Evidentiality in Uzbek and Kazakh

Christopher A. Straughn

10/25/2011

Committee:
Victor Friedman, chair
Kağan Arık
Lenore Grenoble
Salikoko Mufwene

Overview

The purpose of this work is to describe and account for the broad range of phenomena that have been referred to as “evidentiality” in two Turkic languages of Central Asia: Uzbek and Kazakh. In doing so, I hope to answer the following questions:

- How is evidential meaning expressed in Uzbek and Kazakh?
- What other meanings are expressed using these strategies and can we find a way to account for all of these meanings?
- How do Uzbek and Kazakh fit into a broader Eurasian evidentiality belt?

The answers to these questions are:

- Evidential meaning is expressed primarily via copular forms of historical perfects. Other meanings that have been categorized as evidential are expressed by various past tense markers.
- These markers also express admirativity (i.e. unexpected information); some past tense forms also express doubt and non-volitionality. These meanings can be accounted for under an umbrella category of non-confirmativity (Aronson 1967, Friedman 1978, 1980).
- Uzbek and Kazakh resemble other languages of the Eurasian evidentiality belt (e.g. Turkish, Macedonian, Georgian, Tajik) in that the origin of evidential forms is the past tense, and
that non-confirmativity, rather than pure evidentiality, is the primary meaning of these forms.

The structure of today’s presentation:
1 Background: Turkic, Previous Analyses, and the Turkic Verb
2 Past tenses in Uzbek and Kazakh
3 Ekan/Eken: the Copular Perfect and its Four Meanings
4 Analysis
5 Conclusions

1 Background

Why Uzbek and Kazakh?

Uzbek and Kazakh are both Turkic languages, but Uzbek belongs to the Southeastern branch, and Kazakh to the Kipchak branch.

There exists a “Eurasian evidentiality belt”, which is roughly coterminous with Turkic-speaking regions, yet little research has been done on the Central Asian portion of this belt.

Uzbek and Kazakh are generally understudied, especially outside of their basic morphology.

Uzbek and Kazakh are remarkably similar in terms of how they express evidentiality and related meanings. It is therefore possible to combine our discussion of them. When both cognate forms from both languages are mentioned, Uzbek comes first, then Kazakh (e.g. -gan/-GAn). Data has been transliterated.

Data Sources

Uzbek and Kazakh data comes primarily from:

Native speakers of both languages, with data collected from speakers in Chicago and the surrounding areas, in Almaty, Kazakhstan, and online correspondence.

Literature, both translated literature and literature originally written in Uzbek and Kazakh.

The Internet, a great source spontaneous, searchable language with better context than mere elicitation or formal literature
**Key Terms**

**evidentiality** grammaticalized information source; can refer to a category or a type of meaning

**(non-)confirmativity** speaker’s degree of confidence in utterance, willingness to vouch for information; a subtype of:

**status** Jakobsonian term, grammatical category concerned with speakers evaluation of the contents of an utterance. Roughly equivalent to:

**modality** term borrowed from logic, grammatical category concerned with speaker’s evaluation of the contents of an utterance as necessary or possible. More formally, it is concerned with the relationships between possible worlds.

**Previous Accounts of Evidentiality in Turkic**

Previous accounts of evidentiality focus either on Turkish, or on the family as a whole.

A distinction is made between a simple past *-DI and a perfect *

*-MIš or *-GAn.

These past-tense morphemes are often described in the following ways:

**Past -di/-DI**: firsthand, direct

**Perfect -gan/-GAn, *-MIš, *-GAn**: non-firsthand, indirect, mediatif

But these analyses fail:

Perfec ts can be used to refer to events that the speaker witnessed...

(1) *Men keše düken-ge bar-ğan-mïn* (Kaz)
     *I yesterday store-DAT go-PST-1SG*
     *I went to the store yesterday.*

...and pasts can refer to events that speakers have not witnessed.

(2) *Huddi shu serial o’tgan oy-lar-da Turkiya kanal-i-da just that serial past month-PL-LOC Turkey station-3-LOC ham ber-il-di, lekin ko’r-ma-di-m. (Uz)*
     *also give-PASS-PST, but see-NEG-PST-1SG*
     *That serial was also shown on the Turkish station, but I didn’t see it.*

Accounts of other languages within this Eurasian evidentiality belt employ the sub-category of (NON-)CONFIRMATIVITY to more adequately capture the relevant features of these past tenses. This is the analysis used here.
The Turkic Verbal System

Predicates may be verbal or non-verbal (i.e. nouns, adjectives). Verbal morphology can be broken down into two types: finite and non-finite.

