From Fables to Newspapers: Turkish and Bulgarian Evidentiality in Written Genres

Vesela Simeonova
University of Graz, Austria

Seckin Arslan
CNRS, France

Abstract   We examine evidentiality in Turkish and Bulgarian across seven written genres (news, history, police reports, fiction, fables, religious texts, and dreams). The results reveal genre-specific patterns and show that similar markers are used differently in each language, highlighting cross-linguistic contrasts in evidential interpretation.

Keywords   Evidentiality. Bulgarian. Turkish. Narrative genres. Dream language.

1   Introduction

Evidentiality is the linguistic encoding of the source of information.[1] While all languages can express evidential meanings lexically (e.g. "reportedly" in English), about a quarter have specialized grammatical evidential paradigms (Aikhenvald 2004).

One of the major theoretical debates is whether evidentiality constitutes an independent grammatical category (see Aikhenvald 2004; Plungian 2002; De Haan 1999), or it is a part of the epistemic modality system.[2] While previous studies have focused on conversational data, written genres, due to their inherent limitations to certain types of evidence and degrees of certainty, provide a new testing ground for the two hypotheses. For example, reputable news reports and historical and encyclopedic texts convey a high degree of certainty (they communicate facts) but are based on indirect evidence. Conversely, narrating dreams describes events of low certainty but for which the speaker has, as it were, first-hand experience.

This paper documents evidential uses across seven narrative genres: traditional fables, modern fiction, religious texts (e.g. the Bible), news (journalism), police reports, texts about historical events, and dreams. The empirical focus of the study is on Bulgarian and Turkish, two languages that have long been in areal contact within the Balkan language area (Friedman 2008) with comparable yet different evidential systems. This allows us, additionally, to reveal finer-grained differences between them that may not manifest in prototypical conversational use.

We found that (i) evidential marking interacts with certainty in genre-specific ways; (ii) Bulgarian and Turkish show similar patterns in some genres, but differ in others; (iii) non-evidential forms interact with the use of evidential markers in complex ways.

2   Evidentiality in Bulgarian and Turkish

This section presents an overview on the evidential systems of Bulgarian and Turkish informed with data from constructed conversational scenarios.

---

[1] See Aikhenvald 2004; Chafe, Nichols 1986; Johanson, Utas 2000; Plungian 2001; Willett 1988.

[2] See Palmer 2001; Chafe, Nichols 1986; Givón 1982; Willett 1988; Izvorski 1997; Smirnova 2013; 2021.

2.1   Evidentiality in Bulgarian

Evidentiality in Bulgarian has been studied from multiple theoretical and methodological perspectives.[3] Below is a brief overview of the main properties relevant for this paper.

Bulgarian evidentiality is expressed by morphology intertwined with tense and aspect on finite verbs. In the past tense, which is the target data of this paper, evidential marking is obligatory and comes in three options illustrated below: direct (1), indirect (2), and reportative (3). The sentences are temporally and aspectually identical – they all use the verb 'to give birth' in the preterite ('aorist') tense (and perfective aspect) but are evidentially distinct. We propose that evidentials in Bulgarian in their prototypical use encode information source, as follows.

The direct evidential is used when the speaker witnessed or, more abstractly, consciously participated in a situation (cf. Korotkova 2020; Simeonova 2021). The sentences "My grandmother gave birth to my father in 1956" (1a) and "I gave birth to my son in 1956" (1b) both feature a very high degree of certainty, but they are evidentially distinct: In (1a) it is objectively impossible for the speaker to have direct evidence/experience of the event because it happened before she was born, and the direct form is infelicitous. In (1b), the direct form is acceptable as the speaker describes her own experience.

(1) Direct evidential
   a. #Baba   mi   rodi   bašta   mi   prez 1956.
       grandmother   my   gave.birth.3SG.AOR.DIR   father my   in   1956
       'My grandmother gave birth to my father in 1956, <*direct*>.'
   b. Rodih   sina   si   prez   1956
       gave.birth.1SG.AOR.DIR   son.DEF   refl.gen in   1956
       'I gave birth to my son in 1956, <*direct*>.'

