

Dissertation Proposal - 22 January 2008
Christopher A. Straughn

Committee:
Victor A. Friedman, Chair
Kağan Arık
Lenore Grenoble
Salikoko Mufwene

Evidentiality in Kazakh and Uzbek

0. Introduction

Evidentiality is generally defined as the linguistic encoding of information source. Although information source may be expressed in any language, it has been claimed that approximately one-quarter of the world's languages possess a *grammatical* means of encoding evidentiality (Aikhenvald 2004:1). Much like other grammatical categories (GENDER, NUMBER ASPECT, TENSE), languages vary with regard to whether the speaker's source of information is obligatorily encoded and in how many different ways a particular language expresses this kind of information. Also like other grammatical categories, evidentiality interacts with other verbal categories, particularly STATUS/MODALITY, MOOD, TENSE, and ASPECT. The Turkic languages, and those which have grammaticalized evidentiality as a result of contact with the Turkic languages, possess rather unusual means of encoding evidentiality (c.f. Friedman 1981, 1999, 2000, inter alia). Rather than having morphemes exclusively devoted to expressing evidentiality, these languages express evidentiality either through a speaker's commitment or non-commitment to confirmation (in which case evidentiality is merely a highly salient contextual meaning) or through more complex pluperfect-like constructions. Based on current descriptions of Uzbek and Kazakh, it is expected that

they, too, will behave in certain respects like Turkish and other genetically or areally related languages.

Before any study of Kazakh and Uzbek may be undertaken, it is first necessary to examine what has previously been written. Although Uzbek and Kazakh do not appear to fit perfectly into certain previously posited definitions, it is necessary to understand what evidentiality is if it, or what passes for it, is to be identified in these languages. We will therefore start with a general definition of evidentiality. Following that, a thorough examination of evidentiality in a number of genetically and areally proximate languages will be given. As many grammars employ analyses similar to that for Turkish (e.g. Coşkun 2000, Koç & Doğan 2004), it is important to understand the data given for Turkish. And as much of what has been written about Turkish refers to the similar systems in Bulgarian and Macedonian, which Turkish is often claimed to have influenced, these languages will also be used as points of reference. As it may be possible to posit that the evidentiality found in Turkish and Balkan Slavic is similar to that in other parts of Eurasia, the possibility of a broader, Eurasian style of evidentiality will also be discussed.

Having examined the systems of other languages, we will turn to Uzbek and Kazakh themselves. What has been written about the verbal systems of these two languages is frequently vague or contradictory, but some insights may be made into what constructions are likely to bear evidential meanings and deserve further study. The prospects for further study and areas to be investigated will follow, along with a tentative outline for the structure of the dissertation.

1. What is Evidentiality?

As stated above, evidentiality is generally defined as the linguistic encoding of information source. However, it is rare that any linguistic phenomenon is so simply defined, and evidentiality, in particular, has had a contentious history. A brief account of these definitions and their typological implications is given below.

The first known mention of anything resembling evidentiality (and, in fact, almost identical to Jakobson's [1957] definition nine centuries later) was given in reference to the Turkic languages. Al-Kāšġarī, in his 11th Century *Dīvān Luġāt at-Turk*, distinguished two forms of the past tense in Turkic, namely those forms corresponding to modern Turkish *-DI* and *-mİŝ*, stating that “the difference between these two forms is that the *dāl yā*’ [-*DI*] on preterite verbs indicate that the action occurred in the presence of the speaker. The action is verified by its occurrence in his presence” whereas “*Mīm ŝīn* - [*mİŝ*], on the other hand, indicate that the action occurred in the absence of the speaker” (Al-Kāšġarī 1982: 297, as cited in Friedman 2003: 189). While this description includes elements of information source (firsthand vs. non-firsthand), it also covers an important distinction that is more likely the primary one for these morphemes: verificational (confirmative) vs. non-confirmative (Friedman 1978).

In modern times, evidentiality was first explicitly noted by as a distinct category by Franz Boas in his grammar of Kwakiutl. He states: “To the suffixes expressing subjective relation belong those expressing the source of subjective knowledge -- as by hearsay, or by a dream” (1911: 443 - emphasis added). The term evidential as a verbal category was first employed by Jakobson in 1957, who, drawing on the work of Boas, defines it as a speaker’s report of an event “on the basis of someone else’s report (quotative, i.e. hearsay evidence), of a dream (revelative evidence), of a guess

(presumptive evidence), or of his own previous experience (memory evidence)”. In Jakobson’s system this is expressed as $E^n E^{ns} / E^s$ - the characterization of a narrated event and a narrated speech event (the source of information) with respect to a speech event (1957/1971: 135). This definition is the basic starting point for virtually all subsequent studies of evidentiality.

As noted above, most, but not all later works usually base their definition of evidentiality on that of Jakobson. A major, fairly recent publication is that of Chafe and Nichols (1986). While Chafe and Nichols opt not to define evidentiality themselves, they implicitly endorse one of their authors, Jacobsen (1986:3), who expands Jakobson’s definition of evidentiality by deeming it “a linguistic category which applies to predications that the speaker assumes have a reasonable likelihood of being true, but which he cannot vouch for out of direct observation or experience” This definition, which reflects that already given in Aronson (1967) and Friedman (1977), diverges from Jakobson’s formulation of evidentiality as being strictly a matter of information source, and expands it to include speaker confidence. This notion of speaker confidence is one that Aronson (1977) and Friedman (1977) argued should be considered part of what Jakobson included under the category of STATUS (E^n - qualifying a narrated event without reference to the speech event or involving the participants)¹ (1957/1971: 134).

Another article in Chafe and Nichol’s volume discusses a number of criteria for identifying evidentials. The criteria given by Anderson (1986: 274-5) are generally in accord with Jakobson’s notion of information source:

¹ Jakobson mentions Macedonian and Bulgarian in discussing evidentiality, and, citing Lunt (1952) he notes that the opposition in those languages is not one of firsthand and non-firsthand source, but rather one of “vouched for” and “distanced”.

- (1a) Evidentials show the kind of justification for a factual claim which is available to the person making that claim, whether
 - direct evidence plus observation (no inference needed)
 - evidence plus inference
 - inference (evidence unspecified)
 - reasoned expectation from logic and other facts
 and whether the evidence is auditory, or visual, etc.
- (1b) Evidentials are not themselves the main predication of the clause, but are rather a specification added to a factual claim ABOUT SOMETHING ELSE.
- (1c) Evidentials have the indication of evidence as in (a) as their primary meaning, not only as a pragmatic inference.
- (1d) Morphologically, evidentials are inflections, clitics, or other free syntactic elements (not compounds or derivational forms)

Criterion (1a) relates specifically to information source, while (1b) is necessary to rule out complement constructions, such as *I saw that...* or *He inferred that...* In these examples, information source is specified (vision, inference), but the verbs *see* and *infer* bear predicational force and should not be considered evidentials.

Johanson and Utas' (2000) volume represents a shift in studies of evidentiality, as it focuses on only two language families: Turkic and Iranian. The distinct forms of evidentiality in these languages, and the interplay of evidentiality and other categories (ASPECT, MOOD), leads to their rather vague characterization of evidentiality as "the expression of subjective 'experience', more specifically the presentation of a situation 'by reference to its reception by a conscious subject'" (2000:v). The notion that evidentiality in Turkic is somehow special is reinforced in Johanson's own contribution to this volume; he sees the 'indirectivity' in Turkic as distinct from other forms of evidentiality:

The source of knowledge--the way in which the event is acknowledged by the conscious subject in question--is not criterial; it is unessential whether the reception is realized through hearsay, logical conclusion or direct perception. The

receiver P is not necessarily the speaker, P^s, but may also be a participant of the narrated event, Pⁿ. Consequently, markers of indirectivity do not fit into narrow evidential schemes primarily distinguishing between the speaker's non-first-hand and first-hand information (Johanson 2000: 61).

