

Evidentiality in Turkic

Lars Johanson

1. Introduction

The present chapter aims at summarising the essential features of the grammatical categories of evidentiality found in Turkic languages. Other epistemic categories such as presumptives and dubitatives will not be included. The data will be dealt with along the principles suggested in Chapter 1, with special regard to the distinctive devices of each particular system type. Systems will be taken to be paradigmatic sets of evidential forms, fitting together in a contrastive way, systematically opposed to each other with respect to their functions.

Since the survey concerns a whole language family, numerous language-specific details will be omitted. 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.

Turkish examples will be cited according to the official script of Turkey. Quotations from other languages will be given in traditional Turcological transcription. In the standardised morpheme notation, capital letters and segments in brackets indicate morphophonemic variation. Thus, *I* and *A* stand for harmonic variation of high and low vowels, respectively. In the examples, morphemes will be divided by hyphens, though this does not correspond to the orthographic practice in the languages concerned.

2. Indirectivity as an evidential category

Despite system differences, almost all known older and recent stages of Turkic possess the grammatical means of expressing *indirectivity*. This firmly integrated cognitive category covers various notions traditionally referred to as ‘hearsay’, ‘inferential’, etc. It implies that the statement is ‘indirect’ in the sense that the narrated event Eⁿ is not stated directly, but in an indirect way, by reference to its reception by a conscious subject, a recipient R. This core meaning may be paraphrased ‘as R becomes (became, etc.) or is (was, etc.) aware of Eⁿ’. The result is two-layered information: ‘it is stated that Eⁿ is acknowledged by R’. 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.

Examples: Turkish *Eşek öl-üyor-muş*, Uyghur *Işek öl-üwèt-iptu* [donkey die-INTRA-1C] ‘The donkey is/was obviously dying’, Turkish *Eşek öl-müş*, Uyghur *Işek öl-üptu* [donkey die-IPAST] ‘The donkey has/had obviously died’.

The notion of indirectivity is in accordance with the crosslinguistic definition of evidentiality as ‘stating the existence of a source of evidence for some information’. The crucial element of indirectivity, the presentation of an event ‘by reference to its reception by a conscious subject’, is certainly not unique to Turkic. It may be basic to many other evidentiality systems, and even qualify as a crosslinguistic definition of evidentials (see Comrie 2000: 1).

3. Sources

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

- i. Reportive uses: ‘Eⁿ or its effect is reported to R’. The basis of knowledge is a foreign source, reported speech, hearsay, e.g. Turkish *Bakan hasta-ymış* (minister sick-1C) ‘The minister is reportedly sick’, said by somebody informed about the sickness. English translation equivalents include *reportedly*, *allegedly*, *as they say/said*, etc.
- ii. Inferential uses: ‘Eⁿ or its effect is inferred by R’. The basis of knowledge is pure reflection, logical conclusion, e.g. Turkish *Uyu-muş-um* (sleep-IPAST-1SG) ‘I have obviously slept’, said by somebody just waking up. English translation equivalents include *as far as R understands/understood*, etc.

- iii. Perceptive uses: 'E' or its effect is perceived by R'. The basis is first-hand knowledge, direct sensory perception of the event or indirect perception on the basis of traces or results, e.g. Turkish *İyi çal-ıyor-muş* (good play-INTRA-IC) '(S)he is, as I hear, playing well', said by somebody just listening. English translation equivalents include *it appears/appeared that, it turns/turned out that, as R can/could see, hear, etc.*

All these readings can be translated by *evidently, obviously*, etc. If the Turkish sentence *Ali bu-nu bil-iyor-muş* [A. this-ACC know-INTRA-IC] is rendered as 'Ali evidently knows/knew this', *evidently* may be interpreted as judging from (i) hearsay, (ii) inference, or (iii) appearance.

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. These uses cannot be derived from reportive or inferential meanings or be subsumed under 'non-first-hand knowledge'. Indirectivity markers thus do not fit into evidential schemes distinguishing between 'the speaker's non-firsthand and firsthand information'. Their primary task is not to express the foreign origin of the speaker's knowledge.

Some more elaborated 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.

4. Marked and unmarked terms

Turkic languages display basic contrasts between marked indirectives and their unmarked counterparts. Functionally marked terms, expressing the evidential notions explicitly, stand in paradigmatic contrast to non-evidentials. Thus, Turkish exhibits items signalling indirectivity, e.g. *gel-miş* [come-IPAST] 'has obviously come/obviously came' and *gel-iyor-muş* [come-INTRA-IC] 'is/was obviously coming, obviously comes'. It has corresponding unmarked items such as *gel-di* [come-DPAST] 'has come/came', *gel-iyor* [come-INTRA] 'is coming/comes'.

Although the relations between marked and unmarked terms vary across languages, the unmarked ones always exhibit neutral uses in cases where the speaker considers the evidential distinction unessential and thus chooses not to use it. The widespread opinion that unmarked items such as *gel-di* 'has

come/came' consistently signal 'direct experience' or 'visual evidence' is incorrect. Unmarked items 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.

