Turkic indirectivity

Lars Johanson

1 Introduction

This survey summarizes the essential features of the grammatical categories of evidentiality found in Turkic languages, with special regard to the distinctive devices of particular system types. It follows the principles laid down in Aikhenvald (2004 and 2015).

Turkic evidential categories state the existence of a source of evidence for a given propositional content. The specific kind of evidentiality typical of Turkic is indirectivity, translatable as ‘obviously’, ‘evidently’, ‘apparently’, ‘reportedly’, ‘as it appears/appeared’, ‘as it turns/turned out’, etc.

Turkic evidential categories do not express epistemic modality in the sense of the addressee’s attitude towards the truth of this content. Nonetheless, brief sections on epistemic and rhetorical stance markers will conclude this survey.

2 Turkic

Since the survey concerns evidentiality as expressed in a whole language family, numerous language-specific details will be omitted. The reader may have to be reminded of the fact that today’s Turkic-speaking world extends from Turkey and its neighbours in the Southwest, to Eastern Turkistan and further into China in the Southeast. From here it stretches to the Northeast, via Southern and Northern Siberia up to the Arctic Ocean, and finally to the Northwest, across Western Siberia and Eastern Europe. Most Turkic languages may be classified as belonging to a Southwestern, a Northwestern, a Southeastern or a Northeastern branch. Khalaj in
Central Iran and Chuvash in the Volga region constitute separate branches. Of the peripheral languages in the Southeast, Yellow Uyghur and Fu-yü are related to dialects of the Northeastern branch, and Salar to the Southwestern branch.

3 Indirectivity

Despite system differences, almost all known older and recent stages of Turkic possess grammatical means of expressing indirectivity, covering various notions traditionally referred to as ‘hearsay’, ‘inferential’, etc. Evidential statements are indirect in the sense that the narrated event is not stated directly, but in an indirect way, by reference to its reception by a conscious subject, a recipient. This seems to be basic to many evidentiality systems and may even qualify as a partial crosslinguistic definition of evidentiality (Comrie 2000: 1). The recipient may be the speaker as a participant of the speech event or a participant of the narrated event, e.g. a protagonist in a narrative. The result is two-layered information: ‘It is stated that narrated event is acknowledged by a recipient’.

3.1 Sources of information

Specification of the source of information, the way in which the event is acknowledged by a recipient, is not criterial for indirectivity as such. The reception may be realised through (i) hearsay, (ii) inference, or (iii) perception.

(i) Reportive (or quotative) uses: ‘The narrated event or its effect is reported to the recipient’. The basis of knowledge is a foreign source, reported speech, hearsay. English translation equivalents include reportedly, allegedly, as they say/said, etc.
(ii) Inferential uses: ‘The narrated event or its effect is inferred by the recipient’. The basis of knowledge is pure reflection, logical conclusion. English translation equivalents include as far as one understands/understood, etc.

(iii) Perceptive (or experiential) uses: ‘The narrated event or its effect is perceived by the recipient’. The basis is first-hand knowledge, direct sensory perception of the event or indirect perception on the basis of traces or results. English translation equivalents include it appears/appeared that, it turns/turned out that, as one can/could see, hear, etc.

All these readings can be translated by evidently, obviously, etc.

Indirectivity markers do not fit into evidential schemes distinguishing between ‘the speaker’s non-first-hand and first-hand information’. Their primary task is not to express the external origin of the addresser’s knowledge.

In their perceptive uses, indirectives express that the event or its effect turns out to be the case, becomes manifest, visible, or apprehended through one of the senses and thus open to the recipient’s mind. Note that these usages cannot be derived from reportive or inferential meanings or be subsumed under ‘non-first-hand knowledge’.

Some more elaborate Turkic systems distinguish between ‘reported’ and ‘non-reported’ (inferential/perceptive) indirectivity. There are, however, no systematic differences relating to other types of sources, e.g. visual versus other kinds of sensory information.

3.2 Marked and unmarked terms

Turkic displays basic contrasts between marked indirectives and their unmarked counterparts. Marked terms, expressing evidential notions explicitly, stand in paradigmatic contrast to non-evidentials. Thus, Turkish Gel-miş ‘X has obviously
come/obviously came’ and Gel-iyor-muş ‘X is/was obviously coming, obviously comes’ have corresponding unmarked items such as Gel-di ‘X has come/came’ and Gel-iyor ‘X is coming/comes’.

The unmarked terms exhibit neutral uses in cases where the distinction in question is inessential. The widespread claim that unmarked items such as Gel-di ‘X has come/came’ consistently signal ‘direct experience’ or ‘visual evidence’ is clearly fallacious. Clauses unmarked for evidentiality do not necessarily denote situations that are personally known to the addresser. They simply do not signal that the event is stated in an indirect way, i.e. acknowledged by a recipient by means of report, inference or perception.

### 3.3 Formal types of markers

The coding of indirectivity in Turkic is scattered, i.e. morphologically realised by two types of markers. One type consists of postterminals that tend to vacillate between evidential and non-evidential readings. The other type consists of copular particles that are stable markers of evidentiality.

- Inflectional markers are suffixes occurring after verbal stems, comprising the types MIŞ, GAN and IB-DIR. The Turkish simple inflectional marker { -miş } carries high pitch and has mostly past time reference, e.g. Gül-müş (laugh-miş) ‘X (has) evidently laughed’.