<table>
<thead>
<tr>
<th>Finite</th>
<th>Non-Finite</th>
</tr>
</thead>
<tbody>
<tr>
<td>occur only in predicative positions or act as nouns, adjectives, or conversbs</td>
<td>occur in predicative positions or act as nouns, adjectives, or conversbs</td>
</tr>
<tr>
<td>take either possessive agreement or idiosyncratic agreement</td>
<td>take pronominal agreement</td>
</tr>
</tbody>
</table>

Converbs are V+V constructions found in Turkic and other languages. They may indicate simultaneity or sequence of events. Sometimes, verbs in converb constructions take on grammatical meaning and indicate aspect, taxis, or ability, much like English modal verbs.

Finite Non-Finite

- Finite
- Non-Finite

Examples

<table>
<thead>
<tr>
<th>PST</th>
<th>PRF</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>-di/-DI</td>
<td>-gan/-GAn</td>
<td></td>
</tr>
<tr>
<td>COND:</td>
<td>AOR:</td>
<td></td>
</tr>
<tr>
<td>-sa/-SA</td>
<td>(-a)r/-(A)r</td>
<td></td>
</tr>
<tr>
<td>OPT:</td>
<td>CVB:</td>
<td></td>
</tr>
<tr>
<td>-sin/-sln</td>
<td>-(i)b/-(I)p</td>
<td></td>
</tr>
</tbody>
</table>

Non-finite forms of the verb and non-verbal predicates both take pronominal agreement and may co-occur with copular forms of the verb.

Examples

<table>
<thead>
<tr>
<th>PST</th>
<th>PRF</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>-di/-DI</td>
<td>-gan/-GAn</td>
<td></td>
</tr>
<tr>
<td>COND:</td>
<td>AOR:</td>
<td></td>
</tr>
<tr>
<td>-sa/-SA</td>
<td>(-a)r/-(A)r</td>
<td></td>
</tr>
<tr>
<td>OPT:</td>
<td>CVB:</td>
<td></td>
</tr>
<tr>
<td>-sin/-sln</td>
<td>-(i)b/-(I)p</td>
<td></td>
</tr>
</tbody>
</table>

In Old Turkic, the copula had the form *er- and was fairly productive. In the modern Turkic languages, the copula has either been lost entirely, or heavily restricted in the types of morphology it can take. There also exists a fully productive verb bo’l-/bol- that means roughly be or become that is sometimes used where we might expect a copula.

Past Tenses in Uzbek and Kazakh

Uzbek and Kazakh have three past tense forms which differ primarily in markedness for confirmativity.

<table>
<thead>
<tr>
<th>PST</th>
<th>PRF</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>-di/-DI</td>
<td>-gan/-GAn</td>
<td></td>
</tr>
<tr>
<td>COND:</td>
<td>AOR:</td>
<td></td>
</tr>
<tr>
<td>-sa/-SA</td>
<td>(-a)r/-(A)r</td>
<td></td>
</tr>
<tr>
<td>OPT:</td>
<td>CVB:</td>
<td></td>
</tr>
<tr>
<td>-sin/-sln</td>
<td>-(i)b/-(I)p</td>
<td></td>
</tr>
</tbody>
</table>

Definiteness is a feature only in Uzbek.
Confirmativity is often interpreted as indicating firsthand/witnessed evidentiality. This is incorrect.

(5) **Huddi shu serial o’tgan oy-lar-da Turkiya kanal-i-da**
    just that serial past month-PL-LOC Turkey station-3-LOC
    **ham ber-il-di, lekin ko’r-ma-di-m.** (Uz)
    also give-PASS-PST, but see-NEG-PST-1SG
    ‘That serial was also shown on the Turkish station, but I didn’t see it.’

Non-confirmativity is likewise not equivalent to non-firsthand information source.

(6) **Biraq, ökiniške oray kešig-ip qal-îp-pîn** (Kaz)
    But unfortunately be.late-CVB PFV-CPST-1SG
    ‘But, unfortunately, I was late.’

The **confirmative** feature is concerned with the expression of a speakers willingness to vouch for the contents of an utterance. Forms unmarked for this feature (as in *gan/-GAn*) require context to interpret. [-CONFIRMATIVE] markers (*-(i)b/-(I)p*) express:

- doubt
- surprise
- non-firsthand information source
- non-volitionality

Perhaps the most interesting non-confirmative forms, however, are the converbial forms *ekan/eken*.