5. Formal types of markers

The coding of indirectivity in Turkic is 'scattered', morphologically realised by two types of markers:

- i. Inflectional markers: Verbal suffixes filling obligatory slots and opposed to other inflectional morphemes, e.g. -MIŞ, -GAN, -IBDIR.
- ii. Copula particles: Enclitic or suffixed elements, occurring after predicate cores, filling non-obligatory slots and opposed to zero forms, e.g. ERMIŞ, ERKEN.

6. Turkish examples

Turkish possesses the simple inflectional marker *-miş*, which is suffixed to verbal stems and marks main clauses. It is capable of carrying high pitch. It is a rather stable marker of indirectivity, mostly with past time reference, e.g. *gül-'müŝ* [laugh-IPAST] '(has) evidently laughed' (*gül* 'to laugh' + *-miş*). It is homonymous with the postterminal participle suffix *-miş*, which does not, however, express indirectivity.

Turkish also possesses the copula particle *imiŝ*, which cannot be added to verbal stems, but may follow nominals, including nominal stems of the verb. It is not capable of carrying high pitch, a characteristic of enclitic particles. Like its equivalents in other Turkic languages, it is temporally indifferent, i.e. ambiguous between past and present time reference, and a stable marker of indirectivity. Since *imiŝ* normally exhibits suffixed allomorphs in the shape of *-(y)müş*, some of its written realisations coincide with those of the inflectional marker *-miş*, e.g. *'gülmüş* [rose-1C] 'it is/was evidently a rose' (*gül* 'rose' + *imiŝ*). In spoken language, the variants of the two morphemes are distinguished by different pitch patterns (cf. Csató 2000a).

This copula particle may, for example, form indirective intraterminals, i.e. presents and imperfects, e.g. *gel-iyor-muş* [come-INTRA-1C] 'is/was evidently arriving', prospectives such as *gel-ecek-miş* [come-PROSP-1C] 'will/would evi-

dently arrive' and unequivocally indirective postterminals such as *gel-miş-miş* [come-POST-IC] 'has/had reportedly arrived'.

7. Aspectual values

The deceptive similarity of certain allomorphs has led linguists to confuse the two markers *-miş* and *imiş*, referring to both as 'the suffix *-miş*', allegedly attachable to both verbal and nominal stems, or to speak of an 'evidential perfect' that would include both markers. In view of such misunderstandings, some comments on the aspectual values of the respective items may be appropriate. The perspectives of *intraterminality* and *postterminality* – two marked aspectual ways of envisaging events with respect to their limits – are grammaticalised in Turkic as well as in many other languages (Johanson 1996a, 2000).

The intraterminal perspective envisages, at a given aspectual vantage point, an event within its limits, e.g. Azerbaijani *gel-ir-em* [come-INTRA-1SG] 'I am coming, I come'. Intraterminals are marked 'imperfectives' with 'relative present' meanings of 'the state of verb-ing'.

The postterminal perspective implies that, at a given aspectual vantage point, the decisive limit of the event is already passed over, e.g. Azerbaijani *gel-miş-em* [come-PPAST-1SG] 'I have [already] arrived'. Postterminals express 'the state of having verb-ed' and can thus refer to preexisting events in an indirect way. The event is totally or partly absent from the view, depending on the actional content, but it is still relevant at the vantage point, possibly through observable results or traces.

In many languages, postterminals, e.g. 'perfects' expressing past events with present relevance, tend towards indirective readings (Johanson 1971: Chapter 8, 2000: 121–123). 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 vantage point. Postterminals may thus suggest that information becomes available post factum, that a preexisting state is discovered, etc. These secondary meanings are pragmatic side effects that can be used as 'evidential strategies'. What indirectives share with 'perfects' as a common core meaning is not the notion of 'distance', but rather that of postterminality.

The functions of Turkic indirective markers such as *-miş* derive from the indirect postterminal perspective. However, they should not be equated with pure postterminals used as 'evidential strategies'. Although some of them may vacillate with respect to their semantics, they are, as a rule, more explicit and stable evidentials, whose indirect aspectual value has been reinterpreted as in-

directivity. On the other hand, copula particles such as *imiş* have, through grammaticalisation processes, completely lost their relationship to the post-terminal value and cannot possibly be considered 'perfect' markers, as some linguists have assumed.

8. Semantic types of inflectional markers

Two or more of the following semantic types of inflectional markers are generally found in Turkic languages:

- IPAST-1. Unequivocal indirective pasts, e.g. -IBDIR: Kazakh *-(I)ptI*, etc.
- IPAST-2. Less stable indirective pasts, e.g. -MIŞ, -GAN: Turkish *-mİş*, Tatar *-GAn*, etc.
- PPAST. Postterminal pasts with secondary indirective readings, e.g. -MIŞ, -GAN: Azerbaijani *-mİş*, Kazakh *-GAn*, etc.
- DPAST. Direct pasts, not signalling indirective meanings: -DI.