- Copula particles are enclitic elements added to nominals, the main types being AR-MIŞ and AR-KAN. Turkish i-miş has the suffixed variant { -(y)-miş }. The copula particles are unable to carry high pitch and are ambiguous between past and non-past time reference, e.g. Turkish Hasta-y-miş ‘X is/was evidently sick’, Türkiye’de-y-
miş 'X is/was obviously in Turkey', Gel-iyor-muş 'X is/was evidently arriving', Gel-ecek-miş 'X will/would evidently arrive'.

Some written shapes of the Turkish copula particle coincide with those of the inflectional marker {-miş}. Thus Gül-muş (laugh-miş) is written in the same way as Gül-muş (rose i-miş) 'It is/was evidently a rose'. In spoken language, the allomorphs are distinguished by different pitch patterns. The deceptive similarity of certain allomorphs has led linguists to confuse the two markers, referring to both as “the suffix -miş”, allegedly attachable to both verbal and nominal stems. Uzbek e-kân is frequently cliticized as -kân, sometimes also written as a bound element.

4 Inflectional markers as postterminals

The inflectional markers are of postterminal nature. Postterminality is a marked aspectual way of envisaging events with respect to their limits, grammaticalised in Turkic as well as in many other languages (Johanson 1996a, 2000). It is typical of perfects in British English or Scandinavian languages, expressing past events of present relevance.

It is possible to distinguish degrees of focality depending on the focus of attention. High-focal postterminals focus on the aspectual orientation point and the relevance of the event at this point, whereas low-focal postterminals are more event-oriented, stressing the relevance of the event at the time of its realization (Johanson 2000: 106-136).

High-focal postterminals often tend towards indirective readings (Johanson 1971: Chapter 8, 2000: 121-3). Even if the event is wholly or partly outside the range of vision, traces, results or other forms of present knowledge of it may obtain at the aspectual vantage point. These secondary meanings are pragmatic side effects that
can be used as ‘evidential strategies’. The development of more stable indirective meanings may be seen as a semantic extension in the sense of conventionalised implicatures. Their indirect kind of envisaging events has been reinterpreted as indirectivity.

The oldest known postterminal marker type is MIŠ, and the second one is GAN. In certain languages, the expression of focal postterminality has later been renewed by means of IB-DIR and some other markers.

4.1 The type MIŠ

East Old Turkic {\text{-miš}} has clearly indirective functions, mainly expressing past actions known from hearsay, e.g. Ölür-miš ‘X reportedly killed’. The marker {\text{-mlî}} is still used in West Oghuz, South Oghuz, Khorasan Oghuz, Khalaj, Salar and North Siberian Turkic. It has strong evidential connotations, covering hearsay, inference, surprise, etc. (‘reportedly’, ‘obviously’, ‘surprisingly’), e.g. Turkish Gel-miš ‘X obviously came/has obviously come’, İç-miš-im ‘I obviously drank/have obviously drunk’, Gagauz Gör-muş-ük ‘We obviously saw it/have obviously seen it’. The markers {\text{-mlî}} and {\text{-DI}} exclude each other.

Under the influence of the Persian present perfect, Azeri {\text{-mlî}} exhibits more perfect-like functions without evidential connotations, e.g. Gâl-miš-âm ‘I have come’, corresponding to Turkish ‘Gel-di-m’ rather than to Gel-miš-im (Johanson 1971, 289-90). The same is true of many {\text{-mlî}} forms in Old Anatolian Turkish and Old Ottoman. Standard Azeri has a mixed perfect paradigm, with {\text{-mlî}} in the first person and {\text{-y}lîb} in the second and third persons. This paradigm is also found in other dialects of Iran (Johanson 1998). Certain South Oghuz varieties have a perfect with {\text{-miš}} in all persons, whereas some other varieties use {\text{-ylîb}} for all persons. Khalaj has a {\text{-mlî}} perfect for
all persons, e.g. Kāl-mūš-ām ~ Kāl-mīš-ām 'I have come', corresponding to the Persian present perfect *Man aː mad am*.

The Yakut postterminal marker {-Blt} is an archaic feature, etymologically corresponding to {-mlš}. It forms postterminals with evidential connotations, e.g. Kāl-bit 'X has obviously come', negated {-BA-tAG}, e.g. Kāl-bā-tāx-čīt 'You (Pl) have obviously not come' (Buder 1989).

4.2 The type GAN

The type GAN, which is lacking in East Old Turkic, replaced Miš in East Middle Turkic (Chaghatay). It is now used in the Northwestern and Southeastern branches as well as in the West and South Siberian languages. It corresponds phonetically to the Turkmen participant nominal marker {-An}//{-n}, e.g. okoː-n ‘having read’ ← oko- oka- ‘to read’. The {-GAN} perfects of Tatar and Bashkir have evidential connotations, e.g. Bashkir Al-yan-dar 'They have obviously taken it', Tatar Miːn al-yan-man 'It turns/turned out that I have/had taken it'.

West and South Kipchak {-GAn} expresses past events of current relevance, often on the basis of results or indirect evidence, e.g. Karachay-Balkar Ayt-yān-sa 'You have said', Jāz-yān-ma 'I have written', Kumyk Bar-yān-man 'I have gone', Bar-ma-yān-man 'I have not gone', Crimean Tatar Al-yān-mīz 'We have taken it', Kirghiz ḏal-yān 'X has stayed', Kazakh Meːn kōr-gen-min ~ kōr-ge-m 'I have seen it', Meːn oḵi-yān-mīn 'I have read', Meːn bul kītāp-tī oḵi-yān-mīn 'I have read this book', Meːn özger-ge-m 'I have changed'. It has perfect, resultative, experiential and constative (summarizing) functions.