3 Ekan/Eken

Likely derived from older *er-GAn*, *ekan/eken* have similar distribution to other copular forms. These forms perform four major functions:

<table>
<thead>
<tr>
<th>Evidential</th>
<th>Declarative</th>
<th>Interrogative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotive</td>
<td>Evidential</td>
<td>Evidential Q</td>
</tr>
<tr>
<td></td>
<td>Admireative</td>
<td>Rhetorical Q</td>
</tr>
</tbody>
</table>

**Evidential Statements**

Evidential statements express non-firsthand information source, as shown in this table from Willet (1988):


<table>
<thead>
<tr>
<th>Firsthand</th>
<th>Non-Firsthand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal</td>
<td>Indirect</td>
</tr>
<tr>
<td>Direct</td>
<td>Hearsay</td>
</tr>
<tr>
<td>Experience</td>
<td>Evidence</td>
</tr>
</tbody>
</table>

This range of meaning encompasses both indirect evidence (assumption, inference)... 

(8) *Etistik-ter-di žaqsï kör-e-di eken-siq.* (Kaz)
Verb-PL-ACC good see-PRES-3 EVID-2SG
‘You must really like verbs.’

...and hearsay or report.

(9) *Aytishicha u juda boy ekan.* (Uz)
Reportedly he very rich EVID
‘He’s reportedly very rich.’

*Emish* in Uzbek expresses a similar range of meaning and has similar distribution, but is restricted to expressing hearsay, rather than all kinds of non-firsthand evidence. The combination of hearsay and non-confirmativity often results in gossip:

(10) *Go’yoki men “hezalak” emish-man.* (Uz)
although I impotent REPORT-1SG
‘Although I’m apparently “impotent”.’

The Kazakh cognate *-mIs* has become a marker of pure reportativity with no non-confirmative meaning left. It has become a clitic and attaches to any sentence-final element.

(11) *Olar-diʃ uʃa˘ ğ-a aspan-da bir ʒer-de ʒoʃal-ip ket-ip-ti-mis.* (Kaz)
they-GEN plane sky-LOC a place-LOC disappear-cvb go-CPTST-3-REPORT
‘Their plane (reportedly) disappeared somewhere in the sky.’

**Evidential Questions**

Evidential questions are concerned with questioning the hearer’s knowledge or with the exchange of hearsay. They appear not to be marked as non-confirmative, but instead relate more to the knowledge of the hearer. They are often employed to create open questions...

(12) *Baqiţţan-di kör-gen-der bar ma eken?* (Kaz)
Baqiţţan-ACC see-NMLZR-PL EXIST Q EVID
‘Is there anyone who has seen Baqiţţan?’

---


Uzbek *emish* and Kazakh *-mIs* are relics of Old Turkic *-mIš*, which was fully productive and could directly attach to verbs stems.

This is no longer possible in either modern language: Uz. *’kel-mish*, Kaz. *’kel-mis*.


...or in questions concerned with continuing discourse about events that neither the speaker nor hearer has witnessed.

(13)  

–*Kechasi Saymon Munen bilan Tis Boylni hojatxonada ko‘rib qolishibdi.*

(14)  

–*They were caught* with Simon Moonan and Tusker Boyle in the square one night.

–They were caught

Bolalar unga hayron bo‘lib qarashdi:

–Ko‘rib qilishibdi?

–Nima qilayotgan *ekan* ular?

The fellows looked and him and asked:

–*Caught?*

–*What doing?*

Etti aytldi:

–*Smuggling.*

(15)  

(16)  

(17)  

A major difference between evidential statements and questions is that in questions, the otherwise non-confirmative *ekan/eken* may co-occur with confirmative past tense forms *di/-DI*. This suggests that in questions, *ekan/eken* do not express non-confirmativity.

There’s some debate over what Joyce means by *smuggling*. Glossed, the verb’s Uzbek translation is:

(18)  

£opi-sh-iib-apotgan *ekan*

‘They were mutually covering each other.’

A major difference between evidential statements and questions is that in questions, the otherwise non-confirmative *ekan/eken* may co-occur with confirmative past tense forms *di/-DI*. This suggests that in questions, *ekan/eken* do not express non-confirmativity.

(14)  

£opi-sh-iib-apotgan *ekan*  
cover-coop-recip-prog evid  
‘They were mutually covering each other.’