9. Semantic types of copula particles

The temporally indifferent indirective copula particles express indirectivity without any synchronic association with postterminality. In certain systems, two copula particles divide the area of indirectivity between themselves according to the pattern reportive versus non-reportive (inferential + perceptive) uses. We thus find the following semantic types of indirective items.

- IC-1. General indirectives, e.g. ERMIŞ: Turkish *-(y)mİş*, etc.
- IC-2. Indirectives with reportive meanings, e.g. ERMIŞ: Turkmen *-mİş*, etc.
- IC-3. Indirectives with non-reportive meanings, e.g. ERKEN: Kazakh *eken*, etc. A similar type, frequent in spoken registers, is BOLIBDIR 'evidently is/was': Kazakh, Noghay *bolipti*, Uzbek *bolip*, Altay *boluptir*, *boltir*, Khakas *poltir*, Tatar *bulyan*, etc.

10. System types

The types of evidentiality systems and their organisation vary considerably across the modern Turkic languages. Most systems possess clearly evidential

categories with indirectivity as the primary meaning. The following survey will begin with the more comprehensive systems and proceed to the simpler ones.

10.1 Three-term systems

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

Languages of this type exhibit three-term subsystems of inflectional markers, consisting of an indirective past in *-IBDIR*, a postterminal in *-GAN*, and a direct past in *-DI*.

The marker *-IBDIR* is a stable indirectivity marker of the type *IPAST-1*: 'has evidently done, evidently did', e.g. Uyghur *yèz-iptu*, Uzbek *yâz-ibdi* [write-*IPAST*] 'has evidently written, evidently wrote', Kazakh *tüs-ipti* [fall-*IPAST*] 'has evidently fallen, evidently fell', Turkmen *gid-ipdir* [go-*IPAST*] 'has apparently gone'.

The marker *-GAN* is a postterminal of the type *PPAST*, displaying perfect-like meanings with occasional indirective readings, e.g. Uyghur *yaz-yan*, Uzbek *yâz-yan* [write-*PPAST*] 'has written', Kazakh *öltir-gen* [kill-*PPAST*] 'has killed', Turkmen *öylön-ön* [get.married-*PPAST*] 'has married', 'is married'.

In languages of this type, two indirective copula particles, *ERKEN* and *ERMIŞ*, combine with various predicates to express 'is/was obviously'. *ERMIŞ* tends towards reportive uses (type *IC-2*), and *ERKEN* towards non-reportive, i.e. inferential and perceptive, uses (*IC-3*).

ERKEN combines with intraterminals, i.e. presents and imperfects, prospectives, non-verbal predicates, etc, e.g. Uyghur *yèz-ivatqan iken* [write-*INTRA IC*] 'is/was evidently writing', Kazakh *kel-edi eken* [come-*INTRA IC*] 'is/was evidently writing', *üyde eken* [at.home *IC*] 'is/was obviously at home'. It combines with postterminals to express 'has/had obviously been or done', 'turns out/turned out to have been or done', e.g. Uyghur *tüget-ken iken* [finish-*POST IC*] 'has/had obviously finished', Uzbek *yâz-yan eken* [write-*POST IC*] 'has/had obviously written', Kazakh *tüs-ken eken* [fall-*POST IC*] 'has/had obviously fallen'.

ERMIŞ expresses corresponding reportive meanings, e.g. Uyghur *yèz-ivatqan-miş* [write-*INTRA-IC*] 'is/was evidently writing', Kazakh *kel-edi-miş* [come-*INTRA-IC*] 'is/was reportedly arriving'. With postterminals, it forms unequivocally indirective past items, e.g. Uyghur *yaz-yan-miş* [write-*POST-IC*] 'has/had allegedly written', Turkmen *gid-ip-miş-in* [go-*POST-IC-3SG*] 'has/had reportedly gone'.

There are also items of the structure *-IBDIR + ERMIŠ*, which apply reportive meaning to inferential or perceptive statements, ‘is/was reportedly evident’, e.g. Kazakh *kel-ipti-mis* [come-IPAST-IC]. Uyghur *yèz-ipti-miš* [write-IPAST-IC] ‘has/had allegedly written’ is thus interpretable as a more reliable statement than *yaz-yan-miš* [write-POST-IC], since it suggests the existence of something perceived or inferred.

The distribution discussed here allows three-term subsystems of copula particles of the type ‘indirective’ [‘reportive’ versus ‘non-reportive’] versus ‘unmarked’. However, the opposition ‘reportive’ versus ‘non-reportive’ is often limited to certain dialects or registers. Thus, *ERMIŠ* is not used in all varieties of Uyghur and Uzbek, and its role in Kazakh is also rather limited.

10.2 Split three-term/two-term systems

Some languages exhibit three inflectional markers, indirective, postterminal and direct past, but a simplified subsystem of copula particles: ‘indirective’ versus ‘unmarked’.

