.ACC=DAT give.PST=REL] money 1SG.POSS study.INF fee
 jatoq rasqut, tamoq rasqut=ir fropt
 dorm fee food fee=DAT reach.3SG.NPST
 ‘The money [that the school gave me] is enough for tuition, room, and board fees.’
 (matrix: S, RC: O)

- (56) ju [waz parus tuijd=dʒɛndʒ] ar maktab tuijd
 3SG.NOM [1SG.NOM last.year go.PST=REL] to school go.PST
 ‘He went to the school [that I went to last year].’ (matrix: oblique; RC: oblique)

⁴ In this context, this preposition is used to introduce the comparative construction.

The common argument may also function as the indirect object within the RC, as in the following example.

- (57) [*waz mon paraðud=dzɛndz*] *kaðo vurdi woxt*
 [1SG.NOM apple sell=REL] boy trip fall.PST
 ‘The boy [I sold apples to] fell.’ (matrix: S; RC: indirect object)

Semantically, the common argument may be a possessor in the RC, but this is not shown structurally with a genitive marker because it is not a possessor in the matrix clause:

- (58) [*ato ano jet na-tci-tfəwɣ=dzɛndz*] *oqʷutci-xejl laka xu*
 [dad mom come.INF NEG-can-do.NPST=REL] student-NOM.PL let own

bob mom vor=in
grandpa grandma bring.NPST=3PL.NPST
 ‘Students [whose parents cannot come] may bring their grandparents.’ (matrix: A;
 RC: possessor)

3.2 The RC

Most Sarikoli RCs are marked by a relativizer, usually with the enclitic =*dzɛndz* or =*itfuz*. The difference between these two relativizers will be discussed in sections 3.3.1 and 3.3.2. Unlike relative pronouns, these relativizers do not indicate anything about the reference of the common argument. The relativizer occurs at the end of the RC, attaching itself to the verb.

3.2.1 Placement of the RC

As a head-final language, Sarikoli places the RC before its head. If the common argument is stated in the matrix clause, the RC always precedes the common argument. If the RC is headless, it still precedes the slot where the common argument normally occurs, even though the common argument is not explicitly stated. For example, the RC in (34) has an external head and the one in (36) is headless, but they both precede the slot where the common argument is normally placed.

3.2.2 Grammatical marking within the RC

Some Sarikoli RCs can include the same grammatical marking for tense as in an independent clause. The finiteness of verbs within the RC will be described in the section about the relativizers (3.3). However, while finite RCs contain past or perfect verb stems, they do not show subject-verb agreement through pronominal clitics. As mentioned previously, Sarikoli only uses the gap strategy within RCs, and does not use any kind of resumptive pronouns, including pronominal clitics.

As shown in the following examples, if the RC modifies a noun that is in the accusative case, the pronominal subject stated within the RC is expressed in the accusative/possessive form. If the RC-subject is different from the matrix clause-subject, then the appropriate accusative/possessive pronoun is used, as in (60). If it is the same as the matrix clause-subject, as in (62), it is expressed as *χu* ‘own’, the reflexive pronoun in the accusative/possessive form. *χu* can only be coreferential with the grammatical subject of the clause or sentence, and can take as its antecedent any argument regardless of the antecedent's person or number.

(59) *təw = at* *kutub navift*
 2SG.NOM = 2SG.PST book write.PST
 ‘You wrote a book.’

(60) *waz = am* [*ta* *navift = tʃendz*] *kutub xojd*
 1SG.NOM-1SG.PST [2SG.POSS write.PRF = REL] book read.PST
 ‘I read the book [that you wrote].’ (no clitic in RC)

(61) *waz = am* *xipik tʃəwɣ*
 1SG.NOM = 1SG.PST flatbread make.PST
 ‘I made flatbread.’

(62) *waz = am* [*χu tʃəwɣ = dʒendz*] *xipik χuɣ*
 1SG.NOM = 1SG.PST [own make.PRF = REL] flatbread eat.PST
 ‘I ate the flatbread [that I made].’ (no clitic in RC)

3.2.3 Determiners, adjectives, and adpositions within the matrix clause

Determiners and RCs are both modifiers that precede the head noun. When they co-occur, the order of determiner and RC are flexible. The determiner may either precede or follow the RC.

- (63) *(uji) [m-oto xjɛb zuxt = tʃɛndʒ] (uji) mon-ɛf*
 (that) [1SG.POSS-dad yesterday buy.PRF = REL] (that) apple-ACC.PL

mo-χor
 PROH-eat.NPST

‘Don’t eat those apples [that dad bought yesterday].’

- (64) *waz (uji) [noj χid = itfuz] (uji) batʃo wazon = am*
 1SG.NOM (that) [eagle.flute play.INF = REL] (that) child know.NPST = 1SG.NPST
 ‘I know that kid [who is blowing the eagle flute].’

- (65) *waz (χu) az [a = dʒam tsiz wazonɔ = itfuz] (χ)-oto adi*
 1SG.NOM (own) from [ACC = all thing know.INF = REL] (own)-dad this

pars = am
 ask.NPST = 1SG.NPST

‘I will ask my father, [who knows everything].’

RCs may also co-occur with adjectives if a head noun is modified by both. When they co-occur, RCs are placed farther away from the head noun and the adjectives are placed closer to it.

- (66) *awagul [χu-jan xjɛb zuxt = tʃɛndʒ] qimat ajoy pamowɔdʒ*
 Awagul [own-GEN yesterday buy.PST = REL] expensive shoe wear.PRF
 ‘Awagul wore the expensive shoes [that she bought yesterday].’

- (67) *zulfia [χ-ono tʃəwɔ = dʒɛndʒ] tazo tuχp poj furd*
 Zulfia [own-mom do.PST = REL] very sour yogurt slurp.PST
 ‘Zulfia slurped the very sour yogurt [that her mom made].’

- (68) [*χu tar tsem ujnak weðd=dzændz*] *aqlin batfo az watça*
 [own toward eye glass put.PRF=REL] smart child from Wacha

iθ-tfændz

come.PRF-PRF

‘The smart student [who wears glasses] is from Wacha.’

If the head noun of a RC is part of an adposition phrase, the order of preposition and RC is somewhat flexible, and postpositions are placed on the other side of the RC and its head noun. If the head noun is an object of a preposition in the matrix clause, the preposition often occurs between the RC and its head, even though it is not a modifier of the noun:

- (69) *putxu [χu duxman a=χu naymuy=dzændz] ar bor deðdz*
 king [3SG.POSS enemy ACC=own hide.PRF=REL] in cave enter.PRF
 ‘The king went into the cave that [his enemies hid themselves in].’

- (70) [*pa tçed tfəwγ=dzændz*] *xipik [pa buzur tfəwγ=dzændz] az xipik*
 [at home do.PRF=REL] flatbread [at market do.PRF=REL] from flatbread

χεg

sweet

‘Flatbread [that is made at home] is more delicious than flatbread [that is made at the market].’

Less commonly, the preposition may be placed before the RC, farther away from the head noun. The following examples would still be grammatical if the preposition is placed between the RC and its head:

- (71) *sarmsok tçi [xipik tsaft=tfændz] purg i zjer awulo tfəwγ*
 Sarmsok at [flatbread steal.PST=REL] mouse one rock throw do.PST
 ‘Sarmsok threw a rock at the mouse [that stole the bread].’

- (72) *waz χu az [a=dzəmə tsiz wazond=itfuz] ato adi*
 1SG.NOM own from [ACC=all thing know.INF=REL] dad this

pars = am

ask.NPST = 1SG.NPST

‘I will ask my father, [who knows everything], about this.’

If the head noun of the RC is an object of a postposition in the matrix clause, the postposition occurs after the RC and its head noun:

(73) *m-oto* [*χu weðd=dzɛndz*] *tɛd tçi tjer a=kalo kaxt*
 1SG.POSS-dad [own put.PST=REL] house at top ACC=sheep kill.PST
 ‘My father killed the sheep on top of the house [that he built].’

(74) *mu jaχ di [urdu bejt lɛvd=dzɛndz] tʃurik az zabu bejt lɛvd*
 1SG.POSS sister this [Urdu song say.PRF=REL] man from after song say.PST
 ‘My sister sang after this man [who sang the Urdu song].’

3.2.4 Non-core elements within the RC

RCs may include peripheral arguments, as well as time, place, and manner words that occur in independent clauses. The RCs in the examples below contain these non-core elements. (75) contains an oblique argument, (76) contains a time word, (77) contains place words, and (78) contains manner words.

(75) *waz=am* [*maf tçi mofin zjer weðd=dzɛndz*] *a=kaðo*
 1SG.NOM=1SG.PST [1PL.POSS at car rock throw.PST=REL] ACC=boy

ðud
 hit.PST
 ‘I hit the boy [who threw a rock at our car].’

(76) [*maf parus tujd=dzɛndz*] *risturon-aθ sjeð tfaktf*
 [1PL.NOM last.year go.PST=REL] restaurant-??? this.year demolished.PRF
 ‘The restaurant [that we went to last year] was demolished.’

(77) [*pa tɛd tʃəwɣ=dzɛndz*] *xipik* [*pa buzur tʃəwɣ=dzɛndz*] *az xipik*
 [at home do.PRF=REL] flatbread [at market do.PRF=REL] from flatbread

χεg
 sweet
 ‘Flatbread [that is made at home] is more delicious than flatbread [that is made at the market].’

(78) [*εη buvand rowixt=itfuz*] *ejrplon utuup kaxt*
 [most high fly.INF=REL] airplane win do.3SG.NPST
 ‘The airplane [that flies the highest] wins/will win.’

3.2.5 Negation in RCs

RCs are negated the in same way as regular clauses. Sarikoli negates regular clauses by attaching the *na-* prefix to the verb, as in (79) & (81), and RCs attach the same negative prefix to the verb, as in (80) & (82).

(79) *itaŋ batfo-χejl ato ano na-veðdz*
some child-NOM.PL father mother NEG-be.PRF
'Some children do not have parents.'

(80) [*ato ano na-veð = dzendz*] *batfo eŋ ivul*
[father mother NEG-be.PRF = REL] child most pitiful
'Children who do not have parents are the most pitiful.'

(81) *itaŋ χal-χejl a = qiqat na-wazon = in*
some person-NOM.PL ACC = truth NEG-know.NPST = 3PL.NPST
'Some people do not know the truth.'

(82) [*a = qiqat na-wazon d = dzendz*] *χalg [χu pond xatu ðud = dzendz]*
[ACC = truth NEG-know.PRF = REL] person [own way lose give.PST = REL]

mawl = ir oχford
sheep = DAT like
'People [who do not know the truth] are like sheep [that lost their way].'

3.3 Types of Sarikoli RCs

Sarikoli uses two main relativizers for creating RC constructions: = *dzendz* and = *itfuz*. This section will introduce these relativizers, as well as headless RC constructions, the genitive marker used as a relativizer, unmarked RCs, non-restrictive RCs, and adjectivizers that form adjective phrases. In Sarikoli RCs, the choice of relativizer is determined by the type of verb stem used within the RC: finite or non-finite.

3.3.1 Finite RCs: the relativizer = *dʒendʒ*

The relativizer = *dʒendʒ* is used with RCs that contain events that have already been completed. It is the only relativizer that attaches to finite verbs, as it occurs with the past or perfect stems of the verbs. It cannot attach to non-past or infinitive verb stems, as shown in the following ungrammatical examples:

- (83) **waz* [χu ka = *dʒendʒ*] xipik χor = *am*
 1SG.NOM [own make.NPST = REL] flatbread eat.NPST = 1SG.NPST
 ‘I will eat the flatbread [that I make].’

- (84) **waz = am* [χu tfejg = *dʒendʒ*] xipik χug
 1SG.NOM = 1SG.PST [own make.INF = REL] flatbread eat.PST
 ‘I ate the flatbread [that I make].’

The surface form of this relativizer is phonologically conditioned by the verb stem it attaches to. If the verb stem ends in a voiced consonant, the relativizer is realized as = *dʒendʒ*; if the verb ends in a voiceless consonant, the relativizer is = *tʃendʒ*.

RCs marked with = *dʒendʒ* may occur in various types of matrix clauses. They may occur in matrix clauses with events that have happened in the past:

- (85) [*duχtur wi = ri* δud = *dʒendʒ*] *duri wi = ri utf fujdo*
 [doctor 3SG.ACC = DAT give.PST = REL] medicine 3SG.ACC = DAT very effective
tʃəwɔ
 do.PST
 ‘The medicine [that the doctor prescribed him] worked very well on him.’

- (86) [*parus xojd adu tʃəwɔ = dʒendʒ*] *batfo-χejl = af tʃardʒ*
 [last.year study.INF finish do.PRF = REL] child-NOM.PL = 3PL.PST good
χuzmat vug
 work find.PST
 ‘[Kids (students) [who graduated last year] have found good jobs.’

- (87) *yubun* [χu bunost = *tʃendʒ*] *a = barqo vug, kazwi zjofat δud*
 shepherd [own lost.PRF = REL] ACC = lamb find.PST so party give.PST
 ‘The shepherd found the lamb [that he had lost], so he gave a party.’

They may also be embedded in matrix clauses with non-past events:

- (88) *waz* [wi *levd=dzɛndz*] *gap=ir ifandz ka=m*
 1SG.NOM [3SG.POSS say.PST=REL] word=DAT believe do.NPST=1SG.NPST
 ‘I believe the words [he said].’
- (89) [*parus tej tɬəwɣ=dzɛndz*] *ɤots padijom batfo veɟ=ir vɛðdɟ*
 [last.year wedding do.PRF=REL] girl twin child bring.INF=DAT be.PRF
 ‘The girl who got married last year will give birth to twins (hearsay).’

They may occur in sentences with non-verbal predicates, either as a modifier of the subject, as in (90), or of the predicate nominal, as in (91):

- (90) [*xwor pjɛxt=dzɛndz*] *anur dinju-jɛndz ɛŋ xɛg anur*
 [Kashgar grow.PST=REL] pomegranate world-ADJ most sweet pomegranate
 ‘The pomegranates that grew in Kashgar are the sweetest pomegranates in the world.’
- (91) *mu malum uji [tɕɛm ujnək wɛðd=dzɛndz] əwrat malum*
 1SG.POSS teacher that [eye glass put.PRF=REL] woman teacher
 ‘My teacher is that female teacher over there [who put on glasses].’

They may occur in matrix clauses with modal auxiliary verbs:

- (92) [*digar xalg zjɛð=dzɛndz*] *a=xalg tfoqom zjɛd luzim*
 [other person kill.PST=REL] ACC=person must kill.INF should
 ‘A person [who killed another person] must be killed.’

They may be embedded within complement clauses:⁵

- (93) *mu puts dil maf=ir [xɯ xojd=dzɛndz] a=maktab*
 1SG.POSS son heart 1PL.ACC=DAT [OWN study.PST=REL] ACC=school

vusond
 show.INF
 ‘My son wants to show us his school [where he studied].’

Finally, they may also occur multiple times within a single matrix clause, each as modifiers of different heads. In the following example, a =*dzɛndz* RC modifies the object, and another one modifies the indirect object.

⁵ *Dil* is a very atypical word. It will be discussed further in chapter 4.

- (94) *m-ono* [pa tɕɛd tʃəwŋ=dʒɛndʒ] a=xipk [nudz katf
 1SG.POSS-mother [at home do.PST=REL] ACC=flatbread [new move

 iθ=tʃɛndʒ] qufni=ri ðud
 come.PRF=REL] neighbor=DAT give.PST
 ‘Mother gave flatbread [that was made at home] to neighbors [who just moved in].’

The relativizer =*dʒɛndʒ* may sometimes be truncated to =*dʒ*, so that the verb looks exactly the same as the perfect stem. In the following example, the verb in the second RC may be realized as either *na-wandʒɛndʒ* or *na-wandʒ*.

- (95) *m-ono* [maf tɕɛd-nɛndʒ-χejl qati woχt=dʒɛndʒ]
 1SG.POSS-mom [1PL.POSS house-ADJ-ACC.PL together take.PST=REL]

 a=rasim χu-jan [ðɛs sul na-wand=dʒ] amru=ri vuɔsɔnd
 ACC=picture own-GEN [ten year NEG-see=REL] friend.F=DAT show.PST
 ‘My mother showed the picture [that our family took together] to her friend [whom she has not seen for 10 years].’

There are some exceptions to the rule that this =*dʒɛndʒ* relativizer is only used with past events. First, some non-restrictive RCs may contain non-past events and still take the =*dʒɛndʒ* relativizer (see (127) and (128), in which the common arguments are already fully identified with proper nouns; even though the relativized events are not in the past, they take the =*dʒɛndʒ* relativizer). Second, manner adverbial clauses, which are formed through relativization, always take the =*dʒɛndʒ* relativizer, regardless of the time of the embedded event. Even in these exceptions, however, =*dʒɛndʒ* always attaches to the past or perfect stems of verbs.