The notion that the speaker is inessential in 'indirectivity' does set this concept apart from evidentiality, which, as defined by Jakobson, requires that the speaker, and no one else, be the recipient of non-first-hand information. However, this notion of indirectivity is closely related to Aronson's reinterpretation of Jakobson's assessment of direct and indirect narration in Bulgarian (based in turn on Friedman's [1977] account of Macedonian), in which a narrated event and a speech event may or may not be equivalent, and no reference to the speaker is necessary (Aronson 1991: 114).

Among the most recent of the typological works on evidentiality are those of Aikhenvald (2003, 2004). Although she approaches evidentiality from an Amerindianist perspective, her typological breakdown of evidentiality is worth examining. Working from the notion that information source is broken down into six basic types, she posits a sort of scale, from most personal to least personal, by which evidentials may be categorized (2004: 65):

(2) *Visual * Non-Visual Sensory * Inference * Assumption * Hearsay * Quotation*²

According to Aikhenvald's system any language that expresses evidentiality will possess means of expressing any number of distinctions within that system. If a morpheme may express multiple sources of information, then those sources will be adjacent on the scale

² Note that this scale of most personal to least personal naturally corresponds to a scale of most reliable to least reliable.

above. An example of how this works is shown below for systems that express two sources of information (2004: 65):

Table 1 - Two Term Systems

	I. Visual	II. Sensory	III. Inference	IV. Assumption	V. Hearsay	VI. Quotation
A1	firsthand		non-firsthand			
A1	firsthand	non-firsthand				
A1	firsthand		non-firsthand		other or <no term>	
A2	<unmarked>		indirect			
A3	<unmarked>				reported	
A4	<no term>	non-visual	<no term>		reported	

A major issue with Aikhenvald's typology can be seen in the table above. In counting "terms", she sets a lower limit of two. However, she is generally inconsistent in counting the unmarked term. If an unmarked form does not interact with the evidentiality system, there should, then be a one-term system. Type A2 above (to which Aikhenvald assigns Turkish, Bulgarian, and Tajik) includes systems in which a statement may be marked or neutral with regard to information source. This issue was indirectly addressed in her 2003 work, in which she divides evidential systems into two types:

- I. Types which state the existence of a source without specifying it (Turkic, Balkan Slavic, Uralic, Iranian)
- II. Types which specify the kind of evidence, e.g. visually obtained, inferred, reported

In a system like Type I or A2, the distinction between sources of information is not as important as it would be in more complex systems; what is important is markedness and non-markedness of a given form.

2. Evidentiality in Central Eurasia

Much of what has been written about evidentiality has focused on the Balkans, the Middle East, the Caucasus, and Central Asia. For lack of a better term, this region will be referred to as Central Eurasia. The type of evidentiality in this region is quite different from that found in other parts of the world, especially the types found in the Americas, New Guinea, and East Asia. The languages spoken in this region possess phenomena that are so marginally evidential that they have received labels like *direct/indirect* (Johanson 2000) and *confirmative/non-confirmative* (Aronson 1967). If one were to apply Aikhenvald's 2004 typology to these languages, they would appear to possess simple two-term evidential systems; however, the distinctions encoded in these systems is much more complex. The examples given here will be drawn primarily from Turkish and Balkan Slavic, as these languages have been best described with regards to the phenomena that will be discussed.

The starting point for the analysis of evidentiality in these languages is the past tense. In Macedonian and Bulgarian, the non-resultative and non-taxic past tenses are divided into paradigms that are traditionally labeled the *definite past* and the *indefinite past*. These correspond quite well to the (rather inappropriately named) *görülen geçmiş zaman* (seen past tense) and *duyulan geçmiş zaman* (heard past tense) in Turkish. These forms have been analyzed extensively in the literature, and have been said to possess corresponding meanings of *confirmative* and *non-confirmative* (although this is not an entirely correct assessment of these forms), as well as *direct* and *indirect* (which is also rather misleading) (Johanson 2000). Examples of the 3rd singular past tense of the verb “do” are given in the table below:

Table 2 - The Past Indicative of “Do”

	definite/seen/confirmative	indefinite/heard/non-confirmative
--	----------------------------	-----------------------------------

Bulgarian	napravi	(e) napravil
Macedonian	napravi	napravit
Turkish	yaptı	yapmış

Note that in Balkan Slavic the indefinite terms are often referred to as the *l-forms* and in Turkish as the *miş-form*³. Where necessary, these morphemes will be highlighted.

The forms have often been regarded as participating in an evidential system, that is, expressing source of information. It is along these lines that Johanson has termed the Turkish forms direct and indirect, referring to information source (direct experience vs. hearsay, inference, etc.). However, this characterization is incorrect, as information that must have been obtained from non-firsthand sources may be expressed with the definite/seen/confirmative form:

- (3) Ben baloda yok**tum**. Baj Ganü gitti (Konstantinov 1972, in Friedman 1981)
I was not at the ball. Baj Ganjo went.
(Turkish)

Moreover, the indefinite/heard/non-confirmative forms may express information that the speaker has obtained directly, although these forms often express something like surprise, awe or sarcasm:

- (4a) Toj bil bogat! (Friedman 1980: 7)
(Bulgarian, Macedonian)
(4b) O zengin **imiş**!
(Turkish)
Why, he's rich!

A strict firsthand/non-firsthand analysis of these forms is therefore ruled out.

The analysis I will adopt here is one employed by Friedman (1978). The most basic difference between these two paradigms is that the simple past

³ Turkish possesses vowel harmony so the relevant morpheme may be manifested as *miş*, *mış*, *muş*, or *müş*, or as *di*, *di*, *du*, or *dü*, plus any number of person suffixes.

(definite/seen/confirmative) is actually the more marked of the two morphemes and bears a *confirmative* meaning. This meaning is evident in that the only place the simple past may not occur is in an antiaffirmative context:

- (5) Se somnevam deka toj go **napravi/napravit* toa. (Friedman 1978: 109)
I doubt that he did it.
(Macedonian)
- (6) Inanmıyorum ki o adam bunu **yaptı/yapmış*. (Friedman 1978: 110)
I don't believe that he did it.
(Turkish)

The so-called perfect⁴ form (indefinite/heard/non-confirmative), then, stands in contrast to the simple past in its unmarkedness for confirmativity. As seen in the examples above, it may occur in clauses subordinated under antiaffirmative verbs. However, its use is not strictly limited to non-confirmative contexts:

- (7) Jas *sum stanal* toku vo dva časot nokešta. (Lunt 1952: 96 in Friedman 1978: 109)
I *got up* at exactly two o'clock last night.
(Macedonian)

The non-confirmative forms derive other meanings (aside from pastness or perfectness) in that the speaker employing them has chosen not to employ a form marked for confirmativity. The basic meaning of these forms, however, remains one of anteriority.