Uzbek {-Gān} and Uyghur {-GAN}, so-called "indefinite past" markers, form a present perfect, presenting the event in a postterminal perspective and signaling its current relevance, sometimes with slight evidential connotations, e.g. Uzbek Kel-gān-mān 'I have
come', Yâz-yân 'X has written', Uyghur Kir-gân-mân 'I have entered', Yâz-il-yan 'It is written', Kir-mi-gân-siz 'You have not entered', Bu kitap-ni mân oku-yan 'I have once read this book'.

The Chuvash so-called perfect in {-nỊ} is an indirective postterminal lacking person-number markers, e.g. Äbi vula-nị (Эпӗ вулӑнӑ) 'I have read'. It is traditionally described as a non-eyewitness form found in narrative styles, especially of folktales, e.g. Kur-nị (Кур-нӑ) 'X has obviously seen it'. Its indirective meaning may be corroborated by the particle mîn мӗн, e.g. Pil-nị mîn (Пӗл-нӗ мӗн) 'X has obviously known it'. The Upper Chuvash counterpart is {-sA}.

4.3 Origins

The types MIŞ and GAN are of unknown origin. They may, however, have emerged in postverbal constructions with auxiliaries developed from lexical verbs, with deletion of the original converb suffix. Thus MIŞ may go back to a form of an original verb bîş- 'to ripen', 'to mature', i.e. to attain a final state as 'ripe', 'cooked' or 'done'; cf. Turkish pis- (Johanson 2003: 287). The type GAN may go back to a postverbal construction with an auxiliary verb developed from the lexical source koːn- 'to be satisfied, satiated, repleted', 'to do/be well (sufficiently) done'. The origin of the Chuvash marker {-nỊ} is unknown. Upper Chuvash {-sA} is connected with the hypothetical marker {-sA}.

4.4 The type IB-DIR

Many evidentials are based on the type IB-DIR. It goes back to the periphrasis *(B) tur-ur, which served to renew the expression of postterminality. It originally consisted of a converb of the lexical verb plus tur-ur 'stands', e.g. Yaz-ib tur-ur (lit. 'stands having written') 'X is in the state of having written', 'X has written'. The auxiliary tur-ur was
reduced to {\-{dU}r}, {\-{dI}} or Ø. This type is predominantly an indirective past, often of inferential and perceptive nature ('as I understand', 'as I observe'). In the traditional grammatical literature it is often mistaken for a pluperfect.

This type includes Noghay Yaz-ip-ti 'X evidently wrote/has written', Kazakh Kel-ip-ti 'X evidently came/has evidently come', Sen özger-me-p-sîn 'You have (as I see) not changed', Men bar-ip-pin 'It turned out that I had gone. Ol kel-ip-ti 'It turned out that X had come'. Men onî kör-ip-pin 'It turned out that I had seen him', Kirghiz Ber-be-p-tîr 'X has evidently not given it it', Uzbek Kel-mâ-p-ti 'X has evidently not arrived', Unut-ip-mân 'I have (as it turns out) forgotten it', Ā-p-ti < Āl-ip-ti 'X has evidently taken it', Uyghur Yez-ip-tu 'X evidently wrote/appears to have written', Tama ꞑoχša-p-tu 'The food is (as I taste) delicious''; cf. Turkish ‹Yemek güzel olmuș›, Altay Bar-ip-tîr 'X has evidently left', Salar Gel-du 'X evidently came', Tuvan Bär-ip-tîr 'X evidently gave', Tuvan De-p-tîr 'X has evidently said it', Khakas Uzu-p-tîr 'X has obviously slept', Par-tîr 'X has obviously gone'. Azeri {\-{y)Ib} < *(b) tur-ur forms a mixed perfect paradigm together with {\-{mIš}}, e.g. Yaz-mîš-am 'I have written', Gâl-ib-sân 'You have come', Gâl-ib ~ Gâl-ib-dir 'X has come', Bil-mâ-yib-lâr ~ Bil-mâ-yib-dir-lâr 'They have not known it'.

A few languages have produced a second renewal of focal postterminality by means of the periphrasis (b) converb + present tense of tur-, e.g. Karachay Ket-ib tur-a-di 'X has gone', Kumyk Gel-ip tur-a 'X has come'. These markers do not convey evidential connotations.

**4.5 The types ÄR-MIŠ and ÄR-KÂN**

The particles ÄR-MIŠ and ÄR-KÂN are derived from the defective verb är- ‘to be’. Both may be of postterminal origin, if är- was originally an initiotransformative expressing (i) an initial dynamic phase ‘to become’ and (ii) a subsequent stative phase ‘to be’ (Johanson
2000: 62-3). The postterminal perspective thus envisages the event as still going on at
the aspectual vantage point, e.g. är-miş ‘has appeared’, ‘has become evident’, ‘is evident’.
The particles är-miş and är-kân have now lost their relationship to the postterminal
value and cannot be considered perfect markers.