3.3.2 Infinitival RCs: the relativizer =*itfuz*

The relativizer =*itfuz* serves two purposes. First, it marks RCs that do not involve tense, such as clauses with events that are timeless, ongoing, or habitual. Since such clauses are naturally without tense, the verbs in this type of RC occur in the infinitive stem. Although the infinitive stem is identical to the past stem in some verbs, it is unambiguous that the verbs in this RC construction are in the infinitive stems, and not in the past stems, because =*itfuz* never occurs with stems other than the infinitive.

Tense-less RCs marked with =*itfuz* may also occur in various types of matrix clauses.

They may occur in past events:

- (96) [*poj at mareb paraðod=itfuz*] *əwrat χu ar mulk tujdz*
 [yogurt and naipizi sell.INF=REL] woman own to hometown go.PRF
 ‘The woman [who sells yogurt and naipizi] went to her hometown.’

They may occur in gnomic statements:

- (97) [*iyubun awudz wazond=itfuz*] *kalo χu pond xatu na-ðid*
 [shepherd sound know.INF=REL] sheep own way lose NEG-give.3SG.NPST
 ‘The sheep [who knows the shepherd’s voice] will not lose its way.’

- (98) [*qurbun ejd=ir kaxt=itfuz*] *kalo [digar ejd=ir*
 [Korban holiday=DAT kill.INF=REL] sheep [other holiday=DAT
kaxt=itfuz] tar kalo-jef arðo na-ðejd
 kill.INF=REL] toward sheep-ACC.PL same NEG-fall.3SG.NPST
 ‘Sheep [that are killed for Korban holiday] are not the same as sheep [that are killed for other holidays].’

They may occur in sentences with non-verbal predicates. Again, they may either modify the subject, as in (99) & (100), or of the predicate nominal, as in (101):

- (99) [*xwor pjext=itfuz*] *anur dinju-jendz eη χεg anur*
 [Kashgar grow.INF=REL] pomegranate world-ADJ most sweet pomegranate
 ‘The pomegranates [that grow in Kashgar] are the sweetest pomegranates in the world.’

- (100) [*tudzik-χejl eη pur χig=itfuz*] *tamoq poj, xipik, qatebin tfoj*
 [Tajik-NOM.PL most many eat.INF=REL] food yogurt flatbread milk tea
 ‘Food [that Tajiks eat the most] is yogurt, flatbread, and milk tea.’

- (101) *mu malum [tsem ujnak weðd=itfuz] əwrat malum*
 1SG.POSS teacher [eye glass put.INF=REL] woman teacher
 ‘My teacher is the female teacher [who puts on glasses].’

They may also modify the subject of a predicate adjective:

- (102) *[tar bajixt tid=itfuz] pond utf tong*
 [toward heaven go.INF=REL] way very narrow
 ‘The road [that goes toward heaven] is very narrow.’

The second purpose that the relativizer =*itfuz* serves is to mark RCs that contain non-past events. This entails verbs that convey present or future events. However, within RCs, present and future events are all expressed with non-finite verbs, in their infinitive stem. So structurally, there is no difference between the tense-less =*itfuz* RCs and the non-past =*itfuz* RCs, despite their semantic difference.

These non-past =*itfuz* RCs may occur within past events:

- (103) *[tej tfejg=itfuz] kots-an wi bob mowy*
 [wedding do.INF=REL] girl-GEN 3SG.POSS grandpa die.PRF
 ‘The girl [who will get married]’s grandfather died.’

- (104) *faridun [χu pugan levd=itfuz] a=bejt mafq tfəwg*
 Faridun [own tomorrow sing.INF=REL] ACC=song practice do.PST
 ‘Faridun practiced the song [he will sing tomorrow].’

They may occur in non-past events:

- (105) *waz [wi levd=itfuz] gap wazon=am*
 1SG.NOM [3SG.ACC say.INF=REL] word know.NPST=1SG.NPST
 ‘I know the words [he will say].’

They may occur in matrix clauses with modal auxiliary verbs:

- (106) *[sulir xojd adu tfejg=itfuz] batfo-χeil az uzir χuzmat*
 [next.year study.INF finish do.INF=REL] child-NOM.PL from now work

xikejg luzim
 search.INF should
 ‘Kids (students) [who are graduating next year] should start job searching now.’

- (107) *[sarmsoq pugan vujojd=itfuz] χutur=ir dam zoxt luzim*
 [Sarmsok tomorrow ride.INF=REL] camel=DAT rest get.INF should
 ‘The camel [Sarmsok will ride tomorrow] needs to get rest.’

Finally, they may also be embedded within complement clauses:

- (108) *mu puts dil [sjeð pidz tid xojd=itfuz] a=maktab*
 1SG.POSS son heart [this.year fall go.INF study.INF=REL] ACC=school
wand
 see.INF
 ‘My son wants to see the school [he will attend this fall].’

3.3.3 Headless RCs

Even though RCs are always externally headed if the common argument is explicitly stated, the expression of the common argument is not always required. If the common argument can be understood on the basis of the situational context in which the utterance occurs, it may be omitted. Headless RCs may be formed from both =*dzendz* and =*itfuz* RCs. (109) below is an example of a headless RC with the =*itfuz* relativizer; (110), (111), and (112) are examples of headless RCs with the =*dzendz* relativizer. In all four examples, the RC modifies the implicit subject of a predicate nominal. Notice that none of these four sentences contain subject-verb agreement clitics, which are required for sentences with a regular verb in the matrix clause. Since the only verbs that occur in these examples are embedded within a RC, there are no verbs in the matrix clause, and the absence of subject-verb agreement clitics is expected.

- (109) [*tudzik-xejl εη pur xig=itfuz] poj, xipik, qatεsin tfoj*
 [Tajik-NOM.PL most many eat.INF=REL] yogurt flatbread milk tea
 ‘Stuff [that Tajiks eat the most] are yogurt, flatbread, and milk tea.’

- (110) *maf-an itaη awlut [az fugnon iθ=tjendz]*
 1PL.NOM-GEN some ancestor [from Shughnon come.PRF=REL]
 ‘Some of our ancestors are those [who came from Shughnon].’

- (111) *mu piʃ [waz dzuli waχt mowγ=dzendz]*
 1SG.POSS cat [1SG.NOM small time die.PRF=REL]
 ‘My cat is the one [that died when I was young].’

- (112) *waz [tudzik milat rasmi dzongwo-jendz ozsonleq milat*
 1SG.NOM [Tajik ethnic truly China-ADJ minority ethnic
seð=dzændz maθ azmud seð=dzændz]
 become.PRF=REL day born become.PRF=REL]
 ‘I am one [who was born on the day that the Tajiks officially became a minority
 group of China].’

The headless RC may occur by itself as a core argument (as in (109) above), or may also take case markers and subject-verb agreement clitics which usually attach to the head nouns. These enclitics may be directly attached to the whole RC. In the following example, the RC takes the nominative plural marker and the past tense agreement clitic for third-person plural subjects.

- (113) *[mur intowum ðod=itfuz]-xejl=af wayt qati pa maktab jot*
 [today test give=REL]-NOM.PL=3PL.PST time with at school come.PST
 ‘Those [who are taking the test today] came to school on time.’

They may also be marked as accusative case, as shown by the accusative plural marker in this example:

- (114) *maf=an [az xwor vəwg=dzændz]-jef-an paraðud adu*
 1PL.POSS=GEN [from Kashgar bring.PST=REL]-ACC.PL-1PL.PST sell.PST finish
tʃəwg
 do.PST
 ‘We finished the selling of the stuff [we brought from Kashgar].’

Finally, they may be marked as dative case.

- (115) *waz [wi levd=dzændz]=ir ifandz ka=m*
 1SG.NOM [3SG.POSS say.PST=REL]=DAT believe do.NPST=1SG.NPST
 ‘I believe what [he said].’
- (116) *[a=otʃku vig=itfuz]=ir waz ðes kuj ðo=m*
 [ACC=key find.INF=REL]=DAT 1SG.NOM ten kuai give.NPST=1SG.NPST
 ‘I will give 10 kuai to the one [who finds the key].’

3.3.4 The genitive marker *-an*

In a limited context, the genitive marker *-an* may attach to an infinitival clause to form another type of RC. This is used in predicative constructions which contain a noun modified by a RC. This type of RC may modify either the subject, as in (117), or the predicate, as in (118).

- (117) *[varçidje tid-an] eŋ χil waχt mjendz*
[Tashkurgan go.INF-GEN] most good time summer
'The best time [to go to Tashkurgan] is summertime.'

- (118) *ju [kalo pwojd-an] tfardz dzuŋ*
3SG.NOM [sheep herd.INF-GEN] good place
'It is a good place [to herd sheep].'

3.3.5 Unmarked infinitival RCs

RCs may be completely unmarked, with no relativizer indicating that a clause is modifying a noun. In this type of RC, the verb is in the infinitive stem and the RC simply precedes the head noun:

- (119) *xjeβ xob olim [awu ðod] awudz xjeðdz*
yesterday night Olim [rain fall.INF] sound hear.PRF
'Last night, Olim heard the sound [of rain falling].'⁶

- (120) *nur [dam zoxt] maθ*
today [rest get.INF] day
'Today is a day [on which one gets rest].' (Arlund & Ibrukhim 2013:136)

Other modifiers of the head noun may co-occur with this type of RC. Adjectives may occur between the unmarked RC and the head noun:

- (121) *olim [maf χumand set] tfardz ulgu*
Olim [1PL.NOM study become.INF] good model
'Olim is a good model [that we study].'

⁶ As in example (49), the bracketed clause in this example is not the clearest example of a RC because it does not have a common argument ('sound' does not play a syntactic role in the bracketed clause).

It may also contain time words and further-embedded RCs:

- (122) *[maf xjeb tujd=dʒendʒ usul sawrun dejd] bawu*
 [1PL.NOM yesterday do.PST=REL dance entertainment.place enter.INF] fee
utf qimat
 very expensive
 ‘The fee [for entering the dance entertainment place that we went to yesterday] is very expensive.’

Depending on the verb in the matrix clause, it may be ungrammatical for unmarked RCs to simply omit the head noun, as in (123). However, if the head noun should be eliminated, this type of RC can be converted into a complement clause by attaching the *-i* nominalizer to the infinitive stem of the verb, as in (124). This will be further described in chapter 4.

- (123) **xjeb xob olim [awu ɖod] xjeɖdʒ*
 yesterday night Olim [rain fall.INF] hear.PRF
 ‘Last night, Olim heard rain falling.’

- (124) *xjeb xob olim [awu ɖod-i] xjeɖdʒ*
 yesterday night Olim [rain fall.INF-NMLZ] hear.PRF
 ‘Last night, Olim heard rain falling.’

3.3.6 Non-restrictive RCs

Sarikoli has non-restrictive RCs, in which the common argument is already fully identified and the RC simply adds more information about it, rather than restricting the referent of the common argument. Even though the meanings of restrictive and non-restrictive RCs differ, Sarikoli makes no formal distinction between the two. The two relativizers used for restrictive RCs, *=dʒendʒ* and *=itʃuz*, are also used for marking non-restrictive RCs.

- (125) *waz xʉ az [a=dʒam tsiz wazond=itʃuz] ato adi*
 1SG.NOM own from [ACC=all thing know.INF=REL] dad this
pars=am
 ask.NPST=1SG.NPST
 ‘I will ask my father, [who knows everything], about this.’

(126) [tarjəwl tɕi χejl undəwd=dʒɛndʒ] sofia xipik tɕi tfejg sut
 [morning at six get.up.PRF=REL] Sofia flatbread at make.INF become.PST
 ‘Sofia, [who woke up at six in the morning], began making flatbread.’

(127) [usul χuf vɛð=dʒɛndʒ] gulnuz puɡan ar tej usul tfejg mejdz
 [dance happy be.PRF=REL] Gulnuz tomorrow at wedding dance do.INF plan
 ‘Gulnuz, [who likes to dance], plans to dance at the wedding tomorrow.’

(128) [az aroj azur mitir tjer vɛð=dʒɛndʒ] varçidje zejn-i
 [from three thousand meter above be.PRF=REL] Tashkurgan winter-TEMP
 utf ij
 very cold
 ‘Tashkurgan, [which is above 3,000 meters], is very cold in the winter.’

In the final two examples above, (127) and (128), the RC heads are modified by finite RCs (with the =*dʒɛndʒ*) even though they do not contain events that have happened in the past. It is not clear why they are not =*itfuz* RCs instead.

Non-restrictive RCs may also be used to modify vocatives. Since speakers already have a fully identified addressee, RCs used to modify them only provide extra information:

(129) o [udil tɕɛr tfejg=itfuz] putxu, təw maf a=ɡap bəwl
 o [just work do.INF=REL] king 2SG.NOM 1PL.POSS ACC=word ear
 wejð
 put.NPST
 ‘O king, [who does fair works], listen to our words.’

(130) o [mu-jan tfardʒ wand=itfuz] batfo, maŋɡun mu qati niθ
 o [1SG.POSS-GEN good see.INF=REL] child forever 1SG.ACC with live.NPST
 ‘O child, [whom I love], live with me forever.’

3.3.7 Adjectivizers: *-ɛndʒ*, *-jɛndʒ*, and *-nɛndʒ*

In addition to the two main relativizers introduced above, there are more enclitics that behave somewhat similarly to them: *-ɛndʒ*, *-jɛndʒ*, and *-nɛndʒ*. However, they only seem to attach to time and place words (often adpositional phrases), rather than full clauses with verbs. Thus, they are glossed as adjectivizers instead of relativizers. (131) is an

example of a time word, which becomes an adjective phrase in (132) through the use of *-nɛndz*:

(131) *ʃitʃ xufs*
now sleep.NPST
 ‘Sleep now.’

(132) *[ʃitʃ-nɛndz] awul*
[now-ADJ] situation
 ‘[present] situation’

Like the RCs, these adjectivized phrases are placed closer to the head noun than determiners and possessors:

(133) *mu [zabu-nɛndz] ɚzɯ-jɛf*
1SG.POSS [later-ADJ] hope-ACC.PL
 ‘my [future] hopes’

(134) *oftuvuz-an [ɛŋ prud-ɛndz] a=dzɯj-ɛf sul lɔwr xalg-ɛf=ir*
bus-GEN [most front-ADJ] ACC=place-ACC.PL year big person-ACC.PL=DAT

ðod luzim
give-INF should
 ‘The [very front] seats of the bus should be given to old people.’

And like the RCs, they are also placed farther away from the head noun than lexical adjectives or other nouns modifying the same head noun:

(135) *[dinju-jɛndz] ɛŋ xɛg anur*
[world-ADJ] most sweet pomegranate
 ‘sweetest pomegranates [in the world]’

(136) *[tɕi radijo-jɛndz] batʃo-jɛf bejt*
[at radio-ADJ] child-ACC.PL song
 ‘children’s songs [in the radio]’ (Arlund & Ibrukhim 2013:198)

The following example contains a modifier that precedes the adjectivized phrase, as well as modifiers that follow it. Regardless of how many different types of modifiers there

are, the Sarikoli head noun occurs after all of its modifiers, being true to the head-final OV type.

- (137) *woð [tɕi kursi-jɛndz] ða ləwr tor kutub*
 those [at chair-ADJ] two big black book
 ‘Those two big black books [on the chair]’ (Arlund & Ibrukhim 2013:18)

Adjectivized phrases can also be headless. As in the RCs, the head noun may be omitted, and the plural suffix can be directly attached to the headless phrase. The following example contains an adjectivized phrase that is marked for nominative case:

- (138) *[tɕɛd-nɛndz]-xɛjl qati woxt=tɕɛndz rasim*
 [home-ADJ]-NOM.PL together take=REL photo
 ‘photo that the people [in the home] took together’

The difference between *-ɛndz*, *-jɛndz*, and *-nɛndz* is unclear. They may partly be phonologically conditioned, as [j] is sometimes inserted between two vowels when they are joined together through affixation.⁷ However, they are also sometimes interchangeable, as in the following two examples.

- (139) *nur [mu idora-jɛndz] xalg qati tamoq xor=an*
 today [1SG.POSS work.unit-ADJ] person with food eat.NPST-1PL.NPST
 ‘Today I will eat with people from my work unit.’

- (140) *nur [mu idora-nɛndz] xalg qati tamoq xor=an*
 today [1SG.POSS work.unit-ADJ] person with food eat.NPST-1PL.NPST
 ‘Today I will eat with people from my work unit.’