Noghay, of the Northwestern branch, has *IPAST-1*, *PPAST*, and *DPAST*, e.g. *kel-ipti* [come-IPAST] ‘evidently arrived’, *kel-gen* [come-PPAST] ‘has arrived’, *kel-di* [come-DPAST] ‘arrived’. Its copula particle *ERKEN* is a general indirective marker of the semantic type *IC-1*, thus covering both reportive and non-reportive meanings. It combines, for example, with intraterminals, e.g. *kel-e-di eken* [come-INTRA-3SG IC] ‘is/was obviously arriving’, and with postterminals and indirective pasts to form indirectives signalling relative anteriority, e.g. *kel-gen eken* [come-POST IC], *kel-ipti eken* [come-IPAST IC] ‘has/had obviously arrived’. *BOLIBDIR* has a similar function, e.g. *qal-yan bolipti* [remain-POST IC] ‘has/had obviously stayed’.

Certain other 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, *-GAN* is used without a competing *-IBDIR*, thus representing the type *IPAST-2*. It displays postterminal uses but may also suggest indirectivity, e.g. *yaz-yan* [write-IPAST] ‘has [evidently] written’. The neighbouring language Chuvash has a similar marker *-nĀ* with postterminal and indirective meanings, e.g. *kala-naĀ* [speak-IPAST] ‘has [evidently] spoken’.

Tatar, Bashkir, and Chuvash possess both *ERMIŠ* (reportive) and *ERKEN* (non-reportive), e.g. Chuvash *kil-nĕ imeš* [come-POST IC] ‘has reportedly arrived’, *kil-nĕ ikken* [come-POST IC] ‘has evidently arrived’. The Tatar parti-

cle *bulyan*, which belongs to the type BOLIBDIR, has functions similar to those of ERKEN, combining, for example, with intraterminals, postterminals, and prospectives: *bar-a bulyan* [go-INTRA IC] 'is/was evidently going', *bar-yan bulyan* [go-POST IC] 'has/had evidently gone', *bar-ačaaq bulyan* [go-PROSP IC] 'will/would evidently go'.

10.3 Two-term systems

A few systems just consist of one inflectional marker and one copula particle plus their unmarked opposition partners. An inflectional marker -MIŠ of the type IPAST-2 is used in the western subgroup of the Southwestern branch, e.g. Turkish *-miş*. The cognate item *-BIT* is used in Yakut, the northernmost language of the Northeastern branch, spoken at the opposite extreme of the Turkic world. The languages in question also possess ERMIŠ particles of the type IC-1, e.g. Turkish *imiş*, Yakut *ebit*.

Since -MIŠ lacks a competing -IBDIR, and ERMIŠ lacks a competing ERKEN, these languages display two-term subsystems of the type indirective versus non-indirective. The inflectional markers allow reportive, inferential and perceptive readings, thus corresponding to several items in more comprehensive systems, e.g. Kazakh *-(I)ptI*, *-(I)ptImIs*, *-GAn eken* and *-GAn*.

Turkish -MIŠ, which mainly displays indirective uses, is opposed to an unmarked simple direct past in -DI, which negates the notion of indirectivity but also displays neutral uses. Since Turkish lacks a competing pure postterminal, this item covers both 'perfect' and 'preterite' functions, e.g. *gel-di* [come-DPAST] 'has come/come'. The copula particle is a stable general indirectivity marker with reportive, inferential and perceptive uses. A complex item -MIŠ + ERMIŠ applies an explicitly indirective type of evidentiality to a postterminally envisaged event. It is often used for rumours, gossip, etc., e.g. *gel-miş-miş* [come-POST-IC] 'has/had reportedly arrived'.

Yakut has a similar system (see Buder 1989). The inflectional marker -BIT conveys reportive, inferential and perceptive nuances, e.g. *kel-bit* [come-IPAST] 'has [obviously] arrived'. The temporally indifferent indirective particle *ebit* is similar to Turkish *imiş* and allows similar combinations with intraterminals, postterminals, etc., e.g. *tur-ar ebit* [stand-INTRA IC] 'evidently stands/stood', *kel-bit ebit* [come-POST IC] 'has/had evidently arrived'.

There are still smaller evidentiality systems, in which the inflectional marker represents the type PPAST, i.e. a postterminal with secondary indirective readings. Thus, Azerbaijani possesses, like Turkish, an indirective copula particle ERMIŠ of the semantic type IC-1, *imiş*. However, the status of its inflec-

tional marker *-mİş*, which forms a common mixed paradigm with *-(I)b*, is different from that of Turkish *-mİş*. It represents an ambivalent type with postterminal meanings, e.g. *gel-miş-em* [come-PPAST-1SG] ‘I have arrived’, *yaz-ib-sin* [write-PPAST-2SG] ‘you have written’. The unmarked term *-DI* thus tends towards ‘preterite’ rather than ‘perfect’ functions, e.g. *gel-di* [come-DPAST] ‘came’ versus *gel-ib* [come-PPAST] ‘has come’. The type *-MIŞ + ERMIŞ* unambiguously applies indirectivity to postterminally envisaged events, e.g. *yaz-miş-miş* [write-POST-IC] ‘has/had reportedly written’.