4.6 The type är-miş

The type är-miş is documented in East Old Turkic, where it takes part in various
analytic constructions, e.g. with the aorist, the optative and the prospective. In later
languages it combines with postterminal bases and other nominals, e.g. Kel-gân är-
miş ‘X has reportedly arrived’, Chaghatay Baha:dur e-miş siz ‘You are said to be a
hero’. It often suggests second-hand information in the reportive sense. The Yakut
equivalent of is ä-bit, combinable with various thematic bases. The Turkish marker is
i-miş ~ *(y)mIş, e.g. Zengin-miş ‘X is/was evidently rich’, Čık-iyor-muş ‘X is/was
obviously leaving’, Gel-miş-miş ‘X is said to have come’ (Johanson 1971: 66), Gel-ecek-
miş ‘X will/shall evidently come’, Gel-meli-ymiş ‘X evidently ought to come’. Gagauz
examples are Gid-är-miş-im ‘They say I will go’, Kâl-miş-miş ‘X has evidently
evidently must wait’. The Turkmen particle {*mIş}, which mostly expresses reportive
indirectivity, combines with numerous thematic bases, e.g. Tap-an-miş (Tapanmys) ‘X
is said to have found it’, Gel-ip-miş-in (Gelipmişin) ‘X has reportedly come’,
representing reported past events. Khalaj ä(r)-miş ~ {*A(r)-miş} has non-evidential
perfect and pluperfect functions (‘has/had been’) as a result of Persian influence. It
combines with intraterminal markers, signaling that an intraterminal situation has
been the case, e.g. Ā:t-ayor-amiş, interpretable as ‘It has been the case that X was
doing’; cf. Persian Mi:-karda-ast. Combined with {-miş}, it forms a pluperfect signaling
that a postterminal situation has been the case, e.g. Ā:t-miš ā-miš, interpretable as ‘It has been the case that X had done’; cf. Persian Karda bu:da ast.

Some languages have just preserved remnants of ĀR-MIŠ. For instance, Kazakh possesses the rare form \{-l}p-tl-mls\} < *v tur-ár-miš, which expresses rumours or gossip with mocking overtones, e.g. Ol ayt-ip-tî-mîs ‘X has reportedly said it’; cf. Turkish \{-mlîs-mlî\}.

4.7 The type ĀR-Kân

Many older and more recent Turkic languages display indirective particles of the type ĀR-Kân. The functional development is somewhat unclear, since ā(r)-kän is not a phonetically regular postterminal form in \{-GAn\} (Johanson 1996b: 91). The particles tend to convey the meaning ‘as is/was obvious’ or ‘as it turns/turned out’. Of the older languages, Kuman exhibits the form ā-gän. Modern phonetic variants include Tatar i-kän (икэн), Kazakh e-ken, Uzbek e-kän, Uyghur i-kän, Tofan ār-gän, negated ā-män < ār-mä-gän. Turkmen e-ken tends to express evidentiality in the perceptive sense, such as ‘It turns out that ...’, ‘I recognize/see/understand that ...’, e.g. Muğallîm eken-ðîn ‘I understand you are a teacher’, Gel-en e-ken ‘X has obviously arrived’.

Examples of combinations: Noghay Kele-yat-îr e-ken ‘X is apparently coming’, Kirghiz Ište-çü e-ken ‘X obviously used to work’, Kazakh Bil-e-di eken ‘X obviously knows/knew’, Ol žaman e-mes e-ken ‘X is/was obviously not bad’, Kel-üw-de e-ken ‘X is/was obviously coming’, Kel-gen e-ken-siz ‘You have (as I see) arrived’, Kel-mek-śi e-ken ‘X obviously intends/intended to come’, Kel-etin e-ken ‘X obviously used to come’, Žanbîr žaw-yan e-ken ‘It has (as I see) rained’ (cf. Turkish Yağmur yağ-miş, Uzbek Kásal ekän ‘X is obviously ill’, Yâz-gân e-kän ‘X has/had obviously written’, Bâr-mä-gän e-kän-sän ‘You have/had apparently not gone’.
A marker with functions similar to those of ÄR-KÄN is BOL-IB-DIR, e.g. Noghay bol-ip-tī, Uzbek bol-ip, Uyghur bo-p-ti, Altay bol-up-tīr, bol-tīr, Khakas pol-tīr, Kazakh Kal-yan bol-ip-tī ‘X has/had obviously stayed’. Another marker is BOL-GAN, e.g. Tatar Bar-a bul-yan ‘X is/was evidently going’, Bar-yan bul-yan ‘X has/had evidently gone’, Bar-ačak bul-yan ‘X will/would evidently go’. The verb (bj)ol- is used here in the sense of ‘to turn out to be’.

5 Types of systems

5.1 System type 1

The most comprehensive evidentiality systems are represented by languages such as Uyghur and Uzbek of the Southeastern branch, Kazakh of the Northwestern branch and Turkmen of the Southwestern branch.

They possess an inflectional past in IB-DIR, a stable indirectivity marker, e.g. Uyghur Yez-ip-tu, Uzbek Yâz-ib-di ‘X has evidently written/evidently wrote’, Kazakh Tūs-ip-tī ‘X has evidently fallen/evidently fell’, Turkmen Gida-ip-dir ‘X has evidently gone’. They possess a postterminal in GAN, displaying perfect-like meanings with occasional indirective connotations, e.g. Uyghur Yaz-yan, Uzbek Yâz-yan ‘X has written’, Kazakh Öltir-gen ‘X has killed’, Turkmen Öylön-ön ‘X has married/is married’.

Languages of this type possess two indirective copula particles, ÄR-KÄN, which tends towards non-reportive (inferential and perceptive) uses, and ÄR-MIŠ, which tends towards reportive uses, e.g. Tatar i-kăn (икэн) vs. i-miš (имиш), Chuvash i-kăn (иккен) vs. i-miš (имеш), Uzbek e-kăn vs. e-miš, Uyghur i-kăn vs. i-miš.

ÄR-KÄN combines with intraterinals (presents, imperfects), prospectives, non-verbal predicates, etc., e.g. Uyghur Yez-ivat-kan i-kăn ‘X is/was evidently writing’, Kazakh Kel-edi e-ken ‘X is/was evidently coming’, Üy-de e-ken ‘X is/was obviously at
home’. It combines with postterminals, e.g. Uyghur Tügät-kän i-kän ‘X has/had obviously finished’, Uzbek Yāz-yan e-ken ‘X has/had obviously written’, Kazakh Tūs-ken e-ken ‘X has/had obviously fallen’.

ÄR-MIŠ expresses corresponding reportive meanings, e.g. Uyghur Yez-ivat-ḳan-miš ‘X is/was reportedly writing’, Kazakh Kel-e-di-mis ‘X is/was reportedly arriving’, Uyghur Yaz-yan-miš ‘X has/had reportedly written’, Turkmen Gid-ip-miš-in ‘X has/had reportedly gone’.

Items of the structure IB-DIR + ÄR-MIŠ apply reportive meaning to inferential or perceptive statements, e.g. Uyghur Yez-ip-ti-miš ‘X has/had allegedly written’, Kazakh Kel-ip-ti-mis ‘X has/had allegedly come’.

In certain systems, the two copula particles divide the area of indirectivity between themselves according to the pattern reportive versus non-reportive (inferential + perceptive).

The opposition is sometimes limited to certain dialects or registers. Thus, ÄR-MIŠ is not used in all varieties of Uyghur and Uzbek, and its role in Kazakh is rather limited.

5.2 System type 2

Some languages such as Noghay, of the Northwestern branch, exhibit two inflectional markers, e.g. Kel-ip-ti ‘X evidently arrived’ and Kel-gen ‘X has arrived’, but only one indirective copula particle, ÄR-KĀN. The latter is a general indirective marker covering both reportive and non-reportive meanings. It combines with intraterminals, e.g. Kel-e-di e-ken ‘X is/was obviously coming’, and with postterminals to form indirectives signalling relative anteriority, e.g. Kel-gen e-ken, Kel-ip-ti e-ken ‘X has/had obviously come’.
5.3 System type 3

Certain languages exhibit a simplified subsystem of inflectional markers, while maintaining a richer subsystem of copula particles, distinguishing between reportive and non-reportive. In Tatar and Bashkir, of the Northwestern branch, \(G\)AN is used without a competing \(IB-DIR\). It displays normal postterminal uses but may also suggest indirectivity, e.g. \(Yaz\)-\(yan\) 'X has (evidently) written'. As noted above, the neighbouring language Chuvash has a similar marker \(-n\I\) with postterminal and indirective meanings, e.g. \(K\)ala-\(n\j\) (Каланӑ) 'X has (evidently) spoken'. Tatar, Bashkir and Chuvash possess indirective copula particles of the \(ÄR-MIȘ\) (reportive) and the \(ÄR-KÄN\) (non-reportive) type, e.g. Chuvash \(K\)il-\(n\j\) \(i-mäš\) (Килнӗ имеш) 'X has reportedly arrived', \(K\)il-\(n\j\) \(i-kän\) (Кил-нӗ иккен) 'X has evidently arrived'.

5.4 System type 4

A few systems consist of one inflectional marker and one copula particle. An inflectional marker of the type \(MIȘ\) is used in the western subgroup of the Southwestern branch, e.g. Turkish \(-mIš\). The cognate item \(-BIt\) is used in Yakut, the northernmost Turkic language of the Northeastern branch, spoken in the opposite extreme part of the Turkic world.

The languages in question possess particles of the type \(ÄR-MIȘ\), e.g. Turkish \(i-miș\), Yakut \(ä-bit\). Thus \(MIȘ\) lacks a competing \(IB-DIR\), and \(ÄR-MIȘ\) lacks a competing \(ÄR-KÄN\). The inflectional markers allow reportive, inferential and perceptive readings, thus corresponding to several items in more comprehensive systems. A Turkish complex item \(MIȘ + ÄR-MIȘ\) applies an explicitly indirective type of evidentiality to a postterminally envisaged event and is often used for rumours and gossip, e.g. \(Gel\)-\(miș\)-\(miș\) 'X has/had reportedly arrived'. 
The Yakut inflectional marker {-Blt} conveys reportive, inferential and perceptive nuances, e.g. *Kel-bit* 'X has (obviously) arrived’. The temporally indifferent indirective particle à-bit allows combinations with intraterminals and postterminals, e.g. *Tur-ar à-bit* 'X evidently stands/stood', *Kel-bit à-bit* 'X has/had evidently arrived’.

### 5.5 Smaller systems

There are still smaller evidentiality systems. The status of the Azeri inflectional marker {-miš}, which forms a mixed paradigm with {-(I)b}, differs considerably from that of Turkish {-miš}. It represents a type with mainly postterminal, non-evidential perfect meanings, e.g. *Gäl-miš-äm* 'I have arrived', *Yaz-ib-sin* 'You have written’. It is a postterminal with occasional secondary indirective readings. The unmarked term {-DI} thus tends towards preterite functions, e.g. *Gäl-di* 'X came' versus *Gäl-ib* 'X has come’. However, Azeri possesses, like Turkish, an indirective copula particle of the type *AR-MIŠ*, namely *i-miš*. The combination *miš + AR-MIŠ* thus unambiguously applies indirectivity to postterminally envisaged events, e.g. *Yaz-miš-miš* 'X has/had reportedly written’.

### 6 Contextual interpretations and semantic extensions

The motives for using Turkic indirectives may vary. They may get various contextual interpretations and display various pragmatic extensions of their central meaning.