More complex forms can be created by extending the use of the non-confirmative to create pluperfect forms. In Turkish, the suffix *-miş* may be combined with dummy copula *i-* (*imiş*) or cliticized to essentially any predicate complex as well. These forms have a similar range of meanings to the regular perfect or pluperfect, although “it appears that the combination of taxis (marking for anteriority) and resultativity, i.e. the double marking of a past auxiliary and a past participial form, can result in the pluperfect’s

⁴ These forms are so called because they are historically derived from perfect morphemes; while perfectivity or resultativity may still be expressed by these morphemes, a number of other meanings are available and must be contextually determined.

functioning as a genuinely pure reported form. This is to say that that which in the perfect may be the chief contextual variant meaning occurring in the least marked contexts becomes the invariant meaning in the more highly marked pluperfect” (Friedman 1979: 345, c.f. Friedman 1977: 120)⁵. The combination of two markers of anteriority in Turkish result in the following contrasts in meaning:

- (8) Ayşe gelmişti.
Ayşe has/had come.
- (9) Ayşe gelmişmiş.
Ayşe has/had (reportedly) come.

This distinction is found in Turkish grammars, which distinguish between the simple non-confirmative past, which is referred to *duyulan* (heard), and the more complex forms, which are referred to as *rivayet* (hearsay).

There are, then, three primary meanings that are borne by the historical perfect markers in Turkish and Balkan Slavic:

- (10) Anteriority and Resultativity (Pastness, Perfectness)

Reportedness - this most evidential-like meaning is not, as been claimed, the invariant meaning of these forms, it is rather the “chief contextual variant meaning” of these forms (Friedman 1981: 14).

Admirativity and Dubitativity

The last of these three meanings deserves some further explanation. Consider the example below:

- (11) Ti si bil maž! (Friedman 1980: 15)
You're quite a man!
(Macedonian)

⁵ Note that this is not the case in Georgian, where the pluperfect's primary meaning is modal. Although the Caucasus falls within what I have termed 'Central Eurasia', the positions of these languages in this area is much too complex to discuss here.

Example (11), and the examples in (4) express neither strict anteriority (although they could) nor reportedness. Rather, these examples express admirativity. In Albanian, which has a marked admirative series but no marked confirmative, the admirative paradigms are “markedly nonconfirmative...which, while taking ‘surprise’ as its most common meaning, can also express sarcasm, inference, reporting etc” (Friedman 2003: 192). Admirativity (under the alias of ‘mirativity’) has been claimed to be a separate verbal category expressing only surprise (DeLancey 1997). However, there appears to be no language that possesses a purely admirative morpheme, so it will be assumed here that admirativity is merely one of a number of variant meanings that can develop from a perfect.

2.1. Verbal Categories

Much of the confusion over what is and is not evidentiality is due, in part, to a proliferation of terms related to verbal categories. Various proposals for defining verbal categories will be discussed here, as well as the implications of grouping various phenomena under one category or another.

Perhaps the most influential attempt at classifying verbal categories is that of Jakobson (1957/1971). In this classification he employs a four-term system:

- (12) E^n - narrated events
 E^s - speech events
 P^n - participants of narrated events
 P^s - participants of speech events

The term *evidentiality* was first defined in this system as “ E^nE^{ns}/E^s ” (Jakobson 1957/1971: 135); evidentiality, then, is characterized as source of information (a narrated speech event E^{ns}) and a narrated event in relation to the speech even. Another important category defined by Jakobson at this time is *status*, “which characterizes [qualifies] the

narrated even without involving its participants and without reference to the speech event” (E^n) (Jakobson 1957/1971: 134). Another relevant category is aspect, which is also symbolized as E^n , and which quantifies a narrated event. A final category is *mood* (P^nE^n/P^s), which “characterizes the relation between the narrated event and its participants with reference to the participants of the speech event (Jakobson 1957/1971: 135). A complete table of the categories defined by Jakobson is as follows:

Table 3 - Jakobson’s Verbal Categories

	<i>P involved</i>		<i>P not involved</i>		
	<i>Designator</i>	<i>Connector</i>	<i>Designator</i>	<i>Connector</i>	
<i>Qualifier</i>	P^n	Gender	P^nE^n Voice	E^n <u>Status</u>	E^nE^n Taxis
<i>Quantifier</i>		Number		E^n <u>Aspect</u>	
<i>Shifter</i>	P^n/P^s Person		P^nE^n/P^s Mood	E^nE^s Tense	E^nE^{ns}/E^s Evidential

Aronson (1967) proposed an alteration to Jakobson’s typology. Whereas Jakobson had characterized the indefinite past in Bulgarian as expressing evidentiality, Aronson considers the confirmativity expressed by this verb form a type of status. Aronson later suggested a number of changes to Jakobson’s typology. One of the most important of these changes was to redefine status as the “*subjective* evaluation of the narrated event by the speaker, i.e., E^n/P^s ” (1990: 114). Mood is then redefined as the qualification of E^n , in as much as mood has nothing to do with the speaker’s opinion, and more to do with the “*objective* evaluation of the narrated event” (1990: 115). Aronson carries on Jakobson’s characterization of aspect as quantifying a narrated event E^n . The category of evidential was completely eliminated in Aronson’s typology. Because what had traditionally been regarded as “evidential” appeared to be more closely related to the speaker’s evaluation of the utterance, Aronson suggested that the evidential “should be regarded as closely related to, or, better, a subvariety of, STATUS (E^n/P^s)” (1990: 116). He goes on to claim

that he knows “of no language that has a grammatical category that has evidential as its invariant meaning” (1990: 130). A summary of Aronson’s verb categories is as follows:

Table 4 - Aronson’s Verbal Categories

	Characterizing the participants of the narrated event				Characterizing the narrated event itself			
	With reference to the narrated event		Without reference to the narrated event		With reference to another narrated event		Without reference to the another narrated event	
	Qualita- tively	Quantita- tively	Qualita- tively	Quantita- tively	Qualita- tively	Quantita- tively	Qualita- tively	Quantita- tively
E ^s							E ⁿ /E ^s Tense	
P ^s			P ⁿ P ^s Person				E ⁿ /P ^s Status	[E ⁿ /P ^s] Evidential?
-E ^s , P ^s	P ⁿ /E ⁿ Version	P ⁿ /E ⁿ Oblique Valence	P ⁿ Class	P ⁿ Number	E ⁿ /E ⁿ Taxis		E ⁿ Mood	E ⁿ Aspect, Resultativity

Part of what appears to have complicated the discussion between those proposing that Balkan Slavic and Turkish possess evidentiality and those who actually work on Balkan Slavic may be the use of the term “status”. For all intents and purposes, status is equivalent to what is more commonly known as modality. Both terms refer to the speaker’s evaluation of the narrated event, including evaluations of truth, likelihood, potentiality, and necessity. However, scholars working on status or evidentiality argue that it is insufficient to suggest that evidentiality be treated as a subtype of modality. De Haan (1999), for example, distinguishes between the two phenomena on semantic and syntactic grounds. Semantically, epistemic modality is “the degree of commitment a speaker places in his/her utterance” while evidentiality encodes the source of information. Syntactically, modals and evidentials behave differently with regard to the scope of negation: evidentials cannot occur within the scope of negation, whereas modals can. Matthewson et al. (2007), disagree, however, and argue that evidentiality cannot be claimed to be distinct from epistemic modality and that epistemic modals encode a twofold distinction: they must choose either to encode quantificational strength (as they

are traditionally argued to do) or information source. Some recent work written from a generativist standpoint takes an intermediate stance. Cinque (1999), for example, posits dedicated syntactic slots for mood, modality, and evidentiality which fall into a sort of hierarchy:

(13) [Speech Act Mood [Evaluative Mood [Evidential Mood [Epistemological Mode]]]]

The highest projection, speech act mood, refers to the status of an event as realis or irrealis, evaluative mood expresses the speaker's evaluation of an event as good or bad, evidential mood gives the speaker's evidence for the truth of a proposition, and epistemological mode indicates a speaker's degree of certainty as to the truth of a proposition.