11. 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.

11.1 Testimony, involvement, control

Indirectives may evoke the impression that the recipient does not or did not witness the event, perceive it, or participate in it consciously; that (s)he is or was not present at the event, not in control of it, not directly involved in it. However, despite the indirect way of presentation, these meanings are not signalled explicitly. The indirectly marked event may indeed be apprehended by the recipient through the senses, consciously taken part in, etc. Lack of participation or control is limited to certain contexts and cannot be the common core meaning of indirectives. The source of information may be direct evidence, personal, even visually obtained knowledge. A sentence such as Turkish *Ahmet gel-miş* or Uyghur *Exmet ke-ptu* [A. come-IPAST] ‘Ahmed has [as I note] arrived’ can be felicitously uttered by a speaker who has witnessed the arrival in reality. The indirective statement just expresses the conscious reception. As is often the case with linguistic devices, it does not tell us how something is in reality, but rather how the speaker chooses to present it.

On the other hand, evidentially unmarked terms may suggest that the source of information is direct experience, that the speaker takes/has taken part in the event consciously, is/was in control of it, etc. Again, the unmarked terms do not signal these meanings explicitly, but may also be used for unwitnessed, uncontrolled, reported or inferred events, e.g. Turkish *Çok büyü-dün* [much grow-DPAST-2SG] ‘You have become very big’, *Kemal Paşa, Selânik’te doğdu* [Atatürk(Kemal Paşa) Salonika+LOC be.born+DPAST] ‘Atatürk was born

in Salonika'. They just lack the two-layered information typical of indirectives, and may thus be used whenever this specific information seems unessential.

11.2 Factuality

Turkic indirectives may also 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 and be interpreted as non-testimonial reference. Indirectives, in particular reportive items, can be used to disclaim direct responsibility for the truth of the statement, suggesting that the speaker is not the originator of the information or does not vouch for it. By contrast, unmarked terms may suggest that the speaker is certain of the truth of information and even responsible for it.

'Supposition' is sometimes claimed to be the main meaning of Turkic indirectives. The corresponding unmarked items are said to signal that the speaker regards the event as certain. However, indirectives are not presumptives or dubitatives reducing the factuality of the statement. Their task is not to express the speaker's attitude to the truth of the content, to signal doubt or conjecture concerning the information conveyed.

Turkic languages have other ways of expressing supposition or conjecture, e.g. the particle *-DIR*, added to postterminals, intraterminals and other forms, e.g. Turkish *yaz-mış-tır*, Uzbek *yâz-yan-dir*, Uyghur *yaz-yan-du* [write-POST-SUPP] 'has presumably written', Turkish *oku-yor-dur*, Turkmen *oqo-ya:n-nir* [read-INTRA-SUPP] 'is presumably reading', Uyghur *işle-vatqan-du* [work-INTRA-SUPP] 'is presumably working'.

11.3 Distance

'Distance' is another possible contextual realisation. It has sometimes even been suggested as the common core meaning of Turkic indirectives. Thus, Turkish *-MIŞ* and its counterparts, e.g. in Old East Turkic, have been referred to as 'preterites of distance'. It is true that some kind of distance is likely to be involved if one does not refer directly to the event itself, but rather to the reception of it. By virtue of their central value, indirectives may imply cognitive or emotional distance from the event. Speakers may use them in order to distance themselves from the event, to distinguish themselves from the responsibility for it, or as a way of being vague about sources that they do not want to lay open to view. The notion is, however, too vague to possess any explanatory force.

11.4 Irony

Indirectives may, in particular contexts, display pragmatic extensions of an attitudinal nature. One kind of dissociation from the event may be an ironic relation to it, reservation interpretable as sarcasm, disdain, etc. e.g. *Bunu yap-acak-mış-sın* [this-ACC DO-PROSP-IC-2SG] ‘You will obviously (as you think) do this’.

11.5 Caution, modesty, summarising

The limitation to an indirective statement may be motivated by caution, modesty, need for a summarising view, etc. Thus, indirective inflectional markers may represent events in a complexive, experiential, summarising way, frequently with nuances of modesty, e.g. Turkish *Ben her zaman vazife-m-i yap-mış-ım* [I each time duty-POSS.1SG-ACC DO-IPAST-1SG] ‘I have [as it appears] always done my duty’, cf. Uyghur *Men daim vezipe-m-ni ada qi-pti-men* [I always duty-POSS.1SG-ACC fulfilment DO-IPAST-1SG], Turkish *Önemli bir konu ele al-mış-sın* [important a topic hand-DAT take-IPAST-2SG] ‘You have [if I may summarise] addressed an important topic’. Readings of these kinds derive from the indirect postterminal perspective.