Indirectives may evoke the impression that the recipient does not/did not witness the event or participate in it consciously, not being in control of it or directly involved in it. However, despite the indirect way of presentation, these meanings are not signaled explicitly.

The recipient may apprehend the event through the senses or take part in it consciously. Lack of participation or control is limited to certain contexts and cannot
be the common core meaning. The source of information may be direct evidence, personal, even visually obtained knowledge. Uyghur Äxmat kā-p-tu 'Ahmed has (as I note) arrived' can also be uttered by somebody who has witnessed the arrival. The indirective statement just expresses the conscious reception. It does not tell us how something is in reality, but rather how the addressee chooses to present it.

Evidentially unmarked terms may suggest that the source of information is direct experience, but they may also be used for unwitnessed events, e.g. Turkish Büyüdün 'You have grown'. They just lack the two-layered information typical of indirectives, and may be used whenever this information seems unessential.

Turkic indirectives may have epistemic connotations in the sense of reservations about the validity of the event as a fact. The indirect way of referring may create uncertainty concerning the realisation of the event. Indirectives can be used to disclaim direct responsibility for the truth of the statement, suggesting that the addressee does not vouch for the information. By contrast, unmarked terms may suggest that the addressee is certain of the truth of the information and responsible for it. However, indirectives are not presumptives or dubitatives reducing the factuality of the statement.

As a pragmatic extension of their central value, indirectives may suggest a certain dissociation from the narrated event, i.e. a cognitive or emotional distance to it. Some kind of distance is certainly involved if the addressee does not refer directly to the event, but rather to its reception. Thus miş and its counterparts, e.g. in Old East Turkic, have been referred to as 'preterites of distance'. One kind of dissociation from the event may be an ironic relation to it, a reservation interpretable as sarcasm or disdain. An indirective statement may be motivated by caution, modesty, need for a summarising view, etc., e.g. Turkish Ben her zaman vazife-m-i yap-miş-im, Uyghur
Män daim væzipä-m-ni ada ḳi-pto-män 'I have (as it appears) always done my duty',

Turkish Önemli bir konu el-e al-miş-sin 'You have (if I may summarise) addressed an important topic'. Readings of these kinds derive from the indirect postterminal perspective.

Indirectives of the types IB-DIR, Miš, ÄR-Kän and ÄR-Miš may, in particular contexts, convey mirative connotations, i.e. be interpretable in terms of new knowledge, discovery, sudden awareness of revealed facts, surprise, mental unpreparedness, perception contrary to one's expectations, admiration, etc. Such readings naturally follow from the notion of indirectivity; what the recipient turns the mind to may come as a surprise. The conscious reception may be sudden or unexpected. The statement that Turkish indirectives may convey new information that is not yet part of the speaker's integrated picture of the world (Aksu-Koç & Slobin 1986) is compatible with the central value of indirectivity. This does not mean that mirativity is their central meaning from which the other uses may be derived (DeLancey 1997). Surprise, novelty and contrariness to the speaker’s expectation are not necessary elements of indirectivity. On the contrary, so-called “hot news” is typically expressed by the direct preterite marker DI. The indirective marker just adds the meaning ‘as I am/become aware of’, e.g. Kazakh Ol ket-ip ḳal-ić-tï, Uyghur U ket-ip ḳa-p-tu ‘X has left (as I note)’, Turkish Bu kız ne güzel-miş! ‘How beautiful this girl is!’ Uyghur Bu kız čiraylık i-kän! ‘This girl is beautiful!’.

7 Differences between grammatical persons

Though evidential specifications are possible in all grammatical persons, certain interdependencies with the person systems may be observed. The semantic interpretations vary according to the degree of the recipient's involvement in the
event. There are often differences between the first person singular and other persons. Reportive or inferential uses are naturally most common with third persons.

The narrow definition of indirectivity as the expression of ‘the speaker’s non-first-hand information’ is obviously incorrect. The use of indirectives when speaking of oneself would then necessarily imply lack of awareness, consciousness or control due to inattention, sleep, drunkenness, coma, etc. However, a definition based on the presentation of the event ‘by reference to its reception by a conscious subject’, is by no means contradictory to the use of indirectives with first-person subjects.

In Yellow Uyghur, a small Turkic language spoken in Western China, the second and third persons of the past tense tend to take on the evidential marker \textit{ib-DIR}, whereas the first person takes on the non-evidential marker \textit{di} (Tenişev 1976: 92-3). Roos (2000: 105-6) suggests a unified past tense paradigm, in which first and non-first persons take on different suffixes, e.g. \textit{Män păr-tï} ‘I went’ (\textit{di}), \textit{Sän part-t-tï} ‘You went’ (\textit{ib-DIR}).

\textbf{8 Correlations with other grammatical categories}

Turkic indirectives are limited to main clauses with a contradictable content, e.g. Turkish \{-mIš\} in \textit{Git-miş} ‘X has apparently gone’. Other uses of \{-mIš\}, e.g. in the pluperfect marker \{-mIš-tI\}, cannot express evidentiality. In certain constructions governed by postpositions such as \textit{gibi} ‘like’, ‘as’, evidentials can, however, occur as non-finite forms, e.g. \textit{git-miş gibî} ‘as if ... having gone’.

Indirective copula particles do not combine with the preterite in \textit{di} and the related copula particles \textit{e-di}, \textit{i-di} ‘was’ etc. They are at variance with each other: it would be contradictory to combine indirective markers with items conveying a direct perspective.
Combinations with imperatives are excluded since they would indicate that a direct appeal is expressed in an indirect way, as based on some source. Evidentials may, however, co-occur with necessitatives or debitives, e.g. Turkish Git-meli-ymış-sin, Uyghur Sän ket-iş-iñ keräk i-k än 'You evidently ought to go'.