Regardless of whether or not one agrees with the notion of dedicated positions for given categories, there is a clear connection between a speaker's objective evaluation of a proposition (mood), subjective evaluation (modality), and source of information for realizing these evaluations (evidentiality). All three categories must be considered to understand evidentiality in Uzbek and Kazakh.

2.2. Further Areal Implications

As previously stated, it is important to understand the phenomena in Turkish and Balkan Slavic in order to understand Kazakh and Uzbek. Although the Kazakh and Uzbek are quite distant from the Balkans, the "evidentiality" in all of these languages appears to exist within a larger, areal framework. Prior studies have noted some sort of areal phenomenon in this region, dating at least as far back as Conev (1910/1911). Haarmann (1970) proposed the broadest linguistic area, suggesting a Eurasian isogloss defined by an "indirect experience-form" that includes Turkic, Mongolic, Tungusic, Uralic, and some Caucasian and Paleo-Siberian languages. Friedman (1979) focuses on a selection of

representative languages: Balkan Slavic, Turkish, Albanian (to some extent), Azerbaijani, Tajik, Avar, and Georgian (to some extent). Comrie (2000:1-2) also noted this areal phenomenon, noting that evidentiality in Eurasia (encompassing Turkic, Tungusic, Iranian, eastern Uralic, Balkan, and Caucasian languages) lacks some of the phenomena found in evidential systems elsewhere, such as fine-tuned distinctions based on source of information. He further notes that language contact may have played a role in spreading evidentiality between these languages. And Aikhenvald identifies a larger area within Eurasia characterized by a small evidential system located within a “largish ‘evidentiality belt’ spreading across the Balkans, the Caucasus, and Central Asia into Siberia (2004:290). She and others have suggested that the influence of the Turkic languages is responsible for the spread of this feature. Kehayov (2008) broke down Haarman’s larger linguistic area into four regions. He focuses the Balkans and the eastern Baltic, and also identifies two further regions: the Caucasus and the Volga-Kama area.

If, as Aikhenvald has posited, the Turkic languages are responsible for the spread of “evidentiality” throughout Central Eurasia, what has been called evidentiality in these languages deserves further study. Essentially very few studies have been devoted to evidentiality in most of the Turkic languages, the data for the most part being limited to brief mentions in grammars or in broader works, such as that of Haarman (1970). So far, there has been one volume devoted to evidentiality in the Turkic languages: Johanson and Utas (2000). Numerous works exist discussing evidentiality in Turkish (there are three papers in Johanson and Utas alone). Shorter works have been devoted to Salar (Dwyer 2000), Gagauz (Menz 2000), the Mishär dialect of Tatar (Tatevosov 2007), Altay

(Skribnik and Osnova 2007), and Cypriot Turkish, which is considerably different from Anatolian Turkish (Demir 2003).

Among the most important languages to consider in examining the areal context of evidentiality in Uzbek is Tajik. It has been speculated that evidentiality in Tajik has come about as a result of contacts with Uzbek, as the two languages have exhibited a great amount of influence upon each other and still coexist in diglossic situations in Uzbekistan and Tajikistan (Lazard 1996: 29). However, so little has been written on evidentiality in Uzbek that it is difficult to substantiate that claim. It is worth noting, however, that the system of evidentiality in Tajik is quite similar to that of Turkish and Balkan Slavic, which further supports the claim that the system of evidentiality in Uzbek is similar to that of Turkish and Balkan Slavic.

Per Friedman (1979), evidentiality in Tajik functions similarly to other languages of the region. There is a contrast between the simple past, which is confirmative, and the perfect, which is not marked for confirmativity. The complex perfect forms bear similar meanings to those in Turkish and Balkan Turkic, namely, reportedness admirativity, and dubitativity. Lazard (2000) illustrates the similarities between the normative description of Bulgarian and the Tajik system, which are remarkably similar. Note that he employs a neutral/mediative distinction, rather than a confirmative/non-confirmative distinction:

Table 5 - Bulgarian and Tajik Compared

	Bulgarian		Tajik	
	“Neutral” (Confirmative)	“Mediative” (Non-Confirmative)	“Neutral” (Confirmative)	“Mediative” (Non-Confirmative)
Present	<i>čéte</i>	<i>četjal</i>	<i>mekunad</i>	<i>mekarda-ast</i>
Imperfect	<i>četeše</i>		<i>mekard</i>	
Aorist	<i>četé</i>	<i>čel</i>	<i>kard</i>	<i>karda-ast</i>

Perfect	<i>čel e⁶</i>	<i>čel bil</i>	<i>karda-ast</i>	<i>karda buda-ast</i>
Pluperfect	<i>čel beše</i>		<i>karda bud</i>	

Examples of the uses of the Tajik forms are shown below:

- (14) xola pogo meomaday⁷. -ki guft? -rajab (Lazard 1996:29)
 aunt tomorrow come.COMPLEXPERF who say.PST Rajab
 Aunt is (reportedly) coming tomorrow. -Who told you? -Rajab.
- (15) pul-am na-buday (Lazard 1996: 29)
 money-my NEG-be.PERF
 Look, I have no money!

The Persian of Iran differs from Tajik in that the confirmative/non-confirmative distinction is found only in the past. As seen in (14), complex forms in Tajik are allowed in the present as well, much as in Turkish (where a reportative reading is also found). Moreover, the perfect in Tajik allows for an admirative reading (as in 15), which is not found in Persian (Lazard 1996).

3. Uzbek and Kazakh

Kazakh and Uzbek were chosen for this study for three reasons. The first is that the evidentiality systems in these languages have been scarcely studied outside of reference grammars and a few mentions in broader works on evidentiality. Any work on evidentiality in Uzbek and Kazakh will therefore be novel. The second reason for their study is that they play a central role in the evidentiality belt that has been proposed by Friedman (1979), Aikhenvald (2004), and others. As they are spoken in the heart of Central Asia, any further examination of evidentiality as an areal feature must take them into account. The third reason they have been selected is their place within the genetic

⁶ Note that the “neutral” perfect and the “mediative” aorist are not formally distinct in both Bulgarian and Tajik. See Aronson (1967) and Friedman (2002) for details on auxiliary loss in the third person in Bulgarian.

⁷ These forms are dialectal. The suffix *-ay* here is *-ast* in the standard language.

classification of Turkic. Although they are spoken in geographically contiguous areas, Kazakh and Uzbek belong to different branches of Turkic, making them prime candidates for understanding how evidentiality functions within a larger Turkic context. Kazakh belongs to the Kipchak branch of Turkic, and Uzbek belongs to the branch variously known as Eastern, Turki, or Chagatay (Tekin 2005).

