

Language and Transformation in Samarqand

Nicholas Kontovas

Tajik is a variety of Persian, a language which has been widely spoken in Central Asia in some form or another since at least the 7th century AD. It is an Indo-European language, distantly related to English. Uzbek, by contrast, is a Turkic language, thus genealogically unrelated to Tajik, with a vastly different grammar and core vocabulary. While Turkic has also been spoken in the vicinity of Samarqand since around the middle of the first millennium, the spaces it has occupied have long been separate – both socially and, often, physically – from those in which Tajik has been dominant. While Turkic made some inroads as a language of literature and administration during the reign of the Timurids (mid-14th to early 16th centuries), Tajik remains a marker of identity and a prestige language for the inhabitants of the Old City.

Yet, faced with centuries of increasing pressure from Turkic through interaction and intermarriage with other parts of the city and beyond, Tajik speakers have unconsciously introduced more and more elements from Uzbek into their Tajik. While words are easy to weed out as insufficiently Tajik, aspects of grammar seem to have been less hard to filter out and, over time, Samarqandi Tajik has evolved into a sort of hybrid: a Turkic soul in a Persian body, to paraphrase Dawkins' seminal study on contact-induced language change.¹

Note on linguistic examples

This article contains examples from several languages: Samarqandi (Tajik), Standard Tajik, Early New Persian, Uzbek, and Qarakhanid (Turkic). For ease of reading, all examples are presented in Latin script. For Uzbek, Samarqandi, and Standard Tajik, I have used a regularised version of the mode of transcription used colloquially in Samarqand itself. For Early New Persian, I have used the recommended transcription system of the International Journal of Middle Eastern Studies; for Qarakhanid, for which there is no standard scholarly Latinisation, I have modified the system recommended for Ottoman Turkish.

This article also makes use of interlinear glossing, a convention for making linguistic examples more comprehensible to non-speakers of the relevant language(s). I have tried to keep linguistic jargon to a minimum and explain it where present, however the following abbreviations could not be avoided:

1S	1st person singular
2P	2nd person plural
3S	3rd person singular
ABL	Ablative
ACC	Accusative
EZ	Ezafé
INF	Infinitive
IPFV	Imperfective
PRES	Present
PTCP	Participle
REL	Relative particle



Fig. 1: The Sherdor Madrasa – a Timurid centre of learning at the heart of Samarqand's Old City. (Photo by the author)

Context

In the only large-scale systematic study of the impact of language contact on the grammar of Samarqandi Tajik, Soper identified the extensive structural isomorphism between Uzbek and Samarqandi Tajik as what linguists now call “metatypy”: the wholesale alignment of two grammatical systems through contact.² Yet, while Soper documented the phenomenon of grammatical isomorphism described above at the level of the individual noun and verb phrase, his untimely death meant that he was never able to delve deeper into the impact of contact on other aspects of grammar.

When I set out to conduct three months of field research on Samarqandi Tajik, my plan was to examine the structure of a lesser examined point of contact: relative clauses. This area of grammar is ripe for investigation from the perspective of language contact, not only because the standard structure of relative clauses in Turkic and Persian are so different, but because there are a variety of intermediate forms between the standard structures that may represent different stages of development in the transition from a more conservative Tajik system to a more Uzbek-influenced one.

After my first interviews, it became apparent that retention of conservative grammatical elements – either alongside or to the exclusion of an possible Uzbek-like alternative – existed in numerous common areas of Samarqandi grammar. Moreover, outside my initial object of study, I found other conservative or otherwise non-Uzbek-like structures which told an interesting story of the timing and nature of the language contact which otherwise makes Samarqandi Tajik so different to most other Tajik varieties. Given the extent to which Samarqandi has been reshaped on the model of Uzbek, it is surprising that there should be any elements which have escaped this otherwise nearly wholesale morphosyntactic restructuring. What follows is an attempt to explain why some of the most obvious conservative elements examined during fieldwork have remained.

The city of Samarqand in modern day Uzbekistan has long been one of Central Asia's most attractive and diverse cities. Nestled in the heart of Eurasia at the nexus of the Silk Routes, Samarqand has drawn conquerors, traders, artisans, scholars, and missionaries from settled and nomadic populations far and wide for the better part of two millennia. It is perhaps no surprise, then, that Samarqand today is something of a linguistic oddity. Although surrounded by numerous Uzbek-speaking towns and villages, the historic centre of this second most populous city in Uzbekistan is predominantly – and staunchly – Tajik speaking.

In my field research, use of the more conservative version (a) was restricted to those interlocutors who reported having received education in or significant exposure to standard Tajik, but even these interlocutors only used this version some of the time. Preliminary analysis suggests that the choice of a structure closer to the standard was conditioned by the formality of the interview environment, in which I as a foreigner speaking Tajik was considered less likely to understand local forms. As conversation migrated into the area of personal anecdotes, or as other native speakers of Samarqandi joined the conversation, the likelihood of these speakers to switch to other versions seems higher.

The reasons behind variation between versions (b) and (c) seem less clear. While most interlocutors without heavy exposure to standard Tajik preferred version (c), version (b) seems to appear more frequently where the length of the relative clause is shorter.

Interestingly, a similar distribution can be observed with the position of underived adjectives, which can appear both postposed with an ezafé and preposed. In the case of underived adjectives, however, frequency of the adjective seems to play a larger role in deciding one's preference in terms of position than length of the adjective phrase. High frequency, basic adjectives such as *kalon* “big” can appear in both types of constructions, but most frequently appear with the ezafé:

Variation in adjectival phrase and relative clauses

My initial investigation into Samarqandi relative clauses revealed a staggering variety already, conditioned not only by various semantic and syntactic environments, but also by social factors.

Briefly, interlocutors employed three different structures for the building of relative clauses:

- (a) Postposed with relative particle and finite clause
- (b) Postposed with ezafé and participle
- (c) Preposed with participle

Each of these strategies can be demonstrated with an example from the data. Strategy (a) employs a relative particle *ki*, as in standard Tajik:

- (1) Samarqandi
kasi ki Samarqandi nest
 someone REL Samarqandi is.not
 “someone who is not from Samarqand”

Strategy (b) employs a linking suffix, known in linguistic literature by its Iranian Persian name, the ezafé, along with a postposed participle or verbal adjective. This is acceptable in standard Tajik and other varieties of literary Persian (albeit without the possessive ending shown below) but it is not the preferred way of forming true relative clauses:

- (2) Samarqandi
hamu savol-i mepursidagem
 (me-pursid-agi-yam)
 that question-EZ IPFV-ask-PTCP-my
 “that question which I will ask”

Finally, strategy (c) in which the participle simply precedes the noun is the most Uzbek like:

- (3) Samarqandi
har röz kor me-kad-agi odam
 every day work IPFV-do-PTCP person
 “the person who works every day”

- (4) Samarqandi
xone (xona-yi) kalon
 house-EZ large
 “large house”

Certain fixed phrases seem also to have been lexicalised with an ezafé and postposed adjective:

- (5) Samarqandi
choy-i fomil
 tea-EZ green (of tea,
 formerly a brand)
 “green tea”

In contrast, adjectives which are more distant from the “core” vocabulary of the language usually appear before the noun they modify:

- (6) Samarqandi
mazador ovqot
 tasty food
 “tasty food”

Samarqandi also has a number of adjectives derived from Russian – a language in which adjectives generally precede the nouns they modify. It is perhaps unsurprising, then, that while some interlocutors reported that it was possible to employ Russian-derived adjectives in ezafé constructions, they displayed a strong preference for pre-nominal position when using them:

- (7) Samarqandi
angliski zuvon
 English language
 “the English language”

The imperfective prefix *me-*

Another conservative element of Samarqandi grammar is its use of a prefix *me-* to indicate imperfective aspect on the verb. In other words, it is used to indicate the quality of a predicate which remains conceptually incomplete by the time being referred to in the context of the utterance in which it appears.

Though verb forms with this prefix correspond semantically to verbs in Uzbek with the imperfective ending *-a*, it is easy to miss that, while many preposed grammatical markers in Samarqandi Tajik have moved to the end of the verb to match their position in Uzbek, this one remains in the same position which it occupies in Standard Tajik:

(8) Samarqandi
me-don-am
IPFV-know-1S
“I know”

(9) Uzbek
bil-a-man
know-IPFV-1S
“I know”

The position of *me-* in Samarqandi Tajik is perhaps even more surprising when one considers that, much like pre-posed adjectival elements, some varieties of Persian as recent as the 12th century have allowed the *hamē* element, from which *me-* is derived, to appear after the rest of the verbal complex:

(10) Early New Persian
bōy-l jūy-i Mūliyān
smell-EZ stream-EZ Muliyan
āy-ad hamē
come-3S IMPV
“The smell of the stream of Muliyan comes”
- Rūdakī, *Qaṣīda* 121 (10th c. AD)

One reason that *me-* has remained at the front of the verb root may be that it seems to have fused to THE root to form a conceptual stem. Attesting to this is its phonetic reduction before verbs beginning with a vowel as well as a type of regressive assimilation we see with the vowel in certain monosyllabic stems:

(11) Samarqandi
my-ovr-am
IMPV-bring-1S
“I (will) bring”

(12) Samarqandi
mōrōt (< me-rav-at)
IMPV-go-3S
“she/he goes/will go”

This suggests at the very least that the type of contact with Uzbek that produced the dramatic changes in position that we see in Samarqandi Tajik affixes must have occurred long after *me-* lost the ability to appear post-verbally and attained its fixed in its current, highly integrated pre-verbal position.



Fig. 2: A statue of 'Ali Šēr Navā'i, the national poet of Uzbekistan who spent time in Samarqand and was a prolific writer in both Turkic and Persian. (Photo by the author)

Prepositions that are still prepositions

While most varieties of Persian employ prepositions exclusively, in Samarqandi, most prepositions have become postpositions on the model of Uzbek case endings:

(13) Standard Tajik
ba Samarqand kay omd-ed
to Samarqand when came-2P
“When did you come to Samarqand?”

(14) Samarqandi
Samarqand ba kay omd-et
Samarqand to when came-2P
“When did you come to Samarqand?”

(15) Uzbek
Samarqand-ga qachon kel-di-ngiz
Samarqand-DAT when come-PST-2P
“When did you come to Samarqand?”

Yet, much like the verbal prefix *me-*, several prepositions in Samarqandi Tajik seem to have escaped the migration of most prepositions to post-nominal position. The three most noticeable gathered during fieldwork are *broy* “for”, *a* “from”, and *qati* “with”:

(16) Samarqandi
broy shumo
for you
“for you”

(17) Samarqandi
a injo dur na-rav
from here far not-go
“don't go far from here”

(18) Samarqandi
qati way
with him/her
“with him”

Curiously the last of these alternates with a much more common postposition *kati* of the same meaning:

(19) Samarqandi
man kati
me with
“with me”

Unlike with the verbal prefix *me-*, however, there is no indication of a fusion between these prepositions and the nouns they precede. Why, then, should these particular prepositions not have become postpositions?

In the case of *broy* and *qati*, this may have to do with a fundamental difference between the meaning of these prepositions and those which became postpositions. Namely, *broy* and *qati* do not indicate a fundamental physical location or motion like most prepositions, but more abstract concepts. Cross-linguistically, these types of meaning are more likely to be encoded in case endings vs. pre- or post-positions, as they are in many Turkic languages, such as Uzbek. Perhaps, movement to the post-nominal position is fundamentally easier for those prepositions which are more “case-like” in meaning, a theory which makes sense at first glance, as they occupy the same slot in the noun phrase as a case ending in Samarqandi.

There are, however, two problems with this theory. Firstly, why, then, would there be an alternative form *kati* “with” which does behave like a postposition? This may simply be because it originated in a variety of Tajik where contact was intense enough to overcome the semantic pressure to remain a preposition. Mixing between dialects on a local level then led to a system in which both forms coexisted, which may incidentally explain why the two forms vary slightly in pronunciation.

Secondly, no such semantic explanation is possible for a “from,” which does express a more fundamental spatial meaning. It may be that a remained at the front of the nominal complex because to move it to the end would have caused confusion with another postposition with a different meaning: the accusative/genitive suffix *-(y)a* derived from earlier *-rā*. While the latter has the form *-ya* after vowels, it is identical to a when used after a consonant, except perhaps



Fig. 3: The Sijob Bazaar in Samarqand, where much of the fieldwork for this research was conducted. (photo courtesy of Yoshi Canopus on Wikimedia, reproduced under a CC license. Accessed 10 March 2023)

for prosodic factors which have yet to be investigated:

(20) Samarqandi
a bozor oma -ysa -s
from market come PRES 3S
“He's coming (back) from the market.”

(21) Samarqandi
Samarqand-a nag'z me-bin-et-mi
Samarqand-ACC good IMPV-see-2P
“Do you like Samarqand?”