In negative sentences, indirectives are not within the scope of negation. The narrated event itself is negated, not its reception by a conscious subject, e.g. Uyghur U käł-mä-ptu 'X has not arrived (as I note)'.

Indirectives may occur in interrogative sentences, e.g. Turkish O böyle de-miş mi?, Uyghur U mundaç de-ptu-mu? ‘Did X reportedly say so?’, Kazakh Kel-e mi e-ken? ‘Is X, as it appears, coming?’, Üy-de mi e-ken? ‘Is X, as it appears, at home?’, Noghay Ne-ge kel-gen e-ken-ler? ‘Why have they, as it appears, come?’. Indirectives may also be used in questions asked on behalf of someone else than the addresser.

9 Indirectivity and discourse

Indirectives play various roles according to different discourse types. Both as genuine indirectives and as indirectly interpretable postterminals the markers MIŞ, GAN and IB-DIR often serve as propulsive (‘plot-advancing’) basic items in certain narrative styles. In traditional story-telling, e.g. in fairy tales and other folklore texts, indirectives tend to create a specific narrative key, e.g. Uyghur Burun bir padişä öt-üp-tu, un-iñ bir bali-si bar i-kän ‘Once there was a king, he had a child’. On traditional MIŞ-based narratives in Turkish see Johanson (1971: 79-80). For similar forms in Shor folk tales see Nevskaya (2002). On the other hand, indirectives are not typically used for recounting dreams or imaginary events in fictional texts.

10 Contact-induced code-copying
Indirectives play a central part in almost all Turkic languages. However, owing to influence from Indo-European languages such as Persian, Greek and Slavic, a few languages only exhibit evidentiality strategies. The tendency of Azeri {-mIš}/-{(I)b} towards pure perfect readings is a result of Persian influence, e.g. Yap-ib 'X has done', cf. Persian Kard-a ast (Johanson 1988: 249). Evidentiality systems are lacking in Karaim of Lithuania, under Slavic and Lithuanian impact (Csató 2000b), and in the Turkish dialects of the Trabzon province on the east Black Sea coast, under the impact of Greek (Brendemoen 1997).

Features of Turkic evidential systems have proven highly attractive in language contact situations and have been copied into non-Turkic languages of Southwestern and Central Asia, Southeastern and Northeastern Europe. Indirective categories similar to the Turkic ones typically appear in contact areas such as the Balkans, Anatolia, Caucasus, the Volga region and Central Asia, e.g. in Bulgarian, Macedonian, Albanian, Kurdish, Western Armenian, Georgian, Tajik and eastern Finno-Ugric. Northern Tajik has developed a comprehensive evidential system on the Uzbek model. Indirective functions have been copied onto postterminals of the perfect type and also onto related participles, on the model of the temporally indifferent AR-KÂN and AR-MIŠ, e.g. Western Armenian eyer and Bulgarian bil (Johanson 1996b). Hungarian igen 'yes' may go back to a Turkic form AR-KÂN 'evidently' (Johanson 2004).

Differences in markedness sometimes seem to speak against the assumption of contact influence. The basic evidential oppositions of Bulgarian and Macedonian are described as relying on marked ‘confirmative’ items indicating unequivocal and direct assertion, whereas the corresponding unmarked items convey indirective meanings in particular contexts. Have systems based on marked confirmatives emerged through areal contact with Turkic systems based on marked indirectives? Comrie considers the
possibility that the semantic distinction can be reduced to a single prototype with
markedness inversion: 'one of the systems, almost certainly the Balkan one, has
undergone a shift whereby an old indirective was reinterpreted as unmarked, with the
originally unmarked non-indirective then becoming a marked confirmative' (2000: 8).

11 Relations to modal categories

Evidential categories are sometimes difficult to distinguish from presumptive categories.
The value of the enclitic element $\text{DIR} < \text{tur}-ur$ frequently oscillates between affirmation
and presumption. For instance, Turkish Alanya güzel-dir may mean 'It is a fact that
Alanya is beautiful' or 'I assume that Alanya is beautiful'. The presumptive meaning of
Turkish $\{+\text{DIR}\}$ is typical of the informal spoken language, e.g. Zengin-dir 'I guess X is
rich', Um-ar-im iyi-siniz-dir 'I trust you are well', iç-miş-tir 'I guess X has drunk'
(Johanson 1971: 294).

The type $\text{DIR}$ is added to postterminals, intraterminals and other forms, e.g. Turkish
Uyu-yor-dur 'X is presumably sleeping', Turkmen Oka-n-nir (Oka-n-dyr) 'X must have
read it', Oko-ya:n-nir (Oka-yan-dyr) 'X is presumably reading it', Bar-an-nir (Barandyr) 'X
has presumably gone', Düş-en-nir (Duş-en-dir) 'X must have fallen', Bashkir Kil-ä-lîr 'X is
presumably coming', Xat-îm-dî al-yan-hîn-dîr 'You have probably received my letter',
Chuvash Pîl-män-dîr (Пëлмëнтëр) 'X probably does not know', Kirghiz Oyyon-yon-dur 'X
has presumably waken up', Uyghur Kir-i-di-yan-di-men 'I am supposed to enter', Išlâ-vat-
kan-du 'X is presumably working', Uzbek Ket-gân-dir 'X has presumably gone'; cf.
Northern Tajik Räftâgîst.