Based on the data from geographically and genetically close languages, certain hypotheses regarding evidentiality in Uzbek and Kazakh may be made. If the verbal systems of these languages resemble those of Turkish and Balkan Slavic, then Uzbek and Kazakh will possess a past tense marked for confirmativity and an opposing perfect or perfect-like form which is unmarked for confirmativity. Complex forms based on the perfect (e.g. the pluperfect, future perfect) will have a stronger tendency to exhibit pure reportative meanings. The Tajik data given in Friedman (1979) and Lazard (1996, 2000) are quite similar to those of Turkish and Balkan Slavic, which lends support to the hypothesis that Uzbek and Kazakh likely fall into an evidentiality belt that extends from the Balkans through Siberia. Johanson (2003), however, complicates matters by stating that a number of Central Asian Turkic languages (Kazakh, Uyghur, Uzbek, and Turkmen) actually possess four morphemes expressing indirectivity and that these morphemes express different information. Two of these morphemes are suffixes that attach to the verb much like tense or aspect: **-Ib(dIr)*, which may be used in past tenses, and **-GAn*, which contributes a post-terminal (perfect or resultative) meaning. These forms are presumably in contrast with the past tense **DI*. The other two forms are based on a now largely defunct copula **er-*, which may combine with verbs in gerund or

participial forms. **Erken* (the copular form of *-*GAn*) is claimed to indicate inferential and reportative meanings while **ermiš* indicates reportative or quotative meanings.

3.1. The Turkic Predicate

Before discussing the specifics of Kazakh and Uzbek, something must be said about predication in the Turkic languages in general. In the Turkic languages, predicates may be nominal, adjectival, or verbal. These categories are not always easily discernable.

Examples of all of these are provided below:

- (16) **Nominal**
Ben doktor-um. (Turkish)
I doctor-1SG
“I am a doctor.”
- (17) **Adjectival**
Bu šarik kyhyl. (Sakha)
This balloon red.
“This balloon is red.”
- (18) **Verbal**
Ut kurtām. (Chuvash)
Horse see-PST-1SG
“I saw a horse.”

The two existential morphemes from Proto-Turkic **baar* (positive ~ there is) and **yook* (negative ~ there is not) deserve special mention. Reflexes of these existential morphemes are present in all Turkic languages. Aside from merely indicating the existence of a thing, they are commonly employed to express what in English what would be expressed using the word “have”. Their categorial status varies from language to language:

- (19) Pul-im bar. (Uzbek)
Money-1SG EXIST
“I have some money.”

- (20) *Ilinoj-da taw žoq.* (Kazakh)
 Illinois-LOC mountain NEG.EXIST
 “There are no mountains in Illinois.”

Another feature of the Turkic verbal system is the ability to sequence verbs in a manner similar to, but distinct from, serial verb constructions. In these constructions, which are usually called “converbs”, the first verb is marked with a converbial marker and is followed by another verb which is marked for tense, mood, aspect, etc. The series of verbs may indicate a sequence of actions, or the first verb may be modified by the second verb, which may indicate manner or cause. In many cases, the second in a series of verbs may grammaticalize into a marker of aspect, modality/status, version, etc. (Straughn 2008).

- (21) *Murodjon gugurt chaq-ib vlyuchkatel-ni top-di.* (Uzbek)
 Murodjon match strike-CVB switch-ACC find-PST.3SG
 “Murodjon struck a match and found the switch.”
 (Bodrogligeti 2003: 582)

In the example above, the first verb indicates the first event in a sequence (lighting a match), and the second verb indicates the following event (finding the switch). Note that in converbial constructions, the verbs share a subject but do not necessarily share an object. In the following examples, the second verbs in the series have grammaticalized and no longer bear the meanings of their lexical equivalents:

- (22) *Er-dan bahra ol-ib balki o’s-ib ket-ar.* (Uzbek)
 Man-ABL profit take-CVB maybe grow-CVB CONTINUE-AOR.3SG
 “He would take a profit from the man maybe it would keep growing.”
 (Straughn 2008)

In the above example, the verb *ket-* is used to indicate action continued into the future.

As a full verb, however, has a meaning of “leave” or “depart”. In some cases these

grammaticalized converbs become phonologically bound to the verb root, becoming fuller members of the verbal paradigm:

- (23) Gel-ebil-ir-im. (Turkish)
Come-ABIL-AOR-1SG
“I can come.”

In the example above, the independent verb *bil-*, which translates as “know” when it is an independent verb, has been grammaticalized into the verbal affix *-Abil-*, which indicates ability or possibility.

3.2. The Uzbek Verbal System

The Uzbek verb is fundamentally Turkic, although the phonetic shape of the various morphemes is not exactly the same as in other Turkic languages. This is due partially to the replacement of forms over time with new forms, and partially to the loss of vowel harmony and consonant alternations in the Uzbek phonological system.

As the past tense is locus of evidentiality (or, rather, confirmativity) in the other languages of the region, that will be discussed first. In Uzbek, there is a form *-di* that is cognate with Turkish *-DI* (the confirmative/simple past tense). Its paradigm is given in Table 6:

Table 6 - The Simple Past in Uzbek

<i>qil-</i> ‘to do’	SG	PL
1	qil-di-m	qil-di-k
2	qil-di-ng	qil-di-ngiz
3	qil-di	qil-di-(lar)

The negative paradigm is formed by inserting the morpheme *-ma-* after the verb stem:

- (24) qil-ma-di-m
“I did not do.”

The simple past, according to Coşkun, expresses witnessed action (2000:133), Kononov merely equates it to a simple past tense (1960: 215). Bodrogligeti’s definition of the

simple past (which he refers to as the “definite past tense”) is more in line with the confirmative readings suggested by Friedman for the simple past of Turkish and Balkan Slavic: “With the definite past tense the speaker indicates that he is sure of the occurrence of an event or the existence of a certain condition or state in the past” (2003: 667).

A number of other forms are presented as being opposed to the simple past. The *-gan-* form is perhaps the most basic of these. The morpheme *-gan-* has a number of functions in Uzbek; it may be used to form attributed participial phrases, to nominalize clauses for the purposes of complementation, and as a marker of past tense:

(25) **Past Participle**

kel-*gan* odam
 come-PTCP man
 “a man who came”

(26) **Nominalizer**

Yaralan-ib yiqil-*gan*-i-ni aniq bil-a-di.
 be.injured-CVB fall-NMLZ-3-ACC sure know-PRES-3SG
 “She knows for sure that he fell and got hurt.”
 (Straughn 2008)

(27) **Past Tense**

yozi-*gan*-man
 write-PRF-1sg
 “I have written.”

Due to its origin as a participle, the past tense in *-gan-* takes person markers which are different from those of the past tense, but equivalent to those used when nouns or adjectives are used predicatively:

Table 7 - The Past Tense in *-gan*

<i>qil-</i> ‘to do’	SG	PL
1	qil-gan-man	qil-gan-miz
2	qil-gan-san	qil-gan-siz
3	qil-gan	qil-gan-(lar)

The past tense in *-gan-* appears to function much like the non-confirmative forms in Turkish and Macedonian. Bodrogligeti and Kononov ascribe to meanings to the *-gan-* past: a perfect reading and something like non-confirmativity or reportativity. Coşkun, continuing in the Turkish tradition, calls the past in *-gan-* one of the “heard” past tenses.

A number of forms related to the *-gan-* form appear to participate in the system of evidentiality and/or status. The first of these is one of the negative forms. As is the case for the simple past, the verb can be negated by inserting the morpheme *-ma-* after the verb root:

- (28) yoz-ma-gan-man
 “I have (apparently) not written.”

However, there exists a second option for negation, in which the verb takes the possessive personal endings and is followed by the negative existential *yoq*:

- (29) yoz-gan-im yoq
 write-PRF-1SG NEG.EXIST
 “I have not written/did not write.”

According to Coşkun, this form is not non-confirmative; he classifies it as a variant form of the simple past tense (2000: 135).

Another related form is modal past, which consists of the verb root plus *-gan-* plus *-dir-* plus a person marker:

- (30) yoz-gan-dir-man
 “I may have written.”

Coşkun describes this form as a variant of the “heard” past tense, but Kononov and Bodrogligeti both describe this as bearing some sort of epistemic modal meaning, that is, one of probability or subjective evaluation.

Another past tense form discussed in grammars is the *-ib-* form. This form is related to the converbial/gerundive suffix *-ib*. It takes the same endings as the *-gan-* past and predicative nouns and adjectives:

Table 8 - The Past Tense in *-ib*

<i>qil-</i> 'to do'	SG	PL
1	<i>qil-ib-man</i>	<i>qil-ib-miz</i>
2	<i>qil-ib-san</i>	<i>qil-ib-siz</i>
3	<i>qil-ib-di</i>	<i>qil-ib-di-(lar)</i>

Coşkun describes this form as a variant of the “heard” past tense. Bodrogligeti and Kononov call it the “subjective past tense”, and ascribe to it meanings of doubt or uncertainty, of admirativity, or of historical fact:

- (31) Boy, handalak pish-ib-di-ya!
 oh, musk.melon ripen-PST-3SG-P
 “Oh, the muskmelons appear to be ripe!”
 (Bodrogligeti 2003: 693)

A final form not discussed in the grammars is the form in *-mish*, which is cognate with Turkish *-miş*. This form appears to be some sort of historical past. However, due to its similarity with its Turkish cognate, it deserves further study:

- (32) Azaldan ming-minglab bayt-lar yoz-mish-lar.
 in.ancient.times by.the.thousands couplet-PL write-PST-3PL
 “In ancient times they wrote couplets by the thousands.”
 (<http://forum.ziyouz.com/index.php?topic=1738.220;wap2>)

3.2.1. Complex Forms

Some of the morphemes above may be combined with the copula *e-*. These forms may combine with forms of the verb based on gerunds and participles. The three forms of interest here are *edi*, *ekan* (*e+gan*), and *emish*.

The forms in *edi* are referred to as narrative (*hikâye*) forms in Turkish grammars.

These forms generally have anterior readings:

- (33) bil-ayotgan e-di-m
 know-PROG.PTCP COP-PST-1SG
 “I was knowing.”
 (Coşkun 2000: 156)

Although the past marker *-di* has a confirmative reading, it is possible to combine it with the forms unmarked for confirmativity, *-gan*, and *-ib*:

- (32) ayt-ib e-di-m
 say-GER COP-PST-1SG
 “I had (apparently?) said.”
 (Coşkun 2000: 161)
- (33) bil-gan e-di-m
 say-PTCP COP-PST-1SG
 “I had (apparently?) known.”
 (Coşkun 2000: 159)

Bodrogligeti and Kononov state that these forms merely bear a pluperfect reading.

The form *ekan* is referred to as the “hearsay” (*rivayet*) form in Turkish grammars. Bodrogligeti states that these forms have more of a modal meaning, indicating a lack of confirmation or presence of doubt:

- (34) U xursand e-kan.
 She happy COP-PTCP.3SG
 “She is (apparently) happy.”
 (Bodrogligeti 2003: 778)
- (35) Yoz-ar e-kan-man.
 Write-AOR COP-PTCP-1SG
 “I apparently wrote.”

As is the case for the form in *edim*, forms in *ekan* may combine with the forms unmarked for confirmativity, *-gan*, and *-ib*:

- (36) O’qi-gan e-kan-man
 Read-PTCP COP-PTCP-1SG
 “I have apparently read.”

It is unclear if these forms have the same strong reportative reading of the equivalent forms in Turkish and Balkan Slavic.

Forms in *emish* appear to bear a modal meaning expressing probability:

- (37) Yoz-gan e-mish-man.
 Write-PTCP COP-PTCP-1SG
 “I have probably written.”
 (Bodrogligeti 2003: 791)

It is unclear whether the non-confirmative sense of the *-gan-* ending in the example above interacts with the modal meaning of *emish*.

There are, of course, an enormous number of other forms in the Uzbek verbal paradigm, expressing the future mood, desiderativity, conditionality, etc., however, a complete description of the Uzbek verb is beyond the scope of this proposal. To the extent that these forms are relevant for the study of evidentiality, they will be considered in the dissertation.

3.3. The Kazakh Verbal System

The Kazakh verbal system expresses similar distinctions to those expressed in Uzbek.

The simple past tense is formed, as in Uzbek and Turkish, with the morpheme *-DI-*⁸:

Table 9- The Simple Past in Kazakh

<i>bar-</i> ‘to go’	SG	PL
1	bar-dy-m	bar-dy-q
2	bar-dy-η, bar-dy-ηyz	bar-dy-η-dar, bar-dy-ηyz-dar
3	bar-dy	bar-dy

As in Uzbek, negation is indicated by the morpheme *-MA-* inserted after the root of the verb:

- (38) bar-ma-dy-m
 go-NEG-PST-1SG
 “I did not go.”

⁸ Kazakh, unlike Uzbek, possesses a rather complex system of vowel harmony and consonant assimilation, which requires the use of capital letters to indicate archiphonemes.

Due to the apparent simplicity of these forms, little has been written about them.

Rysbaeva notes that this form, which she calls the “recent past tense” “indicates only actions completed recently, at the time of speech” (2000).

Kazakh, like Uzbek, possesses past tense forms in *-GAN-*:

Table 10 - The Past Tense in *-GAN*

<i>bar-</i> ‘to go’	SG	PL
1	bar-ğan-myn	bar-ğan-myz
2	bar-ğan-syñ, bar-ğan-syz	bar-ğan-syñ-dar, bar-ğan-syz-dar
3	bar-ğan	bar-ğan

There is disagreement over the exact meanings of these forms. Koç and Doğan consider them to be essentially the same as the basic past tense (i.e. *görülen geçmiş zaman*) (2004: 255). Rysbaeva (2000), however, refers to this paradigm as the distant past tense, stating that it is used to describe past events that the speaker did not witness. Serğaliev et al., however, contrast the *-DI* and *-GAN* forms in different terms. Both forms can, they state, be used to describe past events. However, events described by the *-GAN* form must be completed (i.e. perfect), and forms in *-DI* must be known for certain (1991: 65). For example:

- (39) Ol üj-ge ket-**ti**.
He home-DAT go-PST.3SG
“He went home (but may still be going.)”
“He went home (and I am certain of it.)”
- (40) Ol üj-ge ket-**ken**.
He home-DAT go-PTCP.3SG
“He has gone home (and is done doing so).”
“He went home (or so I am told).”

Another form worth examining is the past tense in *-Ip*, for which there are similarly contradictory descriptions:

Table 10 - The Past Tense in *-Ip*

<i>bar-</i> ‘to go’	SG	PL
1	bar-yp-pyn	bar-yp-pyz
2	bar-yp-syñ, bar-yp-syz	bar-yp-syñ-dar, bar-yp-syz-dar

3	bar-yp-ty	bar-yp-ty
---	-----------	-----------

Koç and Doğan (2004) call this form the “learned past tense” (*öğrenilen geçmiş zaman*) and equate it with Turkish *-miş*. That is, it is used to report non-firsthand information. Rysbaeva (2000) considers this form to be equivalent to the forms in *-GAN*, stating that it is used to report information learned from someone else. Serġaliev et al., however, indicate that this form is used to indicate events that occurred long ago or to indicate action that the speaker does not wish to take credit for or which were performed inadvertently, e.g.:

- (41) Kilt-ti tast-yp ket-**ip**-pin.
 Key-ACC leave.behind-CVB leave-PST-1SG
 “I left without my key (accidentally, because I forgot it).
 (Serġaliev et al. 1991: 65).

3.3.1. Complex Verbal Forms

The copula *e-* may combine with two forms that combine to create more complex forms of the past tense: *e-di* and *e-ken* (*e+GAN*). These forms may combine with verbs in gerund or participial forms or with nouns and adjectives acting as predicates. When the form *edi* is used, the reading is one of anteriority. However, it is unclear whether combination with *-Ib* or *-GAN*, which have otherwise non-confirmative or modal readings preserves that modal reading through the confirmative *-DI* morpheme. Koç and Doğan consider the *edi* to be a purely anterior form with no non-confirmative reading (which they refer to as the *hikâye* “narrative” form), but Rysbaeva considers it to be equivalent to forms without the *edi* attached:

- (42) Qyz e-di-m.
 Girl COP-PST-1SG
 “I was a girl.”
 (Koç and Doğan 2004: 285)

- (43) Žaz-ğan e-di-ŋ.
Write-PTCP COP-PST-2SG
“You have (apparently?) written.”
(Rysbaeva 2000: 105)
- (44) Kel-me-p e-di-k
Come-neg-pst cop-pst-1pl
“We have (apparently?) not come.”
(Rysbaeva 2000: 109)

The form in *eken* appears to have, at the very least, a non-confirmative reading, much like forms in Turkish *emiş*. It is unclear if the combination of a non-confirmative morpheme on the verb stem combined with *eken* produces the same effect of reportedness that is found in equivalent constructions in Turkish and Balkan Slavic (the so-called *rivayet* “hearsay” form):

- (45) Žigit e-ken-siŋ-der.
Boy COP-PTCP-2-PL
“You all are (apparently) boys.”
(Koç and Doğan 2004: 286)
- (46) Qara-ğan e-ken.
Look-PTCP COP-PTCP-3
“He has (reportedly?) looked.”
(Koç and Doğan 2004: 297)

Because so much of what must be learned about Kazakh and Uzbek is dependent upon matters of discourse, field work will be necessary.

4. Research Questions

The verbal systems of Kazakh and Uzbek appear to exhibit the confirmative/non-confirmative described for Turkish and Balkan Slavic. However, there are a greater number of forms participating in the paradigms of Kazakh and Uzbek than in Turkish. The semantics and distribution of these forms needs further investigation. Forms that deserve special attention are those which combine more than one marker of non-confirmativity (*-ib ekan* and *-gan ekan* in Uzbek and *-Ip eken* and *-GAn eken* in Kazakh).

Friedman has stated that these forms have strong reportative readings in Turkish, Balkan Slavic, and other languages of that region (1979: 345).

In order to assess the possible range of meanings of these forms, a number of strategies will be employed. In Uzbek, at least, there exists the possibility of subordinating finite verb forms under verbs expressing doubt or disbelief (c.f. Straughn 2008). Confirmative forms should not be possible in these contexts, whereas forms that are unmarked for confirmativity should be possible, e.g.:

(47) Ishonmayman ki o kelgan.
I don't believe that he came.

vs.

*?Ishonmayman ki o keldi.
I don't believe that he came.

In order to assess these possibilities for Kazakh, some investigation into complementation will be necessary.

Beyond investigating the confirmative/non-confirmative distinction, other aspects of verbal semantics will be investigated. Important categories to examine include tense, aspect, resultativity and taxis, and other meanings related to modality/status. Admirative and dubitative extensions of non-confirmative forms will also be examined. The first person of otherwise non-confirmative forms is a likely place to find variation. Situations in which speakers are unaware of their actions (sleep, coma, various altered states) need to be tested to ensure that non-confirmative meanings are possible in the first person. Aikhenvald (2004) notes that evidentiality in the first person may bear meanings of non-volition, lack of control, lack of responsibility, or distance. While Aikhenvald's definition of evidentiality is quite different from the confirmativity of Turkic and affiliated languages, these possible extensions should be studied.

Once the semantics of various verbal markers have been established, their use in various contexts will be investigated. The use of various verbal forms in the context of folktales, historical narratives, and news reporting deserves special attention. Tentative research indicates that state-run media in Uzbekistan prefers the confirmative form in reporting news, whereas other media tends to use forms unmarked for confirmativity.

Research will be carried out using three main methods. The first method will employ corpora, such as various genres of published texts and freely available data from Internet message boards and websites. A second method will be the use of surveys, in person and on-line. These surveys will consist of various questions requesting speaker judgments on pre-composed examples and requests to answer questions and provide descriptions of constructed situations. The final method will be to conduct in-person interviews with native speakers of Uzbek and Kazakh and also record spontaneous conversation. The collection of oral data will take place both in America and in Kazakhstan, and will likely be supplemented with phone and on-line interviews. Work to secure travel funding is ongoing.

5. A Tentative Outline

I. Introduction

In the introduction, background matter will be discussed, including definitions of evidentiality and some description of how evidentiality/confirmativity function in general. Background information on Uzbek and Kazakh will be given, including relevant genetic, areal, and typological information.

II. Evidentiality in Uzbek

A. The Uzbek Verbal System

In this section, the verbal system of Uzbek will be examined, perhaps using Aronson's description of Bulgarian (1967) and Friedman's

description of Macedonian (1977) and the subsequent developments therefrom as models.

B. Uzbek Evidentiality in Context

This section will discuss the uses of the verbal forms discussed in the previous section, focusing on their distribution in conversation, and various genres of spoken and written language.

III. Evidentiality in Kazakh

A. The Kazakh Verbal System

In this section, the verbal system of Kazakh will be examined, perhaps using Aronson's description of Bulgarian (1967) and Friedman's description of Macedonian (1977) and the subsequent developments therefrom as models.

B. Kazakh Evidentiality in Context

This section will discuss the uses of the verbal forms discussed in the previous section, focusing on their distribution in conversation, and various genres of spoken and written language.

IV. Kazakh and Uzbek in a Larger Context

In this section, the systems of evidentiality/confirmativity in other languages of the region will be discussed.

A. Kazakh and Uzbek Compared

In this section Kazakh and Uzbek will be compared with regard to their verbal systems and what distinctions are expressed: confirmative, non-confirmative, and reportative.

B. The Turkic Languages

This section will consist of a comparison of Uzbek and Kazakh with the other Turkic languages.

C. Evidentiality in Eurasia

Although the evidential systems of Eurasia are not well known, some generalizations may be able to be made. Areal and typological considerations will be discussed here.

V. Conclusion

The final portion of the dissertation will tie together all of the data and analysis of the individual chapters.

Works Cited

- Aikhenvald, Alexandra Y. 2003. Evidentiality in Typological Perspective. In Aikhenvald and Dixon, 1-31.
- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford UP.
- Aikhenvald, Alexandra Y. and R. M. W. Dixon, eds. 2003. *Studies in Evidentiality*. Philadelphia: John Benjamins.
- Al-Kāšgarī, Maḥmūd. 1982. *Compendium of the Turkic Dialects (Dīwān Luyāt at-Turk)*. Ed., Trans. Robert Dankoff with James Kelly. Cambridge: Harvard UP.
- Anderson, Lloyd B. 1986. Evidentials, Paths of Change, and Mental Maps: Typologically Regular Asymmetries. In Chafe and Nichols, 273-312.
- Aronson, Howard I. 1967. The Grammatical Categories of the Indicative in the Contemporary Bulgarian Literary Language. In *To Honor Roman Jakobson*, 82-98. The Hague: Mouton.
- Aronson, Howard I. 1990. Towards a Typology of Verbal Categories. In *New Vistas in Grammar: Invariance and Variation*, eds. Linda R. Waugh and Stephen Rudy: 111-131.
- Boas, Franz. 1911. Kwakiutl. In *Handbook of American Indian Languages*, ed. Franz Boas, 423-557. Washington: Government Printing Office.
- Bodrogligeti, András J. E. 2003. *An Academic Reference Grammar of Modern Literary Uzbek*. Muenchen: Lincom.
- Chafe, Wallace, and Johanna Nichols, eds. 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-linguistic Perspective*. Oxford UP, New York.
- Comrie, Bernard. 2000. Evidentials: Semantics and History. In Johanson and Utas, 1-12.
- Conev, B. 1910/1911. Opredeleni i neopredeleni formi v bălgarski ezik. *Godišnik na Sofijskija universitet: Istorikofilologiĉeski fakultet* 7: 1-18.
- Coşkun, Volkan. 2000. *Özbek Türkçesi Grameri*. Ankara: TDK.
- De Haan, Ferdinand. 1999. Evidentiality and Epistemic Modality: Setting Boundaries. *Southwest Journal of Linguistics* 18: 83-101.
- DeLancey, Scott. 1997. Mirativity: The Grammatical Marking of Unexpected Information. *Linguistic Typology* 1: 33-52.
- Demir, Nurettin. 2003. On *imiş* in Cypriot Turkish. *Turkic Languages* 7, 268-274.
- Dwyer, Arienne. 2000. Direct and Indirect Experience in Salar. In Johanson and Utas.
- Friedman, Victor. 1977. *The Grammatical Categories of the Macedonian Indicative*. Columbus: Slavica.

- Friedman, Victor. 1978. On the Semantic and Morphological Influence of Turkish on Balkan Slavic. In *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*, eds. Donka Farkas et al., 108-118. Chicago: Chicago Linguistic Society.
- Friedman, Victor. 1979. Toward a Typology of Status: Georgian and Other Non-Slavic Languages of the Soviet Union. In *The Elements: A Parasession on Linguistic Units and Levels, Including Papers from the Conference on Non-Slavic Languages of the USSR*, eds. Paul R. Clyne, William F. Hanks, and Carol L. Hofbauer, 339-350. Chicago: Chicago Linguistic Society.
- Friedman, Victor. 1980. The Study of Balkan Admirativity: Its History and Development. *Balkanistica* 5: 7-30.
- Friedman, Victor. 1981. Admirativity and Confirmativity. *Zeitschrift für Balkanologie* 17.1: 12-28.
- Friedman, Victor. 1999. Evidentiality in the Balkan Languages. In *Handbuch der Südosteuropa-Linguistik*, ed. Uwe Hinrichs, 519-44. Wiesbaden: Harrassowitz.
- Friedman, Victor. 2000. Confirmative/Nonconfirmative in Balkan Slavic, Balkan Romance, and Albanian with Additional Observations on Turkish, Romani, Georgian, and Lak. In Johanson and Utas: 329-366.
- Friedman, Victor. 2002. Hunting the Elusive Evidential: the Third-Person Auxiliary as a Boojum in Bulgarian. *Of All the Slavs My Favorites: Studies in Honor of Howard I Aronson on the Occasion of his 66th Birthday*, eds. Victor Friedman and Don Dyer, 203-30. Bloomington, IN: Slavica.
- Friedman, Victor. 2003. Evidentiality in the Balkans with Special Attention to Macedonian and Albanian. In Aikhenvald and Dixon, 189-218.
- Guentcheva, Zlatka, ed. 1996. *L'Énonciation Médiatisée*. Paris: Peeters.
- Guentcheva, Zlatka, ed. 2007. *L'Énonciation Médiatisée II*. Paris: Peeters.
- Haarmann, Harald. 1970. *Die indirekte Erlebnisform als grammatische Kategorie. Eine eurasische Isoglosse*. Wiesbaden: Harrassowitz.
- Jacobsen, William H., Jr. 1986. *The Heterogeneity of Evidentials in Makah*. In Chafe and Nichols 1986: 3-28.
- Jakobson, Roman. 1957/1971. Shifters, Verbal Categories, and the Russian Verb. In *Roman Jakobson: Selected Writings II*, 130-147. The Hague: Mouton.
- Johanson, Lars. 2000. Turkic Indirectives. In Johanson and Utas: 61-87.
- Johanson, Lars. 2003. Evidentiality in Turkic. In Aikhenvald and Dixon: 273-290.
- Johanson, Lars and Bo Utas. 2000. *Evidentials: Turkic, Iranian and Neighboring Languages*. New York: Mouton de Gruyter.
- Kehayov, Petar. 2008. *An Areal-Typological Perspective to Evidentiality: the Cases of the Balkan and Baltic Linguistic Areas*. PhD Diss. University of Tartu. Tartu: Tartu UP.

- Khrakovskij, V. S., ed. 2007. *Ėvidentsial'nost' v jazykax Evropy i Azii*. Saint Petersburg: Nauka.
- Koç, Kenan, and Oğuz Doğan. 2004. *Kazak Türkçesi Grameri*. Ankara: Gazi Kitabevi.
- Kononov, A. N. 1960. *Grammatika Sovremennogo Uzbekskogo Literaturnogo Jazyka*. Moscow: Izdatel'stvo Akademii Nauk SSSR.
- Lazard, Gilbert. 1996. Le Médiatif en Persan. In Guentcheva 1996: 21-30.
- Lazard, Gilbert. 2000. Le Médiatif: Considérations Théoriques et Application à l'Iranien. In Joahson and Utas 2000: 210-228.
- Lunt, Horace. 1952. *A Grammar of the Macedonian Literary Language*. Skopje: Državno Knigoizdatelstvo.
- Matthewson, Lisa, Henry Davis and Hotze Rullmann. 2007. Evidentials as Epistemic Modals: Evidence from St'át'imcets. *Linguistic Variation Yearbook* 7: 201-254.
- Menz, Astrid. 2000. Indirectivity in Gagauz. In Johanson and Utas.
- Rysbaeva, G. K. 2000. Qazaq Tili: Grammatikalyq Anyqtaghysh/Kazaxskij Jazyk: Grammaticheskij Spravochnik. Almaty: Sözdik/Slovar'
- Serğaliev, Myrzataj, Ajtbaj Ajğabylov, and Oral Külkenova. 1991. *Qazirgi Qazaq Ādebi Tili*. Almaty: Ana Tili.
- Skribnik, E. K. and A. A. Osnova. 2007. Sredstva vyraženiya zasvidetel'stvovannosti i mirativnosti v altajskom jazyke. In Khrakovskij: 519-551.
- Straughn, Christopher. 2008. History vs. Theory: The Grammaticalization of the Uzbek *say*-Complementizers. Paper presented at the 39th Poznań Linguistic Meeting, Gniezno, Poland, 14 September 2008.
<http://home.uchicago.edu/~straughn/PLM.pdf>
- Tatevosov, Sergei. 2007. Evidentiality and Mirativity in the Mishar Dialect of Tatar. In Guentcheva: 377-406.
- Tekin, Talat. 2005. A New Classification of the Turkic Languages. In *Makaleler III: Çağdaş Türk Dilleri*, ed. Nurettin Demir and Emine Yılmaz. Ankara: Grafiker: 405-420.