3.4 Comparison with related languages

This section will present some RC examples from major Iranian languages and other Pamir languages in order to highlight some notable differences between them and Sarikoli.

⁷ For example, when the accusative plural suffix *-ɛf* attaches to a vowel-final word like *tɕɛrtɕi* ‘worker’, the pluralized word is *tɕɛrtɕijɛf* ‘workers’. Or when the genitive suffix *-an* attaches to it, it becomes *tɕɛrtɕijan*.

Data from Persian and Tajik, Shughni, Rushani, and Wakhi will be examined. When compared with other languages within the Iranian language family, the Sarikoli RC construction is very unique in terms of the ordering of RC and head noun, as well as the relativizers used.

3.4.1 Persian & Tajik

Persian belongs to the Western-Iranian branch of Iranian languages, and is the largest among the Iranian languages. As shown in the following example, as it is a verb-final language which places the RC after the common argument.

- (141) *ketab-i* *[ke be mæn dad-id]* *gomçode æst*
 book-INDEF [REL to me gave-2SG] lost is
 ‘The book [you gave to me] is lost.’ (Andrews 2007:209)

(This *-i* suffix is glossed as "indefinite" in the source, even though the free translation in English uses the definite determiner. In another source, Taghvaipour describes this *-i* suffix as a “particle that precedes restrictive RCs in Persian” (2005:12).)

Taghvaipour also notes, “Although Persian is a verb-final language, it has certain head-initial constructions such as Noun-Possessor, Noun-Adjective, and Noun-Relative Clause constructions” (2005:12). (142) is another example of the Persian RC, which places the RC after the common argument. In both of these examples, the relativizer *ke* introduces the RC, unlike the enclitic relativizers in Sarikoli.

- (142) *ust* *zæn-i* *[ke mæn dust.daræm]*
 she.be-PRES-3SG woman-RES [COMP I love.PRES.1SG]
 ‘She is the woman [that I love].’ (Taghvaipour 2005:12)

Another difference between Sarikoli and Persian RCs is that Persian makes a formal distinction between restrictive and non-restrictive RCs. The *-i* suffix is required for restrictive RCs, but not for non-restrictive ones (Comrie 1983:139).

- (143) *ali*, *[ke be firaz ræfteh.bud]*, *bærayæm nameh-i neveft*
 Ali [COMP to Shiraz go.PP.3SG] for.me letter-IND write.PST.3SG
 ‘Ali, who had gone to Shiraz, wrote me a letter.’ (Taghvaipour 2005:14)

As in Sarikoli, Persian RCs are externally-headed and may use the gap strategy. The verb within the RC can be marked for tense.

Tajik is another Western-Iranian language that is closely related to Persian. Tajik RCs are also placed after the common argument, even though it is a verb-final language. Relativizers occur at the beginning of the RC. When juxtaposed with the Persian examples above, it seems that the Persian *-i* is comparable with the Tajik *-e*, and the Persian *ke* is comparable with the Tajik *ki*.

(Like the *-i* suffix in Persian, the *-e* suffix is glossed in many different ways by various linguists. This source glosses this suffix as "relativizer".)

(144) *man stol-e, [ki latif soxt] xarid-am*
 I table-REL [COMP Latif build.PST] buy.PST-1SG
 'I bought a table [that Latif built]' (Stump 2012:1)

(145) *rahim mard-e, [ki malika a vay donotar ast], me-boç-ad*
 Rahim man-REL [COMP Malika than him smarter is] NPST-be.NPST-3SG
 'Rahim is the man [that Malika is smarter than]' (Stump 2012:1)

3.4.2 Shughni

Shughni is an Eastern-Iranian language that belongs to the Pamir language family. Among the Pamir languages, it is often grouped within the Shughni-Rushani subgroup (Arlund 2006; Edelman & Dodykhudoeva 2009; Payne 1989), so it is one of the most closely-related languages with Sarikoli. Because it is spoken in Tajikistan and Afghanistan, it has more direct contact with Western-Iranian languages (Tajik and Dari) than Sarikoli does. It is also a verb-final language.

Shughni RCs are optionally marked by the preverbal relativizer *tsa* (Edelman & Dodykhudoeva 2009:811). *Tsa* is a native particle that marks subordinate clauses (Payne 1989:441) and may be used for both restrictive and non-restrictive RCs, as shown below. (146) shows that the RC may contain a resumptive pronoun, which refers to the head noun. The resumptive pronoun takes the appropriate case marker (Edelman & Dodykhudoeva 2009b:811), based on the role that the common argument plays within the RC. As in Persian and Tajik, RCs in Shughni are placed after the head noun.

(146) *jid-ik-u tforik [idi vegaj-um di ar bozor tsa wint]*
this-very-he man [that yesterday-I him at bazaar tsa saw]
 ‘This is the man whom I saw at the bazaar yesterday.’ (Edelman & Dodykhudoeva 2009b:812)

(147) *jid tu puxok, [narm-at safed tsa], dundga xizmat tu-rd kixt, jo naj?*
this your dress [soft-and white tsa] such service you-to make or not
 ‘This dress of yours, which is soft and white, will it serve you this long as well?’
 (Edelman & Dodykhudoeva 2009b:812)

3.4.3 *Rushani*

Rushani is another language that belongs to the Shughni-Rushani subgroup within the Pamir languages. Like Shughni, it uses the native subordination particle *tsa* before the verb within the RC. In addition, when constructing externally-headed RCs, it uses the native subordinating conjunction *dide* as a relativizer (Payne 1989:442), which occurs after the head noun and introduces the RC.

(148) *wað qalam-en-an [dide ta-re-m tsa datfug] unde vid*
those pen-PL-3PL [REL you-to-1SG tsa give.PST] here be.PST
 ‘Those pens [that I gave you] were here.’ (Payne 1989:442)

3.4.4 *Wakhi*

Wakhi is the only other Pamir language spoken in China. It is mutually unintelligible with Sarikoli (Arlund 2006). Larger populations of Wakhi speakers reside in Afghanistan, Tajikistan, and Pakistan; the Wakhi speakers in China are heavily concentrated in Dafdar village (Lewis 2013; Kreutzmann 2003).

In Wakhi RC constructions, the head noun may be placed before or after the RC, as well as being placed within the RC to form an internally-headed RC. RCs contain the native subordinate marker *tsə*, which is like the *tsa* in Shughni and Rushani (Payne 1989:441). When the RC follows the head noun to form a head-initial construction, Wakhi also uses the relativizer *ki*, which is borrowed from Tajik:⁸

⁸ The source of these Wakhi RC examples did not provide morpheme-by-morpheme or word-by-word glosses. Only the free translation was provided.

(149) *a-ja nayd-i [ki a-jəm mərda-v gur tsə kart-əv]*
'that night [when they buried this corpse]' (Bashir 2009:850)

Unlike the other Iranian languages presented above, Wakhi also allows the head-final RC construction by placing the RC before the head noun. The following head-final RC example shows that Wakhi also has the participial RC, which are formed with the present or perfect participle (Bashir 2009:850):

(150) *[sk-a vadek tuk-kizg] xalg-i zj mormor*
'The man [walking along the road] is my friend.' (Bashir 2009:851)

In Wakhi RCs, the fullest statement of the common argument can either appear in the matrix clause or within the RC. If it appears within the RC, a resumptive pronoun is used in the matrix clause, as in (151).

(151) *[pard ja dəraxt dər-əm tsə tu] jaw-i kot-əv*
'They dug up the tree [that was here last year].' (Bashir 2009:850)

3.4.5 Conclusion

The data presented in this section demonstrate that some other Iranian languages, even Pamir languages that are closely related to Sarikoli, construct RCs very differently from Sarikoli.

Most strikingly, in terms of the ordering of the RC and the head noun, all five of the languages included in this section place the RC after the head noun (head-initial),⁹ whereas Sarikoli places the RC before the head noun (head-final). In addition to the languages presented in this section, other Iranian languages, such as Kurdish, Talysh, and Pashto, are languages in which the RC follows the head noun (Dryer 2013; Tegey & Robson 1996:206 for Pashto). Sarikoli, being true to the OV type, may be the only Pamir language in which all NP modifiers precede the head noun. This may also be a very rare phenomenon within the Iranian language family as a whole.

⁹ Wakhi has both head-initial and head-final RC constructions.

Another interesting difference between Sarikoli and the other Pamir languages is the form of the relativizers. As introduced in the descriptions about Shughni, Rushani, and Wakhi, many Pamir languages mark RCs with a particle which has \widehat{ts} as the onset; variations of this particle include *tsa*, *tse*, *tsə*, *tsəj*, and *tʃə* (Payne 1989:441). While Sarikoli has this *tsa* particle as well, it is not used at all for marking RCs, but for marking the protasis in conditional sentences (see section 5.1.4). Also, Wakhi uses a relativizer borrowed from Tajik, *ki*, and Rushani uses a native subordinating conjunction, *dide*. Sarikoli does not use these particles, but has enclitics (= *dzendz* and = *itfuz*) that seem to be unique relativizers among the Pamir languages.

As of now, there is no solid explanation that accounts for why these peculiarities exist in the Sarikoli RC construction. One possible explanation for the ordering of the RC and head noun is that it is a contact phenomenon, because of the linguistic environment surrounding Sarikoli speakers. Sarikoli is spoken in China, where the national language, Mandarin, uses the head-final RC construction. But more significantly, being situated in the Xinjiang Uyghur Autonomous Region, Sarikoli is surrounded by various Turkic languages, primarily Uyghur, Kyrgyz, Uzbek, and Kazakh. All of these Turkic languages are OV languages that place the RC before the head noun. Uyghur is the lingua franca among the minority peoples of Xinjiang, and many Sarikoli speakers speak Uyghur as their second language (Arlund 2006). It may be the case that heavy contact with these Turkic languages has influenced the placement of the RC, as syntactic phenomena can be the result of borrowing through language contact (Thomason & Kaufman 1988).

CHAPTER 4

Complement Clauses

In a complement clause (CC), one proposition takes another proposition as one of its arguments. Dixon (2006) claims three basic properties of CCs: 1) having the internal constituent structure of a clause, 2) functioning as a core argument of a higher clause, and 3) describing a proposition, involving someone involved in an activity or state.

Sarikoli has at least one CC construction which fulfills all three of these requirements: the finite complement, which has the most structural similarity to a main clause. The other two constructions are non-finite complements with more limited grammatical marking, and may be analyzed as complementation strategies (but not CCs) by some. Nevertheless, their internal constituent structure does resemble that of a clause to some extent, and they do fulfill the latter two requirements.

Dixon also points out that, while any verb can be used within a CC, verbs that can take a CC as an argument are very restricted (2006:5). From the restricted set of verbs, the semantic type of the verb seems to have some influence over determining which type of CC construction will be used. Verbs of perception, knowledge, thinking, and speaking generally take nominalized complements, whereas a smaller set of verbs, including verbs of planning, liking, causation, and allowance, take infinitival complements. In this thesis, CCs are enclosed in angle brackets.

4.1 Types of Sarkoli CCs

Sarikoli has four types of CCs: a) nominalized complements with a nominalizer, b) infinitival complements, c) post-verbal complements used for reporting speech, thoughts, dreams, etc., introduced with a subordinate clause particle, and d) finite complements with no marking.

Sarikoli has two constructions for reporting direct speech and one construction for reporting indirect speech. Although both direct and indirect quotation can be used, the direct quotation is preferred and indirect quotation is rarely used (Arlund & Ibrukhim 2013:176). The first option for reporting direct speech will be introduced in the unmarked finite complement section (4.1.3) and the other option will be introduced in the marked finite complement section (4.1.4). The reporting of indirect speech will be introduced in the nominalized complements section (4.1.1).

4.1.1 Nominalized complement

Although Sarikoli does not have a variety of finite complement constructions, it uses what Dixon describes as nominalization as a complementation strategy. Dixon defines nominalization as “a process by which something with the properties of a nominal can be derived from a verb or adjective, or from a complete clause” (2006:36). Because this complementation strategy relies on the suffix *-i*, whose various functions may cause confusion, the *-i* will be introduced first.

4.1.1.1 The *-i* suffix

The *-i* suffix is used for several different purposes. Only four of them will be introduced in this thesis, although it is likely that there are more. First, it is sometimes used like a postposition on nouns that indicate time, and expresses meanings such as ‘at’, ‘on’, ‘in’, or ‘during’ that time. It is often found on parts of the day, names of seasons, and names of holidays. For this usage, the *-i* is glossed as a temporal suffix, as in the examples below.

(152) *tom pejcin-i* *karwo wazevdz*
 then late.afternoon-TEMP crow return.NPST
 ‘Then in the late afternoon, the crow returned.’

(153) *mjendz-i* *kalo pwojd = itfuz dzuj*
 summer-TEMP sheep herd.INF = REL place
 ‘place where sheep are herded during the summer’

- (154) *kur**bu**n-i kalo laka spejd kalo tor tsem vεð*
 Korban-TEMP sheep put.NPST white sheep black eye be.PRF
 ‘On Korban day, the sheep has to be a white sheep with black eyes.’

Second, the *-i* suffix can also be used as a derivative suffix for forming adverbs from adjectives. There are many adverbs that do not have this suffix, but adverbs that are formed from adjectives usually take it. Normally, adjectives cannot immediately precede verbs because they precede the nouns they modify. But the adverbs that are derived from adjectives with the *-i* suffix immediately precede the verbs they modify, as in the following examples.

- (155) *oqubut**ç**i-**ç**ejl=**af** utf tafsun-i **ç**umand sut*
 student-NOM.PL=3PL.PST very enthusiastic-ADV study be.PST
 ‘The students studied very enthusiastically.’

- (156) *nur**æ**un **ç**alg-**ε**f=**ir** tfardz-i ka=**m***
 many person-ACC.PL=DAT good-ADV do.NPST=1SG.NPST
 ‘I will do good to many people.’

- (157) *ju digaru-**ç**εf=**ir** b**ç**ε-adab-i kaxt*
 3SG.NOM other.person-ACC.PL=DAT NEG-polite-ADV do.3SG.NPST
 ‘He treats other people rudely.’

- (158) *i maθ i ruz xiθp rust-i iθtf*
 one day one life wolf true-ADV come.NPST
 ‘One day, a wolf really came.’ (Arlund & Ibrukhim 2013:315)

Third, the *-i* suffix can be used as another kind of derivative suffix, attaching to adjectives to form nouns. When used for this purpose, *-i* is glossed as a nominalizer. After adjectives have been nominalized by the *-i* suffix, they can function like regular nouns; the nominalized element in (159) is the subject, and the ones in (160) & (161) are predicate nominals.

- (159) *sovdz-i maf=**ir** ajut vird*
 green-NMLZ 1PL.ACC=DAT life bring.3SG.NPST
 ‘Greenness brings us life.’ (Arlund & Ibrukhim 2013:315)

(160) *ammo mu djest karim wi tarabex-i*
 but 1SG.POSS friend Karim 3SG.ACC opposite-NMLZ
 ‘But my friend, Karim, is the opposite.’ (Arlund & Ibrukhim 2013:166)

(161) *insun avon ej ləwr baxt salomat-i*
 humankind for most big blessing healthy-NMLZ
 ‘The biggest blessing for human beings is health.’ (Arlund & Ibrukhim 2013:305)

The fourth purpose of the *-i* suffix will be introduced in the next section.

4.1.1.2 Nominalized complement (with the *-i* suffix)

The *-i* suffix can also occur in one of the complementation strategies, and plays a role similar to that of a complementizer. Its function in this complementation strategy seems to be most closely related to the third function presented above: nominalizing. The *-i* attaches to a verb in its infinitive stem so that it can become an argument of the main verb.

The other part of this complementation strategy is the genitive suffix *-an*, which attaches to the subject of the nominalized complement if it is animate. So structurally, the subject of the embedded clause is marked like a possessor of an NP. Because the embedded verb is nominalized, the entire embedded clause after the subject can become the possessed item. This entire nominalized complement can function as a regular argument, just like other NPs. In the following examples, the nominalized complement fills the object slot, between the subject and the verb. Note, however, they do not take the differential object marker (*a=*), presumably because of their length and complexity. Also, the nominalized complement does not have pronominal agreement clitics.

(162) *alima malum <gul-an bədzin tid-i> wazond*
 Alima teacher <Gul-GEN Beijing go.INF-NMLZ> know.3SG.NPST
 ‘Teacher Alima knows <Gul’s going to Beijing>.’

(163) *m-ono <i xalg-an az tjanza dejt-i> wand*
 1SG.POSS-mother <one person-GEN from window enter.INF-NMLZ> see.PST
 ‘My mother saw <a person’s coming in through the window>.’