11.6 Mirative connotations

Indirectives of the types *-İBİR*, *-MIŞ*, *ERKEN* and *ERMİŞ*, etc. may convey mirative connotations. Their use may, in particular contexts, be interpreted in terms of new knowledge, discovery, sudden awareness of revealed facts, surprise, mental unpreparedness, perception contrary to one’s expectations, admiration, etc. Such readings follow naturally from the notion of indirectivity. The conscious reception (‘as it turns out/turned out’) may be sudden or unexpected; what the recipient turns the mind to may come as a surprise. 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. But 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. So-called ‘hot news’ is typically expressed by the direct past marker *-DI*. On the other hand, certain Turkic languages possess other means of expressing unexpected events, e.g. miratives particles such as Altay Turkic *turbay*.

In sentences such as Kazakh *Ol ket-ip qal-ipti*, Uyghur *U kêt-ip qa-ptu* [(s)he go-CONV remain-IPAST] '(S)he has left [as I note]', or exclamations such as Turkish *Bu kız ne güzel-miş!* [this girl what beautiful-1C] 'How beautiful this girl is!', Uyghur *Bu qız çiraylıq iken!* [this girl beautiful 1C] 'This girl is beautiful!', Turkish *Büyü-müş-sün!* [grow-IPAST-2SG], Uyghur *Çoñ bol-up sen!* [big become-IPAST 2SG] 'You have grown!', the indirective marker just adds the meaning 'as I am/become aware of', which is the central value of the category. Contrariness to the speaker's expectation is, however, not part of this value.

12. Differences between persons

Although evidential specifications are possible in all 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 traditional definition of indirectivity as the expression of 'the speaker's non-firsthand information' is problematic because of its frequent incompatibilities with the first person. The use of indirectives when speaking of oneself then necessarily implies lack of awareness, volitionality, consciousness or control due to inattention, sleep, drunkenness, coma, etc. However, with a definition based on the presentation of the event 'by reference to its reception by a conscious subject', it is by no means contradictory to use indirectives with firstperson subjects whose referents are aware of the event in question.

In Yellow Uyghur, a small Turkic language spoken in Western China, the second and third person of certain tenses tends to take the evidential form $-(i)pti$, $-(i)tti$ = -IBDIR, denoting that the event is beyond control of the speaker, whereas the first person takes the corresponding non-evidential form, $-(^h)ti$ = -DI. Thus, a unified past tense paradigm has been suggested, in which the first and non-first persons take different suffixes, e.g. *men pa^hr-ti* 'I went' = -DI, *sen part-tti* 'you went' = -IBDIR (Roos 1999: 105–106). In earlier descriptions, however, these two types were dealt with as two separate tenses (Tenišev 1976: 92–93).

13. Correlations with other grammatical categories

Turkic evidentials are limited to main clauses with an asserted, contradictable content. They are not allowed in non-indicative moods such as optative and imperative.

They may, however, cooccur with necessitives or debitives, e.g. Turkish *Git-meli-ymiş-sin* [go-NEG-IC-2SG], Uyghur *Sen kêt-iş-iñ kèrek iken* [you go-NOM-2SG necessary IC] ‘You evidently ought to go’.

In negative sentences, evidentials are not within the scope of negation. The narrated event itself is negated, not its reception by the conscious subject, e.g. *U kel-me-ptu* [(s)he come-NEG-IC] ‘(S)he has not arrived [as I note]’.

Turkic indirectives may also occur in interrogative sentences, e.g. Turkish *O böyle de-miş mi?* [(s)he so say-IPAST Q], Uyghur *U mundaq de-ptu-mu?* [(s)he so say-IPAST-Q] ‘Did (s)he reportedly/evidently say so?’, Kazakh *Kel-e mi eken?* [come-INTRA Q IC] ‘Is (s)he, as it appears, coming?’, Üyde *mi eken?* [at.home Q IC] ‘Is (s)he, as it appears, at home?’, Noghay *Ne-ge kel-gen eken-ler?* [what-DAT come-POST IC-PL] ‘Why have they, as it appears, come?’. Indirectives may also be used in questions asked on behalf of someone else. They do not express the speaker’s assumption about the addressee’s source of information.

Though evidentiality specifications are relatively independent of aspect-tense choice, there are certain cooccurrence restrictions. Indirective copula particles are incompatible with the simple past *-DI* and the related copula forms *edi*, *idi* etc. ‘was’, e.g. Turkish **-DI-ymlş*, **(y)di-ymlş*.

14. Modal uses of ERKEN

Certain uses of ERKEN seem to be exceptions to the last rule of the preceding section. In some languages, ERKEN may also be used as a modal particle meaning ‘indeed’, etc., in which case it does not take personal suffixes, but is added to complete main clauses. In this function, ERKEN may also cooccur with *-DI*, e.g. Kazakh *Kel-di eken!* [come-DPAST MOD] ‘(S)he has indeed arrived!’. It can also combine with conditional markers to form modal sentences expressing modest, polite or timid wishes, e.g. Noghay *Yaz-sa-ñ eken*, Uyghur *Yaz-sa-ñ iken* [write-COND-2SG MOD] ‘If you would write it up [it would be fine]’.