Admirativities are linguistic marking of unexpected information (Friedman 1981, DeLancey 1997). *Ekan/eken* may mark admirativity in addition to non-firsthand information source.

Admiratives have the same formal properties as evidential statements.

(15)  

(16)  

(17)  

(19)  

(20)  

While formally quite different, admiratives in Uzbek and Kazakh bear a similar range of meaning to the exclamatives found in European languages (Zanuttini & Portner 2003).

Uzbek *enish* may form admiratives, Kazakh -*mIs* cannot.
Rhetorical Questions

The final function of ekan/eken is to create rhetorical questions. Rhetorical questions are often thought of in terms of Sadock’s (1971) queclaratives, in which a question is equivalent to a statement of the opposite polarity.

(19) O’sha inson o’z-i Insof nima-lig-i-ni
that man self-3 fairness what-NMLZR-3-ACC
bil-ar-mi-kan? (Uz)
know-AOR-Q-EMOT
‘Does he know what fairness is?’ ‘He doesn’t know what fairness is.’

But rhetorical questions constitute a broader class than just queclaratives, and broadly conceived are any question that does not anticipate an answer. In addition to queclaratives, there exists a class of introspective questions. These can be translated with the English verb ‘wonder’.

(20) Išinde ne bar eken? (Kaz)
inside what exist EMOT
‘What is inside?’ ‘I wonder what is inside.’

(21) Hozir uy-lar-i-da bu haq-da munozara
now home-PL-3-LOC this claim-LOC dispute
qil-ish-ayotgan-mi-kan. (Uz)
do-COOP-PROG-PROG-Q-EMOT
‘I wondered if they were arguing at home about that.’
(Joyce 2007, 8)

Just like evidential questions, rhetorical questions allow for the combination of confirmative forms (-di/-DI) and ekan/eken:

(22) Kim bar-di eken? (Kaz)
who go-PST EMOT
‘Who (on earth) went?’

5 Analysis

When we look at all of this data, we are left with to main questions:
• How can we account for these disparate meanings/useages (evidential statements, evidential questions, admiratives, and rhetorical questions)?

• Why can ekan and eken co-occur with confirmative forms in interrogatives, but not in declaratives?

Declaratives with Ekan/Eken

The primary meaning of ekan/eken is non-confirmativity; when the speaker has witnessed the event described in the utterance, non-confirmative forms are interpreted as having meanings of surprise, irony, etc.: Admirativity.

When the speaker has not witnessed the event described in the utterance, non-confirmativity is interpreted as indicating a non-firsthand source.

Emotivity

There are three traditional functions of language:

Referential the use of language to describe things (e.g. statements)

Conative hearer-oriented; prompts a response from the hearer (e.g. commands, questions)

Emotive speaker-oriented; the production of “a certain emotion, whether true or feigned”

Exclamatives are the canonical instances of language being used in its emotive function (Andueza & Gutiérrez-Rexach 2010). Exclamatives, however, are just sentence types (à la Sadock & Zwicky 1985)—grammatical correlates to certain speech acts.

I propose that admiratives and rhetorical questions be considered examples of emotive language as well:

• Admiratives and exclamatives both express “a sense of surprise,” “unexpectedness,” “extreme degree,” and the like (Zanuttini & Portner 2003).

• Rhetorical questions have the form of a question, but expect no response; they are speaker-oriented.

• The purpose of exclamatives, admiratives, and rhetorical questions is not to convey information or to contribute to linguistic turn-taking, but to express the speaker’s inner state.

This method of accounting for both evidentiality and admirativity was formulated by Darden (1977) on the basis of Bulgarian data.

Jakobson (1960) also includes the phatic, metalingual, and poetic functions.

Exclamatives in European languages are signaled by the presence of a wh-element:

(23) How beautiful this day is! English

(24) Wie angenehem est ist! German

‘How nice it is!’

(25) Ce om placet este! Romanian

‘What a nice man he is!’

(26) Zein etxe ederra ikusi dudan! Basque

‘What a nice house I saw!’
Interrogatives with Ekan/Eken

Neither type of question with ekan/eken (evidential questions, rhetorical questions) seems to be accounted for by the concept of non-confirmativity; non-confirmativity seems to be incompatible with questions.

(27) Kim Moskou-ni yaxshi kor-a-di ekan? (Uz)
   who Moscow-ACC like-PRES-3 EMOT/EVID
   'Who (the hell) likes Moscow?'
   'Who (reportedly) likes Moscow?

(28) Who (the hell) likes Moscow? ≠ Who (will I not confirm)
   likes Moscow?