(164) *putxu* <*χu askar-ɛf-an wef ratsist-i*> *wand*
king <own soldier-ACC.PL-GEN 3PL.POSS run.away.INF-NMLZ> see.PST
‘The king saw <his soldiers’ running away>.’

(165) *kuud* <*χu sujib-an a=dver at tfejg-i*> *xud*
dog <own owner-GEN ACC=door open do.INF-NMLZ> hear.PST
‘The dog heard <his owner’s door-opening>.’

As shown in these examples, the main verb in the matrix clause is finite and can be marked for any tense, independently from the embedded event. But since the verb in the nominalized complement is in the infinitive stem, it does not indicate when the event is taking place. However, if it is important to specify the time, time words can be added to the nominalized complement. As they do in regular sentences, the time word in the embedded event usually immediately follows the subject:

(166) *qandik* <*olim-an xjeb fand ðod-i*> *wazond*
Qandik <Olim-GEN yesterday lie give.INF-NMLZ> know.3SG.NPST
‘Qandik knows <Olim’s lying yesterday>.’

(167) *m-ono* <*mu-jan fitf kinu tçixt-i*>
1SG.POSS-mom <1SG.POSS-GEN now movie watch.INF-NMLZ>

na-wazond
NEG-know.3SG.NPST

‘My mother does not know <my watching movies right now>.’

(168) *m-ono* <*mu-jan nur waz=ir wux na-weðd-i*>
1SG.POSS-mom <1SG.POSS-GEN today sheep=DAT grass NEG-put.INF-NMLZ>

wand
see.PST

‘My mother saw <that I did not feed the sheep today>.’

In addition to the object role, the nominalized complement can also function as the subject, in both S and A roles. As a subject, the nominalized complement simply fills the subject slot at the beginning of the matrix clause. The following is an example of a nominalized complement in the S role. Notice that it has two subjects, both of which are nominalized complements:

- (169) < *kəwr-an wi tsem tçi wand set-i* >,
 < blind.person-GEN 3SG.POSS eye at see.INF become.INF-NMLZ >
 < *tjuun-an wi bəwl tçi xid set-i* > *putuun pa*
 < deaf.person-GEN 3SG.POSS ear at hear.INF become.INF-NMLZ > all to
dəwlat iw sut
 country one become.PST
 ‘<The blind person’s eyes becoming able to see>, and <the deaf person’s ears becoming able to hear> became known in the whole country.’

The nominalized complement also occurs at the very beginning of the matrix clause when it fills the A role, but the matrix clause also takes a direct object, which is usually marked with the differential object marker (*a=*):

- (170) < *pif-an wi marg-i* > *a=mu jaχ tçi niwd weðd*
 < cat-GEN 3SG.POSS die.INF-NMLZ > ACC=1SG.POSS sister at cry.INF put.PST
 ‘<The cat’s dying> made my sister cry.’

- (171) < *mu-jan pa aftovuuz nalist na-tfejg-i* > *a=mu tçer*
 < 1SG.POSS-GEN to bus sit.INF NEG-do.INF-NMLZ > ACC=1SG.ACC work
tar zabu weðd
 toward late put.PST
 ‘<My not being able to catch the bus> made me late for work.’

In (168) & (171), notice that the nominalized complement is negated. It is negated in the same way as in a matrix clause, by attaching the *na-* prefix to the verb. Even though the nominalized complement functions like an NP as a whole, its internal structure is very similar to that of an independent clause.

Like regular NPs, nominalized complements can also directly take the dative marker, the enclitic *=ri*. Initially, the dative marker may seem to be a part of this complementation strategy because it occurs so much with nominalized complements. However, it is only because the embedded clauses are nominalized that it freely takes the dative marker whenever required by the main verb in the matrix clause. For example, the verbs ‘believe’, ‘be surprised at’, and ‘be happy at’ normally occur with dative objects:

(172) *ju* *χabar-ir* *ifandz* *tfəwg*
 3SG.NOM news-DAT believe do.PST
 'He believed the news.'

(173) *adi* *χabar-ir* *dzam* *χalg* *ejrun* *rejd*
 this news-DAT all people surprise leave.PST
 'Everyone got surprised at this news.'

(174) *adi* *χabar-ir* *tɕəd-nəndz-χejl=af* *dzam* *χuf* *sut*
 this news-DAT house-ADJ-NOM.PL=3PL.PST all happy become.PST
 'The family all became happy at this news.'

Likewise, when taking nominalized complements as clausal objects, these verbs occur with objects that are marked as dative:

(175) *bjewazan* <*χu* *puɕs-an* *wi* *soq* *set-i*> = *ri* *ifandz*
 widow <own son-GEN 3SG.POSS good become.INF-NMLZ> = DAT belief

tfəwg
 do.PST
 'The widow believed <her son's getting better>.'

(176) <*olim-an* *ar* *bedzin* *donifko* *tid-i*> = *ri* *dzam* *χalg* *ejrun*
 <Olim-GEN INF Beijing university go.INF-NMLZ> = DAT all person surprise

rejd
 leave.PST
 'Everyone got surprised at <Olim's going to a university in Beijing>.'

(177) <*m-ono-jan* *poj* *wεðd-i*> = *ri*
 <1SG.POSS-mother-GEN yogurt put.INF-NMLZ> = DAT

tɕəd-nəndz-χejl=af *χuf* *sut*
 house-ADJ-NOM.PL=3PL.PST happy become.PST
 'The family became happy at <my mother's making yogurt>.'

A nominalized complement can also be used as a predicate nominal. As mentioned before, copulae are not required in Sarikoli, so the subject and the predicate may simply be juxtaposed. In the following examples, *jad* 'this' is the subject and the nominalized complement is the predicate.

(178) *jad* < *mu-jan* *awal gudur xipik* *tfejg-i*>
 this < 1SG.POSS-GEN first time flatbread make.INF-NMLZ >
 ‘This is <my first time making flatbread>.’

(179) *jad* < *maf-an* *awal gudur ar urumtçi jet-i*>
 this < 1PL.POSS-GEN first time in Urumqi come.INF-NMLZ >
 ‘This is <our first time coming to Urumqi>.’ (Arlund & Ibrukhim 2013:296)

Interrogative words are not used as complementizers for these nominalized complements. If the embedded clause contains an interrogative word, the suffix *-i* is still used for marking the nominalized complement, and the interrogative word occurs in-situ within the embedded clause. For example, time words usually immediately follow the subject in a regular sentence, and *tfum* ‘when’ is placed in that same slot within the embedded clause. Below are examples of nominalized complements that contain a variety of interrogative words, such as ‘what’, ‘why’, ‘when’, ‘where’, ‘whom’, and ‘to whom’. Notice that, if the complement is long or complicated, it may optionally be pre-posed, so that the constituent order becomes OSV if the nominalized complement is a direct object.

(180) *waz* < *ta-jan* *maður tsejz* *χig-i*> *wazon = am*
 1SG.NOM < 2SG.POSS-GEN noon what eat.INF-NMLZ > know.NPST = 1SG.NPST
 ‘I know <what you ate for lunch>.’

(181) *waz* < *ta-jan* *tsejzir tilfon na-zox-t-i*>
 1SG.NOM < 2SG.POSS-GEN why phone NEG-grab.INF-NMLZ >

wazon = am

know.NPST = 1SG.NPST

‘I know <why you didn’t answer the phone>.’

(182) *waz* < *ta-jan* *tfum ðandun tfutka* *ðod-i*>
 1SG.NOM < 2SG.POSS-GEN when teeth brush.teeth give.INF-NMLZ >

wazon = am

know.NPST = 1SG.NPST

‘I know <when you are brushing your teeth>.’

(183) *waz* < *ta-jan* *χu bojlɛq kudzur naymig-i* >
 1SG.NOM < 2SG.POSS-GEN own treasure where hide.INF-NMLZ >

wazon = am

know.NPST = 1SG.NPST

'I know <where you hid your treasure>.'

(184) *waz* < *ta-jan* *a = tçi* *wand-i* >
 1SG.NOM < 2SG.POSS-GEN ACC = who.ACC see.INF-NMLZ >

wazon = am

know.NPST = 1SG.NPST

'I know <who you saw>.'

(185) *waz* < *ta-jan* *a = samɸut tçi = ri* *ðod-i* >
 1SG.NOM < 2SG.POSS-GEN ACC = gift who.ACC = DAT give.INF-NMLZ >

wazon = am

know.NPST = 1SG.NPST

'I know <who you gave the gift to>.'

Even stative clauses, with no action involved at all, can become a nominalized complement with the *-i* suffix. The infinitive stem of the 'be' verb, *vid*, is used as the verb of the embedded clause that receives the nominalizing suffix. Because of the contexts in which the stative nominalized complements are used, they often contain interrogative words; however, as shown in (187), this is not required.

(186) *waz* < *ta-jan* *tfoj* *vid-i* > *wazon = am*
 1SG.NOM < 2SG.POSS-GEN who.NOM be.INF-NMLZ > know.NPST = 1SG.NPST
 'I know <who you are>.'

(187) *waz* < *ta-jan* *malum* *vid-i* > *wazon = am*
 1SG.NOM < 2SG.POSS-GEN teacher be.INF-NMLZ > know.NPST = 1SG.NPST
 'I know <you are a teacher>.'

(188) < *wi-jan* *tsarang uzgurus* *vid-i* > *az* *wi*
 < 3SG.ACC-GEN how situation be.INF-NMLZ > from 3SG.ACC

pars = an

ask.NPST = 1PL.NPST

'We ask her <how her situation is>.'

(189) *waz uzir tamaʃ=ir <tudʒiktʃa tej-an tsarang vid-i>*
 1SG.NOM now 2PL.ACC=DAT <Tajik wedding-GEN how be.INF-NMLZ>

lev=am

say.NPST=1SG.NPST

‘I will now tell you all <how a Tajik wedding is>.’

(190) *<wi-jan ar pa qawat tsejz narsa vid-i> lev*
 <3SG.ACC-GEN every at floor what thing be.INF-NMLZ> say.NPST

‘Tell <what is on each of its floor>.’ (Arlund & Ibrukhim 2013:90)

A nominalized complement may contain a further-embedded infinitival complement. In the following example, the verb ‘eat’ is the infinitival complement, which is embedded in the nominalized complement with ‘be planning’, which is embedded in the matrix clause with ‘know’.

(191) *<ta-jan maður tsejz xig mejdʒ vid-i> waz*
 <2SG.POSS-GEN noon what eat.INF plan be.INF-NMLZ> 1SG.NOM

wazon=am

know.NPST=1SG.NPST

‘I know <what you are planning to eat for lunch>.’

All of the examples of nominalized complements presented so far have had a genitive marker or a possessive pronoun for the subject at the beginning of the embedded clause. However, in very rare cases, there are exceptions to this pattern, and the subject of the embedded clause is not marked in any special way:

(192) *xjeb xob olim <awu ðod-i> xjeðdʒ*
 yesterday night Olim <rain fall.INF-NMLZ> hear.PRF
 ‘Last night, Olim heard <rain falling>.’

(193) *<suktak wejrun set-i> a=mu jaχ tçi niwd weðd*
 <toy break become.INF-NMLZ> ACC=1SG.POSS sister at cry.INF put.PST
 ‘<The toy’s breaking> made my sister cry.’

In these two examples, ‘rain’ and ‘toy’ do not receive the genitive marker, and the infinitive stem of the verb and the nominalizer *-i* suffix are the only indicators of the nominalized complement. One possible explanation for this anomaly is that ‘rain’ and ‘toy’

are both inanimate, as opposed to the human and animal subjects in all of the previous examples. Further analysis is needed to figure out why some subjects of nominalized clauses do not take the genitive marker.

Although indirect quotation is not commonly used, it is possible to indirectly report what someone said by making it a nominalized complement. As with other nominalized complements, indirect speech must also begin with the embedded subject (marked as the possessor) and end with the nominalizer *-i*. While this nominalized form is effective for clearly marking the boundaries of the quoted speech, it may sometimes be difficult to understand because it is limited in expressing tense or modality. As the following examples show, the quoted speech completely relies on time words to indicate time. Thus, it seems natural that quoting direct speech is preferred.

- (194) *<muu-jan xjeb kudzur vid-i-ik> malum az mu*
<1SG.POSS-GEN yesterday where be.INF-NMLZ-???)> teacher from 1SG.ACC
parst
ask.PST
 ‘The teacher asked me <where I was yesterday>.’

- (195) *ju <χu wazɛvd-i> =ri qasam tɕəwɔ*
3SG.NOM <own return-NMLZ> =DAT swear do.PST
 ‘He promised <that he will come back>.’

- (196) *waz = am <χu tɕɛr tɕeig meidɔ vid-i> wi = ri*
1SG.NOM = 1SG.PST <own work do.INF plan be.INF-NMLZ> 3SG.ACC = DAT
lɛvd
say.PST
 ‘I told him that <I am planning to do my work>.’ (Arlund & Ibrukhim 2013:20)

(197) *mu vurud tci tilfon <m-ono-jan wi kasal*
 1SG.POSS brother at telephone <1SG.POSS-mother-GEN 3SG.POSS sick

set-i> am <mu-jan mu zjald jet-i>
 become.INF-NMLZ> and <1SG.POSS-GEN 1SG.POSS quick come.INF-NMLZ>

levd
 say.PST

‘On the phone, my brother said <my mother’s becoming sick> and <my quick coming>.’ (free. ‘... said that my mother became sick and that I should come quickly.’) (Arlund & Ibrukhim 2013:301)

4.1.2 Infinitival complement

Sarikoli also has an infinitival complement construction, which is another non-finite complementation strategy (Velupillai 2012:318). As with the nominalized complement, it has less structural similarity to a main clause, as it is not inflected for tense or modality and does not use pronominal agreement clitics. The main structural difference between this infinitival complement and the nominalized complement is that the infinitival complement is not marked by any special affixes, whereas the nominalized complement always ends with the nominalizer suffix *-i*. In addition, the infinitival complement rarely contains an explicit subject, whereas the nominalized complement always contains a subject, whether it is the same subject or a different subject as the matrix clause. Among the examples presented in sections 4.1.1.2, 33 out of 33 examples of nominalized complements contained an explicit subject, 31 of which were marked as the possessor with the genitive marker *-an* or a possessive pronoun. In contrast, among the examples in 4.1.2, only 6 out of 35 examples of infinitival complements contained explicit subjects; again, the explicit subjects in infinitival complements are also marked as possessors with the genitive marker *-an*.

Apart from this structural difference, the infinitival complement is sometimes determined by where the embedded event is placed in time. If the embedded event has already happened, it may receive the nominalizer *-i*, whereas if the embedded event has not happened yet, the embedded event is simply an infinitival complement. Compare the following example with (175):

(198) *bjewazan* <*χu puts-an wi soq set*> = *ir* *ifandz*
 widow <own son-GEN 3SG.POSS good become.INF> = DAT believe

kaxt

do.3SG.NPST

‘The widow believes <that her son will get better>.’

Some of the infinitival complements also differ from the nominalized complement in the set of main verbs it can be an argument of. Verbs that take infinitival complements as an object are more limited than verbs that take nominalized complements. They include: ‘believe’, ‘be willing’, ‘be satisfied with’, ‘plan’, ‘cause’, ‘allow’, ‘like’, and ‘want’. Infinitival complements can also function as the subject of a predicate adjective or nominal. The following are examples of this construction with ‘be willing’ and ‘be satisfied with’ as the main verb. For the verb ‘be willing’, the presence of the dative marker signals that the embedded subject is different from the matrix clause subject.

(199) *aqlia* <*pa jatoq nalist*> *tjombd*
Aqlia <at dormitory live.INF> be.willing.3SG.NPST
 ‘Aqlia is willing <to live in the dormitory>.’

(200) *wođ=af* <*mu-jan a=wi jod*> = *ir*
 3PL.NOM=3PL.PST <1SG.POSS-GEN ACC=3SG.ACC man.marry.INF> = DAT

na-tjimbđ

NEG-be.willing.PST

‘They were not willing for <me to marry her>.’

(201) *waz* <*a=εη qimat baron zox̄t*> = *ir* *rizu*
 1SG.NOM <ACC=most expensive dress buy.INF> = DAT satisfied
 ‘I am satisfied with <buying the most expensive dress>.’

The verb ‘plan’ may also take an infinitival complement as its object:

(202) *ju* <*ar qotnof mejmunχuno alid*> *mejdz*
 3SG.NOM <at traffic hotel reside.INF> plan
 ‘He plans <to stay at the Traffic Hotel (Jiaotong Bingguan)>.’