With interrogative elements, the modal variant of ERKEN forms rhetorical questions, with readings such as ‘I wonder’, e.g. Kazakh *Ne et-ti-m eken?* [what do-DPAST-1SG MOD] ‘I wonder what I have done’, Noghay *Nege bol-ma-y-di eken?*, Uyghur *Nemişqa bol-ma-y-di-ken?* [why become-NEG-INTRA-3SG MOD]

'I wonder why it does not come about'. Two examples from Noghay may illustrate the difference between the modal and the evidential use of ERKEN. *Ne-ge kel-gen-ler eken?* [what-DAT come-POST-3PL MOD], where the particle is added to the complete main clause, is a rhetorical question meaning 'I wonder why they have come'. In the corresponding indirective sentence, ERKEN takes the personal suffix: *Ne-ge kel-gen eken-ler?* [what-DAT come-POST 1C-3PL] 'Why have they reportedly [etc.] come?' (Karakoç 2001:38).

15. Evidentiality and discourse

Indirectives may play various roles according to different discourse types. In traditional story-telling, e.g. in fairytales, they play the role of detaching the narrator from the narrated events, e.g. Uyghur *Burun bir padiša öt-üptü; un-iñ bir bali-si bar iken* [formerly a ruler live-IPAST; (s)he-GEN a child-POSS.3SG existing 1C] 'Once upon a time there was a king who had a son'. On traditional -MIŞ-based narratives in Turkish, see Johanson (1971:79–80). Independently of their status as genuine indirectives or indirectly interpretable postterminals, -MIŞ, -IBDIR and -GAN may serve as propulsive, 'plot-advancing', basic items of this discourse type. In historical narratives, however, direct pasts of the type -DI are used as the basic items; as mentioned, also for events unwitnessed by the speaker. Evidentials are not typically used for describing dreams or content conveyed by written, auditory or visual media.

16. Origins

Turkic indirective markers of the inflectional type originate in postterminals. The development of more stable indirective meanings may be seen as the grammaticalisation of an evidentiality strategy, a semantic extension in the sense of conventionalised implicatures. The markers -MIŞ and -GAN are of unknown origin, but may have developed from lexemes. For example, it is conceivable, though hardly provable, that -MIŞ goes back to a form of a verb of the type *bış-* 'to ripen, to mature' (attaining a final state such as 'ripe', 'cooked' or 'done').

Markers of the formal type -IBDIR, e.g. Uyghur *yöz-ipti* 'evidently wrote', Salar *gel-du* 'evidently came', Tuvan *ber-iptir* 'evidently gave', go back to the postterminal periphrasis -(I)b *tur-ur*, consisting of a converb of the lexical verb plus *tur-ur* 'stands', e.g. *yaz-ıb tur-ur* ('stands having written') 'is in the state of

having written', 'has written'. The type BOLIBDIR consists of *bol-* 'to become, to be' plus the indirective suffix *-IB[DIR]*.

The copula particles ERMIŠ and ERKEN are derived from *er-* 'to be', though not regular forms in *-MIŠ* and *-GAN*. Both may be of postterminal origin, if *er-* was originally an initiotransformative verb (see Johanson 2000: 62–63) expressing (i) an initial dynamic phase 'to become' and (ii) a subsequent stative phase 'to be'. The postterminal perspective of such a verb may envisage the event as still going on at the aspectual vantage point: *er-miš* 'has become (evident)', 'has appeared' = 'is (evident)', 'appears'.

17. Contact-induced codecopying

Indirectives play a central part in almost all Turkic languages. However, due to influence from Indo-European languages such as Persian, Greek and Slavic, a few languages and dialects only exhibit 'evidentiality strategies'. The tendency of Azerbaijani *-miš/-(I)b* towards pure postterminal ('perfect') readings is most probably due to Persian influence, e.g. *yap-ib* 'has done', cf. Persian *karde(-ast)*. There are also other influences. For example, the Yellow Uyghur system is reminiscent of the Tibetan verbal paradigm (DeLancey 1986). Evidentiality systems are lacking in a few Turkic languages and dialects which have been strongly influenced by Indo-European, e.g. Karaim in Lithuania, under Slavic and Lithuanian impact (Csató 2000b), and 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 in Southwestern and Central Asia, Southeastern and Northeastern Europe, etc. 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. For example, Northern Tajik has developed a comprehensive evidential system on the Uzbek model. Indirective functions have been copied (i) onto postterminals of the 'perfect' type (as opposed to 'aorists'), and (ii) onto related participles, on the model of the temporally indifferent ERKEN or ERMIŠ, e.g. Bulgarian *bil* (Johanson 1996b) and Western Armenian *eyer*. I also suggest that Hungarian *igen* 'yes' goes back to a Turkic form ERKEN 'evidently'.