(29) Who (reportedly) likes Moscow? ≠ Who (will I not con-
   firm) likes Moscow?

Moreover, in questions, ekan/eken can co-occur with the confirmative past tense; this is not possible in declaratives and looks like a clash of semantics:

(30) Kim bar-di eken? (Kaz)
   who go-PST EMOT/EVID
   'Who (on earth) went?'
   'Who (reportedly) went?'

Cross-linguistically, however, it has been noted that subjective types of modality does not appear in a two sentence types: POLAR QUESTIONS and CONDITIONALS (Lyons 1977).

In Uzbek and Kazakh, ekan/eken behave strangely in questions, as seen in the previous examples.

When combined with conditionals, the interpretation is desiderativity or deonticity, not non-confirmativity:

(31) Qaysi kino-ni kor-sa-m ekan? (Uz)
   which movie-ACC see-COND-1SG EVID
   'Which movie should I see?'

What this data suggests is that these markers of non-confirmativity express a subjective type of modality, and that Jakobsonian status should be seen as equivalent to the subjective type of EPISTEMIC MODALITY.

The incompatibility of these non-confirmative forms with conditionals and with certain types of questions has led to the development of strategies to repair this incompatibility.

So, when non-confirmative forms are combined with interrogativity, the secondary meaning of these non-confirmative forms (evidentiality, emotivity) becomes the primary meaning (resulting in evidential questions and emotive questions).

Lyons (1977) claims that this incompatibility arises because subjective modals do not contribute to truth conditions, whereas objective modals do. Papafragou (2006) claims that both types contribute to truth conditions, but that subjective modality is "indexical, in the sense that the possible worlds in the conversational background are restricted to what the current speaker knows as of the time of utterance."

It appears that this strategy for repairing polar questions has become generalized to all questions in Uzbek and Kazakh.

The desiderative result of combining ekan/eken and conditionals is similar to English constructions like 'If you would X' or 'If he might X', where the combination of a modal and a conditional produce a desiderative or deontic reading.
5 Conclusions

Is evidentiality a valid category in Uzbek and Kazakh?

- No. Evidentiality (or, rather, non-firsthand information source) is a secondary meaning expressed by morphemes whose primary meaning is non-confirmativity.

- Non-confirmative morphemes are used in a wide variety of contexts to express a number of things, including admirativity, doubt, and non-volitionality.

Can Uzbek and Kazakh tell us anything about how evidentiality is expressed elsewhere in Eurasia or within the Turkic language family?

- Yes. We know how evidentiality functions in Turkish, Balkan Slavic, and an number of other languages in Eurasia (Friedman 1979, 1988). What they have in common with Uzbek and Kazakh is:
  - “Evidential” markers express admirativity (and sometimes other things) and can best be described as non-confirmative
  - These evidential markers are all past tenses or derived from past tenses

- Cursory examinations of Kyrgyz and Uyghur suggest that they express evidentiality in much the same way as Uzbek and Kazakh. Tajik looks more like Turkish and Balkan Slavic, while other languages require further research.

What can Uzbek and Kazakh tell us about evidentiality in general?

- Uzbek and Kazakh provide further support for the close relationship between “evidentiality” and (subjective) modality/status.

- If, as I claim, ekan/eken are markers of a type of subjective epistemic modality, they provide a starting point for further research into the subjective/objective distinction.

- There is the occasional link made between evidentiality and rhetorical questions: this needs further research. Moreover, it is rare that languages grammaticalize means of creating rhetorical questions; Uzbek and Kazakh provide some insight as to how they function and their scope of meaning.

There appears to be a separate Tibetan evidentiality area, where evidentiality is closely tied to the expression of person and volitionality (DeLancey 2001). The languages of the Baltics (Estonian, Lithuanian, Latvian) also have been claimed to possess grammaticalized evidentiality, but there it’s expressed via an inflected passive form of the verb (Kehayov 2008).

Cusco Quechua (Faller 2002) and ASL (Hoza et al. 1997) also have been claimed to have grammaticalized rhetorical questions. Nogay, a Turkic language, is interesting because it forms rhetorical questions using similar means to Uzbek and Kazakh, but distinguishes them from evidential questions (Karakoç 2005):

(32) a. Nege kel-gen eken-ler?
   why come-PRF EVID-PL
   ‘Why have they (reportedly) come?’

b. Nege kel-gen-ler eken?
   why come-PRF-PL EMT
   ‘I wonder why they have come.’
References