The indirect marker signals indirect evidence of a general nature, underspecified for the exact type of indirect source. It is the form that the granddaughter would use instead of the infelicitous (1a), cf. (2a). It is sometimes suggested that the indirect is just an evidentially neutral general past (e.g. Alexander, Zhobov 2009). If this were so, it would be expected to be possible in an indirect context simply because it would be possible in any context. This would predict it to be possible also in direct evidential contexts, which is not the case, as (2b) shows.

---

3   See Friedman 2003; 2004; Fielder 1995; Izvorski 1997; Fitneva 2001; Alexander, Zhobov; 2009; Sonnenhauser 2013; Smirnova 2013; Rivero, Slavkov 2014; Korotkova 2016; Koev 2017; Arregui et al. 2017; Simeonova 2020; Tarpomanova 2021 a.o.

(2) Indirect evidential

    a. Baba   mi   e=rodila   bašta   mi   prez   1956.
         grandmother   my   gave.birth.3SG.AOR.IND   father   my   in   1956
         'My grandmother gave birth to my father in 1956, *<indirect>*.'

    b. #Rodila=sǔm   sina   si   prez   1956.
         gave.birth.1SG.AOR.IND   son.DEF   refl.gen   in   1956
         'I gave birth to my son in 1956, *<indirect>*.'

The reportative evidential marker is specialized for information whose source is specifically a report. It can be uttered, for example, after returning from having run into the neighbors, as a response to a question about what's new with them.

(3) Reportative evidential
     Kotkata   im   rodila.
     cat.DEF   their   gave.birth.3SG.AOR.REP
     'Their cat gave birth, *<reportative>*.'

As is clear from this context above, the reportative does not conventionally convey a reduced level of certainty (see also Izvorski 1997).

     To summarize, Bulgarian has three morphologically and semantically different evidential markers that encode direct, indirect, and reportative information source, and in the past tense, there is no evidentially unmarked form, i.e. a choice between them is obligatory.

2.2   Evidentiality in Turkish

Turkish distinguishes between two evidential forms, direct and indirect, which are marked on finite verbs and morphologically related to tense and aspect (Şener 2011) and are obligatory in the past tense. The direct evidential, cf. (4a), marked by the suffix *–DI*, signals that the speaker personally witnessed the event or has forms of direct evidence through participation. The indirect evidential, as in (4b), formed with the suffix *–(I)mIş*, indicates that the speaker did not directly witness the event but instead obtained the information through report or inference (see Aksu-Koç, Slobin 1986; Slobin, Aksu 1982; Yavaş 1980).

(4) Turkish direct and indirect evidential

    a. Nurhayat   çok   güzel   bir   şarkı   söyle-di   (Arslan 2020, 3)
         Nurhayat   very   beautiful   one   song   say.DIR
         'Nurhayat sang a very beautiful song.' *<direct>*

b. Nurhayat   çok   güzel   bir   şarkı   söyle-miş
   Nurhayat   very   beautiful   one   song   say.$_{IND}$
   'Nurhayat sang a very beautiful song.' *<indirect>*

It is debated whether Turkish evidentials are part of epistemic modality or a distinct grammatical category in their own right (see Aksu-Koç, Arslan 2024 for an overview). Another debate is whether Turkish marks both direct and indirect evidentiality, or whether the direct evidential functions as a neutral past tense marker (Johanson 2003). We follow the perspective that Turkish distinguishes between direct (firsthand) and indirect (non-firsthand) information source as illustrated above.[4]

2.3   Available Comparative Studies on Evidentiality in the Balkans

There are few comparative studies on evidentiality paradigms in Balkan languages, particularly regarding Bulgarian and Turkish. Makartsev (2014) conducted a comparative analysis of how grammatical (indirect) evidential markers appear in different genres of text across Bulgarian, Macedonian and Albanian. He analyzed excerpts from diverse text genres including folkloric texts (songs and ballads) shared across languages in the Balkans, as well as newspaper reports and literary texts. Makartsev concludes that evidential forms go beyond referring to information sources and, in these genres, they build a narrative frame and establish an epistemic author authority.

Bosnali and Kahraman (2015) conducted a study on the use of evidentials in retelling dreams in Japanese and Turkish. They recorded 12