The East Old Turkic inscriptions display an epistemic particle är-inç, an uninflated
utterance-final presumptive marker. It may follow preterite forms, which is impossible
with evidentials, e.g. Karakhanid Ol kel-di ärinç 'X presumably came/has presumably
come'. Its counterpart in Old Uyghur and Karakhanid is ār-ki(n), expressing speculation and skepticism, e.g. Män kärgäk ār-di-m ār-ki 'I guess I was useful', and often used in interrogative sentences, e.g. Ol kāl-ir mū ār-ki 'I wonder whether X is coming'. This type is reflected by Tuvan ir-qi, e.g. Bar ir-qi bā? 'I wonder if X is there' (Isxaxov & Pal’mbax 1961: 433). Compare Turkish constructions such as Var mı ki?, where ki is preserved as a rhetorical particle.

In some Turkic languages, the type ĀR-KĀN may be used as a modal particle with emphatic uses, meaning ‘indeed’, ‘actually’. In this function, it is an utterance-final stance particle lacking person-number markers. It is a result of contamination with ār-ki(n), e.g. Chaghatay e(r)kin ~ e(r)kān ~ ikin. Modern markers include Uzbek -kin, Uyghur ikin, Tuvan ir-qi. It is highly improbable that ār-ki(n) developed to i-kin and later to i-kān, so that Kuman ā-gān, Uzbek e-kān, etc. are “corrupt” forms of ār-ki(n) (Gabain 1945: 149, 1959: 68). Uzbek -kin cannot possibly be described as a phonetic variant of e-kān. Uzbek mi-kin is a combination with the question particle. Combinations with the preterite are represented by -di-y-kin, -di-mi-kin, etc., e.g. Ket-di-mi-kin ‘I wonder whether X left’. The type ār-ki(n) is clearly represented by Yellow Uyghur ā̇-kin, mi̇-kin mi, utterance-final stance particles expressing subjective evaluation.

The modal type ĀR-KĀN may express speculation and skepticism, and is used in questions with the same rhetorical nuances as expressed by ār-ki(n). Unlike the evidential ĀR-KAN, it combines with preterites, e.g. Kazakh Kel-di e-ken, Uzbek Kel-di e-kān ‘X has indeed arrived’; cf. Turkish Gel-di ki!.

It also combines with conditional markers to form modal sentences expressing polite or timid wishes, e.g. Noghay Yaz-sa-q e-ken, Kazakh Ket-se-m
e-ken ‘I wish I could go’, Uzbek Yäz-sä e-kän ‘If only X would write’, Uyghur Yaz-säŋ i-kän ‘What if you would write it?’.

The modal particle är-kän is commonly used in rhetorical questions with readings such as ‘I wonder’, e.g. Kuman Kay-da ä-gän? ‘Where may X be?’, Uyghur Nämiška bol-ma-y-di-kän? ‘I wonder why it does not come about’. This is an attenuating usage in order to tone down a question, giving it meditative, skeptical or timid connotations of wondering and hesitation, similar to the use of är-ki(n).

Kazakh exhibits constructions with the interrogative suffix {-mA} and question words such as ne? what?’, kim? ‘who?’, kay-sï? ‘which?’, ne-ge, ne üšin? ‘why?’, kalay, kan-day? ‘how?’, e.g. Kel-e-di me eken?’ (I wonder:) Is X coming?’, Kel-di me e-ken?’ (I wonder:) Did X come/Has X come?’, Ne et-ti-m e-ken?’ (I wonder:) What may I have done?’, Ne bol-dï e-ken?’ (I wonder:) What may have happened?’, Xat kim-den e-ken?’ (I wonder:) From whom might the letter be?’, Kašan kel-e-di eken?’ (I wonder:) When might X come?’. Noghay distinguishes rhetorical questions such as Nege kel-gen-ler e-ken?’ (I wonder:) Why have they come?’ from evidential questions such as Nege kel-gen e-ken-ler?’ Why, obviously, have they come?’ (Karakoç 2005: #). The Uyghur rhetorical particle {+mi-kin} can co-occur with the evidential particle {- (i)kän}, e.g. Käl-gän-kän-mi-kin ‘I wonder if X appears to have come’. Uzbek displays rhetorical questions such as Kël-gän mi-kän? ‘I wonder if X has come’, whereas the marker {+mi-kin} rather expresses doubt in the sense of ‘Has X really come?’.

Utterance-final particles of the types i-yin and i-yän have exclamatory, emphatic functions, often with mirative overtones, e.g. Dukhan Gäl-di i-yän ‘X has indeed arrived’.
Tofan i-yān displays both evidential and modal functions, which cannot always be clearly distinguished from each other (Rassadin 1978: 271).

**Notes on transcriptions and translations**

Types of evidential markers are noted in small caps, e.g. *miš, gan*. Quotations from individual languages are given in italics and in traditional Turcological transcription, e.g. Azeri *Gälmis*. Citation of forms in official orthography are placed between chevrons. This is the rule in the case of Turkish citations, e.g. *Gel-miş*. Formulas summarizing bound morphemes are placed between brackets of the type `{}`. Here, capital letters indicate morphophonemic variation, e.g. `{miš}, {gan}`. `{I}` and `{A}` stand for harmonic variation of high and low vowels, respectively. In glosses, indirective inflectional markers are abbreviated as `IND.INFL` and indirective particles as `IND.PART`. In examples, constituent segments such as morphemes are divided by hyphens, contrary to the orthographic practices in the respective languages. In translations, X is used for ‘he/she/it’, e.g. Turkish *Gel-miş* ‘X has evidently come’.

**References**