One piece of evidence to support this theory is that there is one context attested in my fieldwork in which *a* can indeed come after a phrase:

(22) Samarqandi
omdagem (< omd-agi-yam) boz a
came-PTCP-my again from
“since I came”

A follow-up question confirmed that it was also possible to place the *a* before the verbal complex or indeed to leave it out entirely (e.g. *omdagem boz*) without any difference in meaning, confirming that it was indeed equivalent to the preposition *a* and not some other morpheme:

(23) Samarqandi
a omdagem (< omd-agi-yam) boz
from came-PTCP-my again
“since I came”

(24) Samarqandi
omdagem (< omd-agi-yam) boz
came-PTCP-my again
“since I came”

What is revealing about this construction is that the *a* appears not after the nominal element *omdagem* “my having come” in the verbal complex, but after an adverbial element *boz* “since, again.” This is precisely a place after which the accusative/genitive case cannot exist, meaning that the *a* in this construction is wholly unambiguous. If it is indeed due to possible confusion with the accusative/genitive that a normally cannot act as a postposition, it would make sense that in unambiguous cases it could.

Participles in place of infinitives

A final construction in Samarqandi which is unlike Uzbek is one which is unique, inasmuch as it is also unlike more conservative forms of Tajik. In Uzbek, to say that one must do something, one normally employs an infinitive verbal noun:

(25) Uzbek
ket-ish-im kerak
leave-INF-my necessary
“I need to leave”

Indeed, this is also possible in Samarqandi and other varieties of Tajik:

(26) Samarqandi
raft-an-am darkor
go-INF-my necessary
“I need to leave”

There is, however, another construction of similar meaning also attested in our Samarqandi data. Here, however, the verbal noun is not an infinitive, but a perfective participle (i.e. “having gone”):

(27) Samarqandi
raftagem (< raft-agi-yam) darkor
go-PTCP-my necessary
“I need to leave”

Using the equivalent morpheme in the same place is Uzbek is wholly ungrammatical:

(28) Uzbek
**ket-gan-im kerak*
leave-PTCP-my necessary
“I need to leave”

That being said, in older varieties of Turkic attested in the region, one can indicate necessity using an older perfective participle *-miş/miş*:

(29) Qarakhanid
qamuğ tārsiz iş-tin
all iniquitous deed-ABL
yıra-mış keräk
stay.far-PTCP necessary
“one must stay far from all iniquitous deeds”
- Yusūf Xaṣṣ Ḥācib, *Qutadğū Bilig*
(11th c. AD)

Though further research is required to support this theory, it is possible that this particular construction in Samarqandi Tajik is based not on analogy with modern Uzbek but on some older Turkic variety. Alternatively, or in addition to contact pressure from older varieties of Turkic, there may be some reason related to universal tendencies in grammar to form such constructions with perfective participles. The logic behind this, is that when one expresses necessity one normally expresses that a predicate should necessarily be completed. This is supported by data from other languages, but in the absence of a larger such pattern in Samarqandi Tajik it is impossible to say what rule cross-linguistic tendencies may have played in the formation of the *-agi darkor* structure.

Lessons from Samarqand

Samarqand has for centuries been a center of learning for those who would come to study, whether in today's universities, the madrasa's of the Timurids, or the Manichaean monasteries before them. My own lessons from Samarqand are more modest, yet no less important for the study of language contact: that often the things which have not changed as the result of ongoing contact are as interesting as those which have.

Nicholas Kontovas

works as Nizami Ganjavi Subject Librarian for Library for the Caucasus, Central Asia, and Ottoman Turkish at the University of Oxford's Bodleian Libraries, where he also acts as curator for Ottoman Turkish manuscripts. He is currently completing his doctoral dissertation at Leiden University on changes in the verbal system of Medieval Central Asian Turkic as a function of language contact and linguistic prestige. Email: nicholas.kontovas@bodleian.ox.ac.uk

Notes

- Dawkins, R.M. 1916. *Modern Greek in Asia Minor. A study of dialect of Silli, Cappadocia and Phárasa with grammar, texts, translations and glossary.* Cambridge: Cambridge University Press.
- Soper, J.D. 1987. *Loan Syntax in Turkic and Iranian: The Verb Systems of Tajik, Uzbek and Qashqay.* Los Angeles: University of California.