(203) *maf* <*pugan adi t̄fat parađod*> *mejdz*
 1PL.NOM <tomorrow this cow sell.INF> plan
 ‘We plan <to sell this cow tomorrow>.’

Notice that another major difference between the two non-finite complements is the expression of the subject within the embedded clause. In the nominalized complement, the subject may or may not be the same as that of the matrix clause. The subject is usually explicitly expressed and marked as a possessor, whether or not the subject is the same as one in the matrix clause. On the other hand, the subject of an infinitival complement is usually the same as that of the matrix clause, so there is often no need for the subject to be expressed again in the embedded clause. Thus, the subject within the infinitival complement that is coreferential with the subject of the main clause is often omitted by ellipsis. But if the subject within the infinitival complement is different from the subject of the matrix clause, it is expressed and marked as a possessor, as in (200).

In some cases, the infinitival complement has an object of its own and the matrix clause also has a direct or indirect object other than the clausal argument. This does not seem to cause too much confusion because the verb in the embedded clause immediately follows its object, and the verb in the matrix clause also immediately follows its object (i.e., the infinitival complement). Also, the other object in the matrix clause is marked with the differential object marker (*a=*) or the dative marker (*=ir/=ri*). The examples of ‘cause’ below illustrate this:

(204) *malum oquwutçi-jef-ir* <*kutub juðzox*t> *ramud*
 teacher student-ACC.PL-DAT <book memorize.INF> cause.PST
 ‘The teacher caused the students <to memorize the book>.’

(205) *açmat istuð=ir* <*kastun insivd*> *ramej=am*
 Ahmet craftsman=DAT <suit sew.INF> cause.NPST=1SG.NPST
 ‘I will cause craftsman Ahmet <to sew a suit>.’ (Arlund & Ibrukhim 2013:251)

(206) *m-ono* *mu=ri* <*adi tfer nur adu tfejg*> *ramud*
 1SG.POSS-mother 1SG.ACC=DAT <this matter today finish do.INF> cause.PST
 ‘My mother caused me <to finish this work today>.’ (Arlund & Ibrukhim 2013:255)

‘Allow’ is another verb that takes infinitival complements. As shown in the ungrammatical example (209), the verb *latfəwɔŋ* ‘allow’ does not take nominalized complements.

(207) *m-oto* *a=mu* <*qimat* *uzzum araq broxt*>
 1SG.POSS-dad ACC=1SG.ACC <expensive grape wine drink.INF>

na-lakaxt

NEG-put.3SG.NPST

'My father does not allow me <to drink the expensive wine>.'

(208) *malum* *a=mu* <*pa xalo* *tid*> *latfəwg*
 teacher ACC=1SG.ACC <to bathroom go.INF> put.PST
 'The teacher allowed me <to go to the bathroom>.'

(209) **malum* <*mu-jan* *pa xalo* *tid-i*> *latfəwg*
 teacher <1SG.POSS-GEN to bathroom go.INF-NMLZ> put.PST
 'The teacher allowed me <to go to the bathroom>.'

The concept of 'like' can be expressed in two different ways, both of which can take either an NP or an infinitival complement as the object. First, the person who does the liking can be marked as dative, followed by the object of liking, and then followed by the word 'happy', *xuf*. Literally, it means that the object is pleasing to the person who likes it, as illustrated below.

(210) *zulfia=ri* <*poj* *furd*> *xuf*
 Zulfia=DAT <yogurt slurp.INF> happy
 'Zulfia likes <to eat yogurt>.' (lit. '<Eating yogurt> is pleasing to Zulfia.')

(211) *sarmsoq=ir* <*bejð* *levd*> *xuf*
 Sarmsok=DAT <song say.INF> happy
 'Sarmsok likes <to sing songs>.' (lit. '<Singing songs> is pleasing to Sarmsok.')

(212) *šamšer=ir* <*xob-i* *kuutub tfixt*> *xuf*
 Šamsher=DAT <night-TEMP book watch.INF> happy
 'Šamsher likes <to read books at night>.' (lit. '<Reading books> at night is pleasing to Šamsher.')

This construction may also take a complement that has a different subject from the matrix clause. In such cases, the different subject is explicitly expressed and marked as a possessor:

- (213) *mu-ri* < *alima-jan bejt levd* > *χuf*
 1SG.ACC-DAT < Alima-GEN song say.INF > happy
 'I like <Alima's singing>. (lit. '<Alima's singing> is pleasing to me.)'
- (214) *wi-ri* < *qandik-an jet* > *χuf*
 3SG.ACC-DAT < Qandik-GEN come.INF > happy
 'He likes <Qandik's coming>.' (lit. '<Qandik's coming> is pleasing to him.)'
- (215) *zulfia-ri* < *tudzik χalg-an usul tfejg* > *χuf*
 Zulfia-DAT < Tajik people-GEN dance do.INF > happy
 'Zulfia likes <Tajik people's dancing>. (lit. '<Tajik people's dancing> is pleasing to Zulfia.)'

The other way to express 'like' is to use the verb 'see', preceded by the word 'good', preceded by the subject and object of liking. Recall that the verb 'see' usually takes nominalized complements rather than infinitival complements (as in (163), (164), and (168)); but when used together with *tjardz* 'good' to mean 'like', it takes infinitival complements. *tjardz wand* 'like' seems to be regarded as a separate lexical entry from the regular *wand* 'see'. This 'like' construction looks very similar to the other verbs that take infinitival complements as objects. As shown in (219), this construction also allows complements with a subject that is different from the matrix clause, and marks the embedded subject with the genitive marker.

- (216) *waz* < *sajoat tfejg* > *tjardz wejn = an*
 1SG.NOM < travel do.INF > good see.NPST = 1SG.NPST
 'I like <to travel>.'
- (217) *alima* < *pa maktab tid* > *tjardz wand*
 Alima < to school go.INF > good see.3SG.NPST
 'Alima likes <to go to school>.'
- (218) *maf* < *az digaru-jef qusur zuvost* > *tjardz*
 1PL.NOM < from other.person-ACC.PL bad.habits find.INF > good

wejn = an
 see.NPST = 1PL.NPST
 'We like <to find bad habits from other people>.' (Arlund & Ibrukhim 2013:208)

- (219) *waz* <*alima-jan bejt levd*> *tçardz wejn-am*
 1SG.NOM <Alima-GEN song say.INF> good see.NPST-1SG.NPST
 'I Like <Alima's singing>.'

The concept of 'desire' is expressed in a peculiar construction, through the word *dil* 'heart'. In this construction, the infinitival complement is placed at the end of the sentence. The following example literally means, 'Qandik's heart is to eat yogurt':

- (220) *qandik dil* <*poj furd*>
 Qandik heart <yogurt slurp.INF>
 'Qandik wants <to eat yogurt>.'

Initially, the word *dil* may appear to be functioning as a verb that means 'want', especially because there is no finite verb in the sentence. However, I will argue against that analysis, for the following reasons. First, *dil* is a noun; even when it is used to express 'want', it does not come in four different stems, as regular verbs do. Second, when *dil* is used in the 'want' construction, it occurs before the object, not word-finally as regular verbs do. Third, pronouns that precede *dil* are always in the possessive form, which suggests that *dil*, not the pronoun, is the actual subject in this type of construction. The person who does the wanting is expressed as the possessor of the 'heart'. Note, however, that only pronouns are marked as possessors and proper nouns are simply juxtaposed with *dil*, the possessed item (as in (220), (246), and (247)).

For these reasons, I will analyze this *dil* as the subject, and the infinitival complement as the predicate nominal. Although there is no verb or copula in the sentence, this is acceptable in Sarikoli because the subject and the predicate nominal can simply be juxtaposed to form a complete sentence. Below are more examples of this *dil* construction.

- (221) *mu dil* <*tçardz læwr maktab-jef xojd*>
 1SG.POSS heart <good big school-ACC.PL read.INF>
 'I want <to study at big, good schools>.' (lit. 'My desire is <to study at big, good schools>'.)

- (222) *wef dil* <*mofin zoxt*>
 3PL.POSS heart <car buy.INF>
 'They want <to buy a car>.' (lit. 'Their desire is <to buy a car>'.)

(223) *wi dil <χu djest-ef qati towp skit tfejg>*
 3SG.POSS heart <own friend-ACC.PL with ball play do.INF>
 ‘He wants <to play ball with his friends>.’ (lit. His desire is <to play ball with his friends>.)

(224) *mu dil <nurɣun at ɖar dzuj-ef tid>, <sajoat tfejg>*
 1SG.POSS heart <many and far place-ACC.PL go.INF> <travel do.INF>
 ‘I want <to go to many far places> and <travel>.’ (lit. ‘My desires are <to go to many far places> and <to travel>.’)

The above construction with the embedded infinitival construction does not work for expressing desire with a different subject. For example, if a mother is expressing her desire for her son to study in Beijing, she will say:

(225) *mu dil <χu puuts ar bedzin xajond>*
 1SG.POSS heart <own son to Beijing send.to.study.INF>
 ‘My desire is <to send my son to Beijing to study>.’ (for: ‘I want my son to go to Beijing to study.’)

However, a different type of complement is used for expressing desire with a different subject in the embedded clause. It requires a subordinating conjunction and a finite complement, and will be described more in depth in the following section (4.1.4).

When infinitival complements function as the subject of the matrix clause and fill the S role, they do not have coreferential arguments outside themselves. Furthermore, as subjects, they do not seem to be limited to a restricted set of main verbs; rather, they tend to be the subject of predicate adjectives or nominals, so verbs are no longer necessary in the matrix clause. In the following examples, the infinitival complement is the subject of a predicate adjective in (226), the subject of a predicate nominal in (227), and the subject of a predicate adjective with a comparative construction in (228).

(226) *zumustun alo, <zumun tfi waχt ɖod> utf fujdo*
 winter TEMP <snow at time fall.INF> very beneficial
 ‘In the winter, <snow falling on time> is very beneficial.’ (Arlund & Ibrukhim 2013:121)

(227) < *a=watan sovdz tfejg* > *maf dzam-an wi wazifo*
 < ACC=country green do.INF > 1PL.POSS all-GEN 3SG.POSS duty
 ‘<Making the country green> is all of our responsibility.’ (Arlund & Ibrukhim 2013:262)

(228) < *inglestfa xumand set* > *az hansu-jef tudzik-ef=ir usun*
 < English study be.INF > from Han-ACC.PL Tajik-ACC.PL=DAT easy
 ‘<Learning English> is easier for Tajiks than for Han people.’

Note that these sentences, like the *dil* construction, do not have finite verbs. The infinitival complements in the subject role are juxtaposed with their predicate adjectives or nominals.

The infinitival complement can also fill the A role and take an object, as in this example:

(229) < *maθ paqad dzul batfo qati skit tfejg* > *a=χal aluk kaxt*
 < day all.day little child qati play do.INF > ACC=person tired do.3SG.NPST
 ‘<Playing with a little child all day> makes people tired.’

Finally, since the infinitival complement functions as an argument of the main verb, it can also directly take the dative marker =*ir*. This structure is used for several purposes. First, if the main verb normally requires a dative argument, then the infinitival complement will be marked as dative if it fills that role in the matrix clause. Second, this structure with the infinitival complement as a dative argument is also used by the hearsay construction. The hearsay construction uses *veðdz*, the perfect stem of ‘be’, as the main verb, and the information that is heard is expressed with an infinitival complement, which is the dative argument of *veðdz*:

(230) < *tom jad putxu niwd* > =*ir veðdz*
 < then this king cry.INF > =DAT be.PRF
 ‘Then this king cried (hearsay).’

(231) < *χu putxugi χu batfo-jef-ir narzambd* > =*ir veðdz*
 < own kingship own child-ACC.PL-DAT pass.on.INF > =DAT be.PRF
 ‘He gave his kingship to his sons (hearsay).’

- (232) <digar ar jurɥ ɣu dzun dzun tfejg> =ir vɛðdz
 <other INF place own life life do.INF> =DAT be.PRF
 ‘He lived his everyday life in another place (hearsay).’

Note that (231) contains two dative markers because the hearsay information, which is dative, contains a further-embedded dative argument.

Third, the infinitival complement is also marked as a dative argument in the purpose adverbial clause, which will be further described in section 5.1.3.

4.1.3 Unmarked finite complement (pre-verbal)

Pre-verbal finite CCs are not very common in Sarikoli. They contain finite verb stems and pronominal clitics, and no complementizer. While the following examples are not ungrammatical, the nominalized complement construction is preferred for expressing the same content.

- (233) <taw=at maður tsejz xug> waz wazon=am
 <2SG.NOM=2SG.NPST noon what eat.PST> 1SG.NOM know.NPST=1SG.NPST
 ‘I know <what you ate for lunch>.’

- (234) <taw=at a=sambut tçi=ri ðud> waz
 <2SG.NOM=2SG.PST ACC=gift who.ACC=DAT give.PST> 1SG.NOM

wazon=am
 know.NPST=1SG.NPST

‘I know <who you gave the gift to>.’

- (235) <olim xjeb fand ðud> qandik wazond
 <Olim yesterday lie give.PST> Qandik know.3SG.NPST
 ‘Qandik knows <that Olim lied yesterday>.’

Pre-verbal finite CCs are also used as one of the strategies for direct quotation. When reporting direct speech, the quoted speech may simply occur in the slot where objects are normally placed, followed by the verb *levd* ‘say’:

- (236) karɥwo <mejli> levd tujdz
 crow <OK> say.PST go.NPST
 ‘The crow said, “OK”, and left.’

(237) *i xalg wi=ri <joð, mu=ri jurdam ka> levd*
 one person 3SG.ACC=DAT <come 1SG.ACC=DAT help do.NPST> say.PST
 ‘A person said to him, “Come, help me.” (Arlund & Ibrukhim 2013:162)

(238) *ju fand-i <xiθp jot! xiθp jot!> levdz qiv tfəwydz*
 3SG.NOM lie-ADV <wolf come.PST wolf come.PST> say.NPST yell do.NPST
 ‘He falsely said and yelled, “Wolf came! Wolf came!”’ (Arlund & Ibrukhim 2013:315)

4.1.4 Marked finite complement (post-verbal)

Sarikoli also has a post-verbal finite CC construction, whose use is limited to reporting direct perception. This is the alternative strategy for quoting direct speech; the other one was introduced in 4.1.3. In this construction, the quoted speech is placed at the end of the matrix clause, and *iko* inserted between the quoted speech and the rest of the matrix clause. This construction is often preferred to the pre-verbal finite complement if the quoted speech is long, but short quotations may also be post-verbal, as in this example:

(239) *yubun putxu=ri levdz iko <məwl-χejl fand gap*
 shepherd king=DAT say.PRF SC <sheep-NOM.PL lie word
ka=jin>
 do.NPST=3PL.NPST>
 ‘The shepherd said to the king: <The sheep are lying>.’

The *iko* is a subordinating conjunction that is used to introduce object CCs of direct speech after verbs of speaking, thinking, dreaming, etc. This post-verbal CC for reporting direct perception is also commonly used for reporting hearsay or dreams. The embedded clause after *iko* is expressed as a fully finite clause, complete with finite verb stems and pronominal agreement clitics.

(240) *waz=am xid iko <ta xejx prud afto χu*
 1SG.NOM=1SG.PST hear.PST SC <2SG.POSS relative before week own
tej tfəwydz>
 wedding do.NPST>
 ‘I heard: <your relative did his wedding last week>.’

(241) *waz = am* *χυῶμ wandz iko* < *ar duxtur-i maktab = am*
 1SG.NOM = 1SG.PST dream see.PRF SC < to doctor-ADJ school = 1SG.PST

djeðdʒ>

enter.NPST >

'I dreamed: <I went to medical school >.'

(242) *ju* *uj* *tɸəwg iko* < *batfo-χejl = af* *ratsust* >
 3SG.NOM think do.PST SC < child-NOM.PL = 3SG.PST run.away.PST >

'He thought: <the children have escaped >.'

(243) *m-ono* *na-wazond* *iko* < *waz-ik* *kinu*
 1SG.POSS-mom NEG-know.3SG.NPST SC < 1SG.NOM-??? movie

tɸos = am >

watch.NPST = 1SG.NPST >

'My mother does not know: <I am watching movies >.'

(244) *ju* *qasam* *tɸəwg iko* < *ju* *inik* *wazafst*
 3SG.NOM swear do.PST SC < 3SG.NOM certainly return.3SG.NPST

joðd >

come.3SG.NPST >

'He promised: <he will certainly come back >.'

This finite CC construction with the subordinating conjunction *iko* is also used for expressing desire with a different subject in the matrix clause. Recall that the *dil* construction for expressing desire may be followed by an infinitival CC, in which case the subject of the CC is required to be the same as that of the matrix clause. However, if the *dil* is followed by the subordinating conjunction *iko*, the embedded clause is a finite CC and a different subject may be used.

(245) *wi* *dil* *iko* < *waz* *wi = ri* *az* *xwor* *anur*
 3SG.POSS heart SC < 1SG.NOM 3SG.ACC = DAT from Kashgar pomegranate

vor = am >

bring.NPST = 1SG.NPST >

'His desire is: <I bring him pomegranates from Kashgar >.'

(246) *sofia dil iko <wi ano ajoy zozd>*
Sofia heart SC <3SG.POSS mom shoe buy.3SG.NPST>
 ‘Sofia’s desire is: <her mom buys shoes>.’

(247) *fanbje dil iko <alima indizd, zulfia naθt>*
Shanbie heart SC <Alima stand.3SG.NPST Zulfia sit.3SG.NPST>
 ‘Shanbie’s desire is: <Alima stands, Zulfia sits>.’

4.2 Comparison with related languages

This section will present Persian, Tajik, Shughni, Rushani, and Wakhi examples of CC's that are complements of speech and cognition verbs. They will be compared with Sarikoli CC's.

4.2.1 Persian & Tajik

In Persian and Tajik, the CC in the unmarked order is placed at the end of the matrix clause, and is introduced by the complementizer *ke* (Persian)/*ki* (Tajik) (Windfuhr & Perry 2009:511). Note that the *ke/ki* particle was also used as a relativizer in Persian and Tajik (see section 3.4.1). In the following examples, when there are two lines in the source language, the first line is Persian and the second line is Tajik. Note that the CC follows the verb:

(248) *hasan mi-dan-ad <ke ali in xane-ra saxt>*
Hasan knows <that Ali this house-DO built>
 ‘Hasan knows <that Ali built this house>.’ (Windfuhr & Perry 2009:512)

(249) *fahmid-am <ke adam-e xub-i ast>*
fahmid-am <ki odam-i xub-e ast>
understand.PST-1SG <that person-EZ good-INDEF COP.3SG>
 ‘I realized <that he was a good person>.’ (Windfuhr & Perry 2009:512)

(250) *did <ke hame rafte and>*
did <ki hama rafta and>
see.PST.3SG <that all go.PART COP.3PL>
 ‘He saw <that all had left>.’ (Windfuhr & Perry 2009:512)

4.2.2 Shughni

The Shughni CC directly follows the verb in the matrix clause; in this example, the CC is introduced by a pronoun (Edelman & Dodykhudoeva 2009b:812):

- (251) *wuz na fam-um, <ju tar ka rawun>*
I not know <he to where goes>
'I don't know <where he goes>.' (Edelman & Dodykhudoeva 2009b:812)

4.2.3 Rushani

In Rushani, the CC also follows the verb in the matrix clause, and the complementizer (*dide*) introduces the complement clause (Erschler & Volk 2010:3). As introduced in section 3.4.3, (*dide*) also functions as a relativizer in Rushani. According to Payne, it introduces various object clauses after verbs of speaking and thinking (1989:442).

- (252) *ikim kaða bovar kixt <ide zamin fipak>*
this boy belief do.PRS.3SG <COMP earth flat>
'This boy believes <that the earth is flat>.' (Erschler & Volk 2010:3)

4.2.4 Wakhi

Finite CCs in Wakhi also seem to occur after the verb in the matrix clause. Optionally, the *ki* may be used to introduce the CC. The same particle can also be used as a relativizer, as shown in section 3.4.4.¹

- (253) *maz-ər xan <(ki) (wuz) tsoyadər wez-im>*
'Tell me <when to come>.' (Bashir 2009:854)

- (254) *wuz nə dif-əm <jaw rəx-k tbiw-ətk ja nəj>*
'I don't know <whether he has left or not>.' (Bashir 2009:854)

However, the following example shows that infinitival complements may precede the verb in the matrix clause.

¹ The source of these Wakhi RC examples did not provide morpheme-by-morpheme or word-by-word glosses. Only the free translation was provided.

(255) *wuz-eş <angrezi nivis-en>-iş dij-em*
'I know <how to write English>.' (Bashir 2009:854)

4.2.5 Conclusion

Based on the data presented above, languages that are closely related to Sarikoli tend to place the CC after speech and cognition verbs in the matrix clause, at least in their unmarked order. This results in an SVO constituent order for CC constructions. On the other hand, in Sarikoli, it is extremely rare for CCs to follow the main verb. Both finite and non-finite CC constructions are arranged so that the matrix clause has SOV or OSV constituent order. Only in reporting direct speech, thoughts, dreams, etc. does the CC sometimes follow the verb, but this is not required.

Another common pattern among the CC constructions in these languages is that they use clause-initial complementizers. Many of them are identical to the relativizer used in that language, and some of them are optional. However, Sarikoli CCs are generally unmarked or marked with a suffix that attaches after the verb of the embedded clause. The clause-initial complementizer *iko* is also used, but less commonly.

As with the comparison of RCs across these languages, Sarikoli CCs show a stronger tendency of adhering to the OV type compared to CCs in the other Iranian and Pamir languages discussed in this section.

CHAPTER 5

Adverbial Clauses

Adverbial clauses (ACs) function as modifiers of verb phrases or entire clauses (Thompson & Longacre & Hwang 2007). Thompson et al. (2007) list three devices that are typically used for marking ACs: subordinating morphemes, special verb forms, and word order. Sarikoli uses various subordinating morphemes, such as *tsa* ‘if’, *qati* ‘with’, *alo* ‘when’, *avon* ‘for’, *az* ‘from’, and the dative marker =*ir*, as well as some generic words that function as RC heads. Most of these subordinating morphemes are postpositional, occurring at the end of the AC, although some of them are placed immediately before the verb in the AC.

In terms of verb forms, Sarikoli ACs are also marked with special verb forms. In most types of ACs, the verb is in the infinitive stem and lacks subject-verb agreement clitics. Compare the independent clause (256) and the temporal AC in (257) below (ACs are enclosed in curly brackets):

(256) *itaŋ batfo-χejl gap ɛəwl na-wejð = in*
 some child-NOM.PL word ear NEG-put.in.NPST = 3PL.NPST
 ‘Some children don’t listen (disobey).’

(257) {*batfo-χejl gap ɛəwl na-weðd waχt*} *ato ano tfoqom rond tɛgif*
 {child-NOM.PL word ear NEG-put.in.INF time} dad mom must scold should
 ‘{When children don’t listen}, parents must scold them.’

Finally, Sarikoli ACs may be recognized, to some extent, by their position. They usually precede the entire matrix clause or immediately follow the subject.

This chapter is organized semantically by AC usage, rather than by different grammatical structures.

5.1 Types of Sarikoli ACs

In this section, seven types of Sarikoli ACs will be introduced. They include temporal, reason, purpose, conditional, concessive, means and simultaneous, and substitutive ACs. In addition, non-AC strategies for expressing location and manner will be described.

5.1.1 Time

There are various ways to form a temporal AC. Different constructions are used for indicating different temporal relations between the matrix clause and the embedded event, such as ‘before’, ‘during’, and ‘after’ the embedded event. When pointing directly at the time in the embedded clause, two constructions are used interchangeably. The first is an unmarked infinitival RC with the head noun *waxt* ‘time’, with no adpositions:

(258) {*ju pinz sulo vid waxt*} *wi jaχ azmud seð-dzændz*
 {3SG.NOM five year be.INF time} 3SG.POSS sister born become.PRF-PRF
 ‘{At the time when she was five years old} her sister was born.’

(259) {*zow tsid waxt*} *tçer tfejg=itfuz χal luzim sæwd*
 {wheat cut.grain.INF time} work do.INF=REL person need become.3SG.NPST
 ‘{When it is time to cut wheat}, people who do work are needed.’

(260) {*tçed waqil pa tçed jet waxt*}, *ju a=tçardz tçer*
 {house master to house come.INF time} 3SG.NOM ACC=good work
tçawg=dzændz tçertçi stæwd, a=zit tçer tçawg=dzændz tçertçi
 do.PST=REL worker praise.3SG.NPST ACC=bad work do.PST=REL worker
rond
 scold.3SG.NPST
 ‘{At the time when the master of the house comes home}, he will praise the worker who worked well, and scold the worker who worked badly.’

Alternatively, the temporal particle *alo*, which has less semantic content, may replace the *waxt*. The verb in the embedded event is still in the infinitive stem, and no adpositions are used:

- (261) {*zumun ðod alo} mofin qati kuřwor tid qilo*
 {snow fall.INF TEMP} car with Kashgar go.INF dangerous
səwd
 become.3SG.NPST
 ‘{When it snows}, it gets dangerous to go to Kashgar by car.’
- (262) {*dzaŋ set alo} batfo mas darak nist ano mas darak nist*
 {war become.PST TEMP} child also news NEG.be mother also news NEG.be
 ‘{When there was war}, there was no news about the son, nor about the mother.’
- (263) {*baðo batfo ar uvd sulo dejd alo} χandasur ka=jin*
 {boy child at seven year enter.INF TEMP} circumcise do.NPST=3PL.NPST
 ‘{When a boy becomes 7 years old}, they circumcise him.’ (Arlund & Ibrukhim 2013:213)
- (264) {*waχt fript alo} maf ðəw=an ar kinuχuno dejd*
 {time reach.INF TEMP} 1PL.NOM two=1PL.PST INF move.theater enter.PST
 ‘{When the time came}, the two of us entered the movie theater.’ (Arlund & Ibrukhim 2013:271)

To point to a time before the embedded event, the embedded event is expressed as an infinitival clause, followed by the nominal postposition *prud* ‘before/in front of’ with a simple preposition like *tçi* ‘at’ or *tar* ‘toward’:

- (265) {*ju varçidje tid tçi prud}, ar qotnoj mejmunχuno i jatoq*
 {3s.NOM Tashkurgan go.INF at before} INF traffic hotel one room
zuxt
 buy.NPST
 ‘{Before going to Tashkurgan}, (he) reserved a room at the Traffic Hotel.’
- (266) {*tamoq χig tçi prud}, χu ðust znej*
 {food eat.INF at before} own hand wash.NPST
 ‘{Before eating food}, wash your hands.’
- (267) *waz {xovd tar prud} χu pjets χu pjeð-ef*
 1SG.NOM {sleep.INF toward before} own face own foot-ACC.PL
znej=am
 wash.NPST=1SG.NPST
 ‘{Before sleeping} I wash my face and my feet.’ (Arlund & Ibrukhim 2013:517)

To point to a time after the embedded event, the embedded event is again expressed as an infinitival clause, followed by the nominal postposition *zabu* ‘later/after’ with a simple preposition like *az* ‘from’:

(268) {*χig az zabu*} *a=istxon-jef dzam wix=in*
 {eat.INF from later} ACC=bone-ACC.PL all collect.NPST = 3PL.NPST
 ‘After eating, they collect all the bones.’

(269) {*xwor bejrd jet az zabu*} *wi ar xwor seð*
 {city go.around.INF come.INF from later} 3SG.POSS INF city become.PRF

dzan
war
 ‘{After spending time in cities and coming back}, in his city there was war.’

Sometimes *az zabudz* may be used instead of *az zabu*:

(270) {*χamerbuts weðd az zabudz*} *arvi namodz wejð=an*
 {yeast put.INF from later} there salt put.in.NPST = 1PL.NPST
 ‘{After putting in the yeast}, we put salt in there.’

(271) {*ɣots batfo a=pjets at tfejg az zabudz*} *dzam mejmun-χejl*
 {girl child ACC=face open do.INF from later} all guest-NOM.PL

tjedz=in
go.NPST = 3PL.NPST
 ‘{After opening up the girl’s face}, all the guests leave.’

(272) {*ju batfo sjet az zabudz*} *wi vunudz*
 {3s.NOM child give.birth from later} 3SG.POSS umbilical.cord

χtfejg=an
slice.NPST = 1PL.NPST
 ‘{After she gives birth to the baby}, we cut off its umbilical cord.’

5.1.2 Reason

Reason is often expressed through coordination, with the finite reason clause preceding the finite result clause and the conjunction *kazwi* ‘so’ in between. Optionally, the

Persian loanword *tfunki* ‘because’ may be added at the beginning of the reason clause, but it is not commonly used. The following are examples of this coordinate construction.

(273) *nur awu ðud, kazwi maf puğan xwor tjeðz = an*
 today rain fall.PST so 1PL.NOM tomorrow Kashgar go.NPST = 1PL.NPST
 ‘It rained today, so we will go to Kashgar tomorrow.’

(274) *varçidje utf buland, kazwi mjewo na-past*
 Tashkurgan very high so fruit NEG-ripen
 ‘Tashkurgan is very high, so fruits don’t ripen.’

(275) *tej tfejg = itfuz kots-an wi bob mowy, kazwi tej pa*
 wedding do.INF = REL girl-GEN 3SG.POSS grandpa die.PST so wedding to
sulir rejd
 next.year remain.PST
 ‘The grandfather of the girl who’s getting married died, so the wedding was left until next year.’

Alternatively, reason may also be expressed through subordination. The reason AC consists of a nominalized or infinitival clause with the preposition *az* ‘from’ preceding the verb.

(276) *{nur awu az ðod-i} maf puğan xwor*
 {today rain from fall.INF-NMLZ} 1PL.NOM tomorrow Kashgar
tjeðz = an
 go.NPST = 1PL.NPST
 ‘{Because it rained today}, we will go to Kashgar tomorrow.’

(277) *{varçidje utf buland az vid-i} mjewo na-past*
 {Tashkurgan very high from be.INF-NMLZ} fruit NEG-ripen
 ‘{Because Tashkurgan is very high}, fruits don’t ripen.’

(278) *{tej tfejg = itfuz kots-an wi bob az mejg}, tej pa*
 {wedding do.INF = REL girl-GEN 3SG.POSS grandpa from die.INF} wedding to
sulir rejd
 next.year remain.PST
 ‘{Because the grandfather of the girl who’s getting married died}, the wedding was left until next year.’

(279) *ju* {*χu* *otʃu* *az* *bunost-i*}, *az* *χu* *tɛlan* *ðud*
 {3s.NOM 3SG.POSS key from lose.INF-NMLZ} from own fine give.PST
 ‘{Because he lost his key}, he paid a fine.’

(280) {*mofin-χejl-af* *pa imi* *ljɛχ* *az* *ðod-i*}, *pond* *waru**vd*
 {car-NOM.PL-3PL.PST to RECP collision from give.PST-NMLZ} road stop.PST
 ‘{Because cars collided with each other}, the road stopped (i.e. there was a traffic jam).’

(281) {*nur* *mu* *mom-an* *wi* *jetmif* *sulino* *azmud*
 {today 1SG.POSS grandma-GEN 3SG.POSS seventy year born

sɛð=dzɛndz *maθ* *az* *vid-i*}, *i* *tɛd*
 become.NPST=REL day from be.INF-NMLZ} one house

χal-χejl=af *tup* *sut*
 person-NOM.PL=3PL.PST group become.PST

‘{Because today is my grandmother’s seventieth birthday}, the whole family got together.’

5.1.3 Purpose

There are two ways to form a purpose AC. First, the benefactive particle *avon* may be placed after the infinitival clause:

(282) {*tudzik-χejl* *χu* *ziv* *xojd*, *navift* *avon*}, *χu* *harf=ir*
 {Tajik-NOM.PL own tongue read.INF write.INF BEN} own writing=DAT

mjɛxtudz
 need

‘{In order for Tajiks to read and write their own language}, (they) need their own writing.’

In this construction, the main verb in the matrix clause may be in the non-past tense, with the pronominal agreement clitic attached on the verb, if applicable:

(283) {*a=tɛd waqil xuf tfejg avon*}, *tʃardz tɛertɕi-xejl tɕiŋ*
 {ACC=house master happy make.INF BEN} good worker-NOM.PL complete

az dil xuɣzmat ka=jin
 from heart work do.NPST=3PL.NPST

‘{In order to please the master of the house}, the good workers work with all their heart.’

(284) {*xu qad buɫand tfejg avon*}, *gul ara maθ xjevd brozd*,
 {own height high make.INF BEN} Gul every day milk drink.3SG.NPST

ɣuzd
 run.3SG.NPST

‘{In order to make her height tall}, Gul drinks milk and runs every day.’

The main verb in the matrix clause may also be in the past tense, in which case the pronominal agreement clitic will attach to constituents other than the verb, as in (285) and (286). In (287) and (288), there are no agreement clitics because the subject is third person singular and the verb is in the past tense.

(285) *waz=am* {*i sawul parst avon*}, *xu θust tjer tʃəwŋ*
 1SG.NOM=1SG.PST {one question ask.INF BEN} own hand raise do.PST
 ‘{In order to ask a question}, I raised my hand.’

(286) {*adi batʃo fuv tfejg avon*}, *waz=am wi=ri kampuut*
 {this child quiet make.INF BEN} 1SG.NOM=1SG.PST 3SG.ACC=DAT candy

θud
 give.PST

‘{To make the baby quiet}, I gave him candy.’

(287) {*xu radzjen az xats zuwost avon*}, *amerdin ar darju rawuxt*
 {own daughter from water take.out.INF BEN} Amerdin INF river jump.PST
 ‘{In order to take his daughter out of the water}, Amerdin jumped into the river.’

(288) {*izo xalg-ɛf-an kasal tɕixt=ir jurdam tfejg avon*}, *qobil*
 {village person-ACC.PL-GEN sick watch.INF=DAT help do.INF BEN} Qobil

duxtuuri xumand sut
 doctor study be.PST

‘{In order to care for the villagers’ sicknesses and help them}, Qobil studied to become a doctor.’

This *avon* may also be used when the subject of the purpose AC is different from the subject of the matrix clause:

- (289) {*maf zuundo rejđ avon*}, *ju xu dzun đud*
 {1s.NOM alive stay.INF BEN} 3SG.NOM own life give.PST
 ‘{In order for us to stay alive}, he gave his life.’

Negative purpose ACs are also formed with *avon*:

- (290) {*mu xuđm na-jet avon*}, *waz=am futu xats bruxt*
 {1s.POSS dream NEG-come.INF BEN} 1SG.NOM-1SG.PST cold water drink.PST
 ‘{So that my dreams wouldn't come}, I drank cold water.’

The second way to construct a purpose AC is by attaching the dative marker, =*ir*/=*ri*, to an infinitival clause. As shown in the examples below, the dative marker is not only used for marking indirect objects, but also for indicating purpose.

- (291) {*pa tçed xipik tfejg=ir*} *i soat waχt tizd*
 {at house flatbread make.INF=DAT} one hour time go.3SG.NPST
 ‘It takes one hour {to make flatbread at home}.’

- (292) *ju {dam zoxt-ir} varçidje tujđ*
 3SG.NOM {rest get.INF-DAT} Tashkurgan go.PST
 ‘He went to Tashkurgan {to get rest}.’

- (293) *sovdz-i {a=mujit tfardz tfejg=ir} fujđo*
 green-NMLZ {ACC=environment good do.INF=DAT} beneficial
 ‘Greenness is beneficial {for making the environment good}.’ (Arlund & Ibrukhim 2013:262)

- (294) *maf ar i sul=ir đa gudur xu jar makun {xu*
 1PL.NOM every one year=DAT two time own land hometown {own
xejx-ef wand=ir} so=n
 relative-ACC.PL see.INF=DAT} become.NPST=1PL.NPST
 ‘We go to our hometown twice every year {in order to see our relatives}.’ (Arlund & Ibrukhim 2013:181)

- (295) *firoz maf pa tʃed {paronj weðd=ir} iθtʃ*
 Shiroz 1PL.POSS at house {chat put.in.INF=DAT} come.PRF
 ‘Shiroz came over to our house for a chat.’ (Arlund & Ibrukhim 2013:280)

With this construction, the distinction between purpose and reason ACs is not always clear. The following is an example in which this difference is neutralized. Even though it has the structure of a purpose AC, it makes more sense if it is interpreted as a reason AC.

- (296) *yubun {χu bjeð=dzændz barqo vig=ir} zjofat ðud*
 shepherd {own lose.PST=REL lamb find.INF=DAT} party give.PST
 ‘The shepherd gave a party {for finding his lamb that he had lost}.’

An AC may be embedded in another AC. In the following example, a purpose AC (marked with *-ir*) is embedded within a temporal AC (marked with *alo*).

- (297) *{waz {a=wi tar pond weðd=ir} naxtig alo}*
 1SG.NOM {ACC=3SG.ACC toward road put.in.INF=DAT} come.out.INF TEMP

tanuk=am pamowydz
 thin=1SG.PST wear.PRF
 ‘When I came out to see him out, I was wearing thin clothes.’ (Arlund & Ibrukhim 2013:280)

5.1.4 Condition

A conditional AC is formed by adding the *tsa* particle either before or after the verb or auxiliary in the protasis. The Persian conjunction *agar* ‘if’ may occur at the beginning of the protasis in any of the examples below. *Agar* is not commonly used, but occurs more in hypothetical conditional ACs.

Conditional ACs are flexible in terms of verb finiteness; they can be finite or non-finite. The example below contains a non-finite conditional AC:

- (298) *{dam zoxt tʃardz na-vid tsa} χuzmat=ir tabijaθ tasir*
 {rest take.INF good NEG-have.INF COND} work=DAT natural influence

kaxt
 do.3SG.NPST
 ‘{If we don’t rest well}, it will naturally influence our work.’ (Arlund & Ibrukhim 2013:148)

The conditional AC may also be finite:

- (299) {*ju tçij az dil χumand tsa səwd*}, *wi*
 {3s.NOM complete from heart study COND become.3SG.NPST} 3SG.POSS
natidzja tjer səwd
 score rise become.3SG.NPST
 ‘{If he studies with all his heart}, his scores will get higher.’

- (300) {*ju tilfon tsa zozd*}, *ardi most wi-jan pul*
 {3s.NOM phone COND buy.3SG.NPST} this month 3SG.POSS-GEN money
na-rast
 NEG-remain.3SG.NPST
 ‘{If he buys the phone}, he won’t have any leftover money for this month.’

- (301) {*zumun ðejd tsa} mofin qati kuřwor tid qilo*
 {snow fall.3SG.NPST COND} car with Kashgar go.INF dangerous
səwd
 become.3SG.NPST
 ‘{If it snows}, it gets dangerous to go to Kashgar by car.’

The conditional AC may also occur within imperative matrix clauses.

- (302) {*varçidje joð tsa} mu=ri tilfon ka*
 {Tashkurgan come.NPST COND} 1SG.ACC=DAT phone do.NPST
 ‘{If you come to Tashkurgan}, give me a call.’

- (303) {*i χal ta-jan tardi pjets tsa ðid*}, *adi pjets*
 {one person 2SG.POSS-GEN one.of.two cheek COND hit.3SG.NPST} other cheek
mas wi=ri waðor
 also 3SG.ACC=DAT grab.NPST
 ‘{If someone hits your one cheek}, turn the other cheek as well.’

- (304) {*χal az ta narsa tsa parst*}, *təw tçij az*
 {person from 2SG.ACC thing COND ask.3SG.NPST} 2SG.NOM complete from
zord pur wi=ri ðo
 heart many 3SG.ACC=DAT give.NPST
 ‘{If a person asks you for something}, give to him generously with all your heart.’

Conditional ACs are negated in the same way as a main clause:

(305) *{juu odris tsa na-wazond}, xuu pond xatu did*
 {3s.NOM address COND NEG-know.3SG.NPST} own way lose give.3SG.NPST
 ‘{If he doesn’t know the address}, he will get lost.’

(306) *{juu pa duxturxuno tsa na-tizd}, wi kasal garun*
 {3s.NOM to hospital COND NEG-go.3SG.NPST} 3SG.POSS sickness heavy
səwd
 become.3SG.NPST
 ‘{If he doesn’t go to the hospital}, his sickness will get serious.’

(307) *{batfo-xejl gap bəwl na-wəðd=in tsa} ato ano*
 {child-NOM.PL word ear NEG-put.INF.NPST=3PL.NPST COND} dad mom
tfoqom rond tɛgij
 must scold should
 ‘{If children don’t listen}, parents must scold them.’

A single sentence may contain a conditional AC and a temporal AC:

(308) *{muu xuiðm na-past tsa} jo {bjε-xuiðm waχt alo} waz*
 {1s.POSS dream NEG-ripe COND} or {NEG-dream time TEMP} 1SG.NOM
bjε-ruχ so=m
 NEG-energy become.NPST=1SG.NPST
 ‘{If I don’t sleep well} or {when I can’t fall asleep}, I get worn out.’ (Arlund & Ibrukhim 2013:157)

5.1.5 Concession

The concessive AC also uses the *tsa* particle, preceded by *mas* ‘also’. The word *gartfa* ‘although’ may also occur at the beginning of any of the concessive clauses presented in this section.

The verb within the concessive AC may occur after the *mas tsa*. This is often the case with stative events, which uses the verb *vid* ‘be’:

(309) {*ajdzjamol tɕed-an ɛŋ dzuulik-ow mas tsa vid*}, *ju ɛŋ*
 {Aijiamol house-GEN most small-NMLZ also COND be.INF} 3SG.NOM most

aklinj
 smart

‘{Although Aijiamol is the youngest one in her family}, she is the smartest.’

(310) {*pa di tamoqχuno ar χil tamoq mas tsa vid*}, *wi=ri*
 {at this restaurant every kind food also COND be.INF} 3SG.ACC=DAT

χuɸ veð=dzɛndz i mas na-veðdz
 happy be.NPST=REL one also NEG-be.NPST

‘{Although this restaurant has various kinds of food}, there was nothing that she liked.’

(311) {*ju pir mas tsa seðdz vid*}, *wi-jan i puɸts*
 {3s.NOM old.lady also COND become.NPST be.INF} 3SG.POSS-GEN one son

sut
 become.PST

‘{Although she had become an old lady}, she got a son.’

The verb may also precede the *mas tsa*:

(312) {*wi pjeð ðizd mas tsa*}, *wi dil χu djɛst-ɛf qati*
 {3s.POSS foot hurt.INF also COND} 3SG.POSS heart own friend-ACC.PL with

towp skit tfejg
 ball play do.INF

‘{Even though he hurt his foot}, he wants to play ball with his friends.’

Finally, the verb may also occur between the *mas* and the *tsa*:

(313) {*ju i tsiz χatu tɕɛr na mas tɕəwydz tsa vid*}, *likin*
 {3s.NOM one thing wrong matter NEG also do.NPST COND be.INF} but

a=wɪ=jaf ar tuɾmo zuɾt
 ACC=3SG.ACC=3PL.PST INF jail grab.PST

‘{Although he didn’t do anything wrong}, he was thrown into prison.’

5.1.6 Means and Simultaneity

In many situations, the means of performing another action is not expressed as a subordinate AC, but as a finite clause that is part of a sequence. This coordinated construction states the means clause first, and then an optional conjunction *χου* ‘then’, and then the result clause. The subject is stated in the first clause and gapped in the second.

(314) *dʒjɛð pa pandʒjaro paðɛvdʒ (χου), pa tɕɛd djɛvdʒ*
burglar at window climb.PRF (then) at house enter.PRF
‘The burglar climbed through the window and (then) entered the house.’

(315) *ju i pa risturon tɕɛr tʃəwɔŋ (χου), χου xojd rasquɪt*
3SG.NOM one at restaurant work do.PST (then) own study.INF fee
ðud
give.PST
‘She worked in a restaurant and (then) paid her tuition.’

(316) *ju aroj kursi tʃɛr tɕi tʃɛr latʃəwɔŋ (χου), tom wi ðuɪst*
3SG.NOM three chair raise at raise put.PST (then) then 3SG.POSS hand
fript
reach.PST
‘He stacked up three chairs and (then) his hand reached.’

Another way to express means is by the postposition *qati* ‘with’. This *qati* may take a noun phrase as its object:

(317) *zulfia χου djɛst-ɛf-an wi jurdam qati χου tɕɛr adu tʃəwɔŋ*
Zulfia own friend-ACC.PL-GEN 3SG.POSS help with own work finish do.PST
‘With the help of her friends, Zulfia finished her work.’

Or it may take the infinitival clause to form a subordinate AC, which is usually placed immediately after the subject. While this structure can be used naturally in some sentences, it does not work for all of them, as shown in the ungrammatical example (320).

- (318) *qandik {ara maθ woxt soat-i mafq tfejg qati}, iskropka*
 Qandik {every day eight hour-ADV practice do.INF with} violin
musobiqa =ri bjur-intfi sut
 competition =DAT one-ORD become.PST
 ‘{With practicing eight hours every day}, Qandik got first place in the violin competition.’
- (319) *tursun {ara maθ inglēs ziv kinu, kutub tfixt qati}, xuu*
 Tursun {every day English tongue movie book watch.INF with} own
inglēs ziv tjer sut
 English tongue raise become.PST
 ‘{With watching English movies and books every day}, Tursun improved his English.’
- (320) **ju {i pa risturon tçer tfejg qati}, xuu xojd rasqut*
 3SG.NOM {one at restaurant work do.INF with} own study.INF fee
ðud
 give.PST
 ‘{With working in a restaurant}, she paid her tuition.’

This construction with the infinitival clause followed by the postposition *qati* may also be used to form an AC indicating that an event occurs at the same time as another event. Optionally, the word *tang* ‘simultaneously, at the same time’ may be used to lexically mark simultaneity after the *qati* in each example. The AC is usually placed at the beginning of the matrix clause when marking simultaneity.

- (321) *{šakar pa maktab tid qati (tang)}, awu ðud*
 {Šakar to school go.INF with (simultaneously)} rain fall.PST
 ‘{As Šakar walked to school, (simultaneously)}, it rained.’
- (322) *{waz pa tçed dejt qati (tang)}, muu =ri tilfon*
 {1s.NOM to house enter.INF with (simultaneously)} 1SG.ACC =DAT phone
jot
 come.PST
 ‘{As I entered the house, (simultaneously)}, I got a phone call.’

This AC construction may also contain a further-embedded AC. In the following example, a purpose AC is embedded within the AC that marks simultaneity.

- (323) {waz {sawul parst avon} χu ðust tjer tjeɟ qati (tang)},
 {1s.NOM question ask BEN own hand raise do.INF with (simultaneously)}
 az dars xovd=itfuz waχt sut
 from class sleep.INF=REL time become.PST
 ‘{As I raised my hand to ask a question, (simultaneously)}, it became the time to end class.’

If the same subject is performing two actions simultaneously, the AC with the infinitival clause follows the subject:

- (324) jad batfo {niwd qati} ago sut
 this child {cry.INF with} awake become.PST
 ‘This child woke up {crying}.’

5.1.7 Substitution

In order to express the replacing of an expected event with an unexpected event, Sarikoli uses the substitutive AC, a construction that is similar to the locative AC. The substitutive AC is also formed through relativization, but it uses the unmarked RC, in which the infinitival RC simply precedes the head noun, with the head noun *dɟuj* ‘place’. In addition, the preposition *tçi* ‘at’ takes *dɟuj* as the object of preposition, occurring between the unmarked RC and *dɟuj*. This construction literally means, ‘in the place of doing X’, where X represents the unmarked RC. The substitutive AC may occur immediately after the subject of the matrix clause, or if the direct/indirect object is shared by the AC and the matrix clause, it occurs after the direct/indirect object.

Syntactically, the substitutive AC substitutes a whole clause for another clause. Semantically, it may express substitution of various roles, such as direct objects:

- (325) nur m-ono maf=ir {layman tjeɟ tçi dɟuj}
 today 1SG.POSS-mom 1PL.ACC=DAT {pulled.noodles make.INF at place}
 firgirindɟ tɬəwɟ
 rice.milk.butter make.PST
 ‘{Instead of making pulled noodles}, my mom made rice with milk with butter for us today.’

(326) *woð = af* *{puur zit xalg-ef zjed tci dzuj} i tfardz*
 3PL.NOM=3PL.PST {many bad person-ACC.PL kill.INF at place} one good

xalg-af zjed
 person-3PL.PST kill.PRF

‘{Instead of killing many bad people}, they killed one good person.’

It may also express substitution of indirect objects:

(327) *ju a = xat {gul = ir ðod tci dzuj} wi dzul*
 3ACC.NOM ACC=letter {Gul=DAT gve.INF at place} 3SG.POSS small

jax = ir ðud
 sister=DAT give.PST

‘{Instead of giving it to Gul}, he gave the letter to her little sister.’

Or it may express substitution of entire actions:

(328) *sofia {topçuruuq tfejg tci dzuj} aroj soat kinu tfuxt*
 Sofia {homework do.INF at place} three hour movie watch.PST
 ‘{Instead of doing homework}, Sofia watched movies for three hours.’

(329) *waz {pujz qati tid tci dzuj} xu mofin qati xutan*
 1SG.NOM {train with go.INF at place} own car with Hotan

tjedz = am
 go.NPST = 1SG.NPST

‘{Instead of going by train}, I will go to Hotan by my own car.’

And it may also express substitution of goals:

(330) *alima malum {ar zyin tid tci dzuj} xu xejx ar tej tujd*
 Alima teacher {to meeting go.INF at place} own relative to wedding go.PST
 ‘{Instead of going to the meeting}, Teacher Alima went to her relative’s wedding.’

5.1.8 Location

Sarikoli does not make use of a structurally-distinctive AC construction to express location. Instead, uses an adpositional phrase in which the object of preposition is relativized. The head noun of the RC is often *dzuj* ‘place’, which is a generic word with little

semantic content. The adposition immediately precedes or follows the head noun of the RC. In the following examples, the preposition *tçi* ‘at’ in (331) and the complex adpositional expression *pa prud* ‘in front of’ in (332) indicate the spatial relationship between the relativized ‘place’ and the location in the matrix clause:

- (331) *{i kots kamput pataw=dzændz tçi dzuj}, mjerz-χejl=af wixtç*
 {one girl candy drop.PST=REL at place} ant-NOM.PL=3PL.PST gather.PRF
seðdz
 become.NPST
 ‘{At the place where a girl dropped candy}, ants gathered.’

- (332) *{muu tçer tfejg=itfuz dzuj pa prud}, dzul batfo-χejl towp skit*
 {1s.POSS work do.INF=REL place at front} small child-NOM.PL ball play
ka=jin
 do.NPST=3PL.NPST
 ‘{In front of my working place}, little children play ball.’

The adposition, however, is optional; it may be omitted if the head noun modified by the RC directly points to the intended location. If the noun phrase alone can point to somewhere ‘in’ or ‘at’ the place, the adposition is unnecessary:

- (333) *ju {parus xaftuli daraxt weðd=dzændz dzuj}, sjeð xaftuli*
 3s.NOM {last.year peach tree put.INF.PST=REL place} this.year peach
waðordz
 grab.NPST
 ‘{At the place where he planted a peach tree last year}, he took a peach this year.’

- (334) *{mjændz-i kalo pwojd=itfuz dzuj}, χurgu jost*
 {summer-TEMP sheep herd.INF=REL place} yurt have.NPST
 ‘{At the place where sheep are herded in the summer}, there are yurts.’

Note that both finite and non-finite RCs may be used to modify the head noun *dzuj*.

In order to be more specific about the place, the head noun of the RC can be something other than *dzuj* ‘place’:

(335) {*aroj afto tci prud at tfəwg=dzəndz pa risturon*}, *nur ða tej*
 {three week at before open do.PST=REL at restaurant} today two wedding

jost

have.NPST

‘{At the restaurant that opened three weeks ago}, there are two weddings today.’

(336) *ju {waz parus tujd=dzəndz ar maktab} tujd*
 3SG.NOM {1s.NOM last.year go.PST=REL to school} go.PST
 ‘He went to the school {that I went to last year}.’

(337) *m-oto {χu weðd=dzəndz tçed tci tom} a=kalo kaxt*
 1SG.POSS-dad {own put.PST=REL house at roof} ACC=sheep kill.PST
 ‘My father killed a sheep {on top of the house that he built}.’

5.1.9 Manner

Sarikoli also does not have a structurally-distinctive AC construction for expressing manner, and instead relies on a RC construction with the generic word *rəŋg* ‘manner’ as the head noun. This strategy for expressing manner is always marked with the finite relativizer *-dzəndz*, regardless of whether the embedded event has already happened or is a timeless event:

(338) {*mu boŋjeq levd=dzəndz rəŋg*} = *am adi tçer tfəwg*
 {1s.POSS boss say.PRF=REL manner} = 1SG.PST this work do.PST
 ‘I did the work {as my boss instructed}.’

(339) *waz-am adi bejt {tci radio latfəwy=dzəndz rəŋg} = am*
 1SG.NOM-1SG.PST this song {at radio put.PRF=REL manner} = 1SG.PST

levd

say.PST

‘I sang this song {as it was played on the radio}.’

(340) *waz = am {χ-ono levd=dzəndz rəŋg}-am adi*
 1SG.NOM = 1SG.PST {OWN-mom say.PRF=REL manner}-1SG.PST this

çədoji unsuvd

Sarikoli.cap sew.PST

‘I sewed the Sarikoli traditional cap {as my mom told me}.’

Note that, in examples (339) and (340), the subject-verb agreement clitic occurs twice. This is optional; the meaning of the sentence remains the same whether the clitic is used once or twice.

Simile clauses are constructed in the same way. If the point of similarity is the action performed by two different explicit subjects, the RC expressing manner (with the image) usually occurs at the beginning of the clause, and is set off from the main clause (with the topic) with an intonation break:

(341) *{purg a=girindz tfardz wand=dzændz rang}, waz a=ta*
 {mouse ACC=rice good see.PRF=REL manner} 1SG.NOM ACC=2SG.ACC

tfardz wejn=am
 good see.NPST=1SG.NPST
 ‘{As a mouse likes rice}, I like you.’

(342) *{χtsuvd a=xutum zatran weðd=dzændz rang}, soqtçi a=dzjēð*
 {eagle ACC=rabbit chase put.INF-PRF=REL manner} police ACC=burglar

zatran weðd
 chase put.PST
 ‘{As an eagle chases a rabbit}, the police chased the burglar.’

If the sentence only mentions one subject and the action portrayed by the main verb is compared to another action, the subject is usually stated first, followed by the manner RC and the other core arguments, and then followed by the verb. The manner RC can either precede or follow the direct object:

(343) *i tup purg {dzarj tfəwy=dzændz rang} a=tamoq-ik az imi*
 one group mouse {war do.PRF=REL manner} ACC=food-??? from RECP

talof ka=jin
 fight do.NPST=3PL.NPST
 ‘{As fighting in a war}, a group of mice is fighting for the food.’

(344) *ju a=hansutfa harf {rasim tizd=dzændz rang} navift*
 3SG.NOM ACC=Han script {picture draw.PST=REL manner} write.PST
 ‘She wrote the Han script {as if drawing a picture}.’

(345) *ju* *{i* *χalg* *məwγ=dzɛndz* *rang}* *niwd*
3SG.NOM {one person die.PST=REL manner} cry.PST
'He cried {as if someone had died}.'

CHAPTER 6

Conclusion

6.1 Summary

In this thesis, I have described RCs, CCs, and ACs in Sarikoli. These RC and CC structures were briefly compared with those found in Persian, Tajik, Shughni, Rushani, and Wakhi.

Sarikoli RCs are placed before the head noun. They may either precede or follow determiners and prepositions that are linked to the same head noun, but always precede adjectives that modify the same head noun. Unlike independent clauses, they do not show subject-verb agreement through pronominal clitics. They are negated in the same way as independent clauses, with the *na* particle that precedes the verb. Common nouns, proper nouns, demonstratives, and generic terms may be relativized, but pronouns are generally not relativized. A wide range of syntactic functions are allowed for the common argument in both the RC and the matrix clause, including A, S, O, and oblique roles. The two main relativizers used for Sarikoli RCs are =*dʒɛndʒ* and =*itfuz*. The =*dʒɛndʒ* relativizer is used for finite RCs, while =*itfuz* is used for non-finite clauses (including future events with an infinitive verb). Other ways of forming RCs include headless RCs, unmarked RCs, and using the genitive marker *-an*.

In contrast with Sarikoli, some of the other languages in the Iranian and Pamir language family, such as Persian, Tajik, Shughni, Rushani, and Wakhi, tend to place the RC after the head noun and use clause-initial relativizers.

Sarikoli has at least two types of finite CCs and two types of non-finite CCs. The pre-verbal finite CC is unmarked and contains a finite verb stem and a subject-verb agreement clitic. In the nominalized complement, the nominalizer *-i* attaches to the infinitive stem of

the verb. The infinitival complement also contains the infinitive stem of the verb, but is unmarked. The post-verbal finite CC is the only type of CC that is placed after the matrix clause verb; it is introduced by the subordinating conjunction *iko* and contains a finite stem of the verb and a subject-verb agreement clitic. The reporting of direct speech uses either of the finite CCs, and the reporting of indirect speech uses the nominalized complement. The hearsay construction uses the infinitival complement with a dative marker.

While Sarikoli generally tends to place CCs before the verb, some of the other Iranian and Pamir languages tend to place CCs after speech and cognition verbs in their default order. Sarikoli has both clause-initial and clause-final complementizers, as well as unmarked CCs; on the other hand, these other languages use clause-initial complementizers, such as *ke*, *ki*, and *ide*.

ACs are marked by various subordinating morphemes, including *tsa* 'if', *qati* 'with', *alo* 'when', *avon* 'for', *az* 'from', and the dative marker =*ir*, which generally occur at the end of the AC. Most types of ACs are non-finite, containing verbs in their infinitive stem and lacking subject-verb agreement clitics. Like regular adverbs, ACs usually precede the entire matrix clause or immediately follow the subject. Sarikoli has structurally-distinctive AC constructions for expressing time, reason, purpose, condition, concession, means and simultaneity, and substitution. Location is expressed through adpositional phrases with relativized objects, and manner is expressed through RCs.

6.2 Areas for further research

This thesis has focused on synchronically describing the structure of subordinate clauses in Sarikoli. But this analysis has raised some questions that could not be addressed without the help of historical-comparative linguistics, as there were some notable differences between Sarikoli and other closely-related languages. Why does Sarikoli have prenominal RCs and pre-verbal CCs, in contrast with the other Iranian and Pamir languages that were investigated? Have these orderings resulted from contact with the surrounding Turkic languages, which are OV languages that place the RC before the head noun? In-depth historical-comparative analysis is needed in order to provide an explanation for these differences.

On the other hand, there was one Sarikoli construction that was similar to those found in some of the related languages. Sarikoli's post-verbal finite CC introduced by the particle *iko* is similar to the Persian, Tajik, and Wakhi CCs that are introduced with *ke/ki*. It would be interesting to find out, through historical-comparative analysis, whether these two constructions are related.

Another question about RCs that remains unresolved has to do with why the finite relativizer =*dʒendʒ* is sometimes used for events that have not happened in the past. There were two instances in which the =*dʒendʒ* relativizer was used, with the past or perfect stem of the verb, even though the embedded event was not in the past: non-restrictive RCs with the verb *veðdʒ* 'be (PR)', and non-past manner ACs. It would be interesting to find out why they do not use the infinitival relativizer =*itfuz*.

Finally, for the CCs, the use of the nominalizer suffix *-i* needs further investigation. It is not completely clear when it occurs (creating a nominalized complement) and when it does not occur (creating an infinitival complement). Some of the possible triggering factors for the *-i* suffix have been discussed, including the tense of the embedded verb and the semantic type of matrix clause verb. However, it seems that a more consistent rule for the use of *-i* is yet to be discovered.

References

- Andrews, Avery D. 2007. Relative clauses. In Timothy Shopen (eds.), *Language typology and syntactic description: Complex constructions.*, 2nd edn. 2,206-36. New York: Cambridge University Press.
- Arlund, Pam S. 2006. *An acoustic, historical, and developmental analysis of Sarikol Tajik diphthongs*. University of Texas at Arlington dissertation.
- Arlund, Pam S. and Neikramon Ibrukhim. 2013. *A Chinese Tajik reader: An introduction to Sarikoy (Sarikol) Tajik*. Grandview, MO: All Nations Publishing.
- Barie, Amanda E. 2009. *Exploring cleft sentences and other aspects of Shughni syntax*. University of Kentucky thesis.
- Bashir, Elena. 2009. Wakhi. In Gernot L. Windfuhr (eds.), *The Iranian languages, collection_volume*, 825-62. London: Routledge.
- The National Bureau of Statistics of China. 2010. *Sixth National Population Census of the People's Republic of China*. Beijing.
- Comrie, Bernard. 1983. *Language universals and linguistic typology: Syntax and morphology*. 2nd edn. Oxford: Blackwell.
- Dixon, R. M. W. 1972. *The Dyirbal language of north Queensland*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. 2006. Complement clauses and complementation strategies in typological perspective. In R. M. W. Dixon and Alexandra Y. Aikhenvald (eds.), *Complementation, a cross-linguistic typology, collection_volume*, 1-48. Oxford: Oxford University Press.
- Dixon, R. M. W. 2010. *Basic Linguistic Theory: Grammatical topics*. 2. Oxford: Oxford University Press.
- Dodykhudoeva, Leila R. 2004. Ethno-cultural heritage of the peoples of West Pamir. *Collegium Antropologicum* 28. 147-59.
- Dodykhudoeva, Leila R. 2007. Revitalization of minority languages: Comparative dictionary of key cultural terms in the languages and dialects of the

- Shugni-Rushani group. In Peter K. Austin, Oliver Bond, and David Nathan (eds), *Proceedings of conference on language documentation and linguistic theory*. proceedings_volume:69-79. London: SOAS.
- Dryer, Matthew S. 2006. Descriptive theories, explanatory theories, and Basic Linguistic Theory. In Felix K. Ameka, Alan Dench, and Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, collection_volume, 207-34. Berlin: Mouton de Gruyter.
- Dryer, Matthew S. 2013. Order of Relative Clause and Noun. In Dryer, Matthew S. & Haspelmath, Martin (eds.), *The World Atlas of Language Structures Online*, Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/90>.
- Edelman, Džoy I. and Leila R. Dodykhudoeva. 2009a. The Pamir languages. In Gernot L. Windfuhr (eds.), *The Iranian languages*, collection_volume, 773-86. London: Routledge.
- Edelman, Džoy I. and Leila R. Dodykhudoeva. 2009b. Shughni. In Gernot L. Windfuhr (eds.), *The Iranian languages*, collection_volume, 787-824. London: Routledge.
- Erschler, David and Vitaly Volk. 2010. On clause-internal complementizers in Ossetic and Pamiri. *Syntax of the World's Languages IV*. Lyon, France.
- Gao, Erqiang. 1985. Tajikeyu jianzhi. In collection_editor (eds.), *Zhongguo Yuwen*, 123, 161-75. Beijing: Nationalities Publishing House.
- Ibrukhim, Neikramon. 2012. *Chinese Tajik alphabet*. Urumqi, Xinjiang.
- Kim, Deborah. 2013. Sarikoli texts. University of North Dakota, ms.
- Kreutzmann, Hermann. 2003. Ethnic minorities and marginality in the Pamirian knot: Survival of Wakhi and Kirghiz in a harsh environment and global contexts. *The Geographic Journal* 169(3). 215-35.
- Lewis, M. Paul. 2013. title. In Lewis, M. Paul (eds.), *Ethnologue: Languages of the world*, 17th edn. collection_volume, collection_pages. Dallas, TX: SIL International. <http://www.ethnologue.com>.
- Noonan, Michael. 2007. Complementation. In Timothy Shopen (eds.), *Language typology and syntactic description: Complex constructions*, 2nd edn. 2,52-150. New York: Cambridge University Press.
- Pakhalina, T. N. 1960. *Sarykolskii dialekt i ego otnoshenie k drugim dialektum shughnano-rushanskoi iazykovoi*. Moscow: Izd-vo vostochnoi lit-ry.

- Pakhalina, T. N. 1966. *Sarikolsko - Russkii slovar*. Moscow: Institute of Linguistics, Soviet Academy of Sciences.
- Payne, John. 1989. Pamir languages. In Rüdiger Schmitt (eds.), *Compendium linguarum Iranicarum*, collection_volume, 417-44. Wiesbaden: Ludwig Reichert.
- Payne, Thomas E. 1997. *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press.
- Stump, Gregory. 2012. Updated discussion of Tajiki homework #5. University of Kentucky, ms.
<http://blog.as.uky.edu/stump/wp-content/uploads/2012/04/2012-516-Tajiki-homework-5-discussion-updated.pdf>.
- Taghvaipour, Mehran A. 2005. *Persian relative clauses in Head-driven Phrase Structure Grammar*. University of Essex dissertation.
- Tegey, Habibullah and Barbara Robson. 1996. *A reference grammar of Pashto*. Washington, D.C.: Center for Applied Linguistics.
- Thomason, Sarah G. and Kaufman, Terrence. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.
- Thompson, Sandra A., Robert E. Longacre and Shin Ja Hwang. 2007. Adverbial clauses. In Timothy Shopen (eds.), *Language typology and syntactic description: Complex constructions.*, 2nd edn. 2,237-300. New York: Cambridge University Press.
- Velupillai, Viveka. 2012. *An introduction to linguistic typology*. Amsterdam: John Benjamins Publishing.
- Weber, David J. 1983. *Relativization and nominalized clauses in Huallaga (Huanuco) Quechua*. 103. (University of California Publications in Linguistics.) Berkeley and Los Angeles, CA: University of California Press.
- Windfuhr, Gernot L. and John R. Perry. 2009. Persian and Tajik. In Gernot L. Windfuhr (eds.), *The Iranian languages*, collection_volume, 416-544. London: Routledge.