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. The question is whether systems based on marked confirmatives may have 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). This suggestion is in line with the basic principle of the Code Copying framework (see, e.g., Johanson 2002) to the effect that a code copy is never identical with its model.

References

- Aksu-Koç, A. A., & Slobin, D. (1986). A psychological account of the development and use of evidentials in Turkish. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The linguistic coding of epistemology* (pp. 159–167). Norwood, NJ: Ablex.
- Brendemoen, B. (1997). Some remarks on the *-mİş* past in the Eastern Black Sea coast dialects. *Turkic Languages*, 1, 161–183.
- Buder, A. (1989). *Aspekto-temporale Kategorien im Jakutischen*. [*Aspect-tense categories in Yakut*.] (Turcologica 5.) Wiesbaden: Harrassowitz.
- Comrie, B. (2000). Evidentials: semantics and history. In L. Johanson & B. Utas (Eds.), *Evidentials. Turkic, Iranian and neighbouring languages* (pp. 1–12). Berlin and New York: Mouton de Gruyter.
- Csató, É. Á. (2000a). Turkish MİŞ- and İMİŞ-items. Dimensions of a functional analysis. In L. Johanson & B. Utas (Eds.), *Evidentials. Turkic, Iranian and neighbouring languages* (pp. 29–43). Berlin and New York: Mouton de Gruyter.
- Csató, É. Á. (2000b). Some typological features of the viewpoint and tense system in spoken North-Western Karaim. In Ö. Dahl (Ed.), *Tense and aspect in the languages of Europe* (pp. 27–187). Berlin and New York: Mouton de Gruyter.
- DeLancey, S. (1986). Evidentiality and volitionality in Tibetan. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The linguistic coding of epistemology* (pp. 203–213). Norwood, NJ: Ablex.
- DeLancey, S. (1997). Mirativity: The grammatical marking of unexpected information. *Linguistic Typology*, 1, 33–52.
- Johanson, L. (1971). *Aspekt im Türkischen. Vorstudien zu einer Beschreibung des türkeitürkischen Aspektsystems*. [*Aspect in Turkish. Prolegomena to a description of the Turkish aspect system*.] Uppsala: Almqvist & Wiksell.

- Johanson, L. (1996a). Terminality operators and their hierarchical status. In B. Devriendt, L. Goossens, & J. van der Auwera (Eds.), *Complex structures: A functionalist perspective* (pp. 229–258). Berlin and New York: Mouton de Gruyter.
- Johanson, L. (1996b). On Bulgarian and Turkic indirectives. In N. Boretzky, W. Enninger, & T. Stolz (Eds.), *Areale, Kontakte, Dialekte. Sprache und ihre Dynamik in mehrsprachigen Situationen* (pp. 84–94). Bochum: Brockmeyer.
- Johanson, L. (2000). Viewpoint operators in European languages. In Ö. Dahl (Ed.), *Tense and aspect in the languages of Europe* (pp. 27–187). Berlin and New York: Mouton de Gruyter.
- Johanson, L. (2002). *Structural factors in Turkic language contacts*. London: Curzon.
- Karakoç, B. (2001). Studien zum finiten Verbalsystem des Nogaischen. [Studies of the finite verb system in Noghay.] Ph.D. thesis, University of Mainz.
- Roos, M. E. (1999). The Western Yugur (Yellow Uygur) language. Grammar, texts, vocabulary. Ph.D. thesis, University of Leiden.
- Tenišev, É. R. (1976). *Stroj saryg-jugurskogo jazyka*. [*The structure of the Yellow-Uyghur language*.] Nauka: Moskva.

Studies in Evidentiality

Edited by

Alexandra Y. Aikhenvald

R. M. W. Dixon

Research Centre for Linguistic Typology, La Trobe University

John Benjamins Publishing Company
Amsterdam/Philadelphia



™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Library of Congress Cataloging-in-Publication Data

Studies in evidentiality / edited by Alexandra Y. Aikhenvald, R. M. W. Dixon.

p. cm. (Typological Studies in Language, ISSN 0167-7373 ; v. 54)

“Revised versions of presentations at the International Workshop on Evidentiality organized by the Research Centre for Linguistic Typology at La Trobe University, 6–11 August 2001” — Pref.

Includes bibliographical references and indexes.

1. Evidentials (Linguistics)--Congresses. 2. Typology (Linguistics)--Congresses. I. Aikhenval'd, A. IU. (Aleksandra IUr'evna) II. Dixon, Robert M.W. III. International Workshop on Evidentiality (2001 : La Trobe University) IV. Series.

P325.5.E96 S78 2003

415'.01-dc21

2002038237

ISBN 90 272 2962 7 (Eur.) / 1 58811 344 2 (US) (Hb; alk. paper)

© 2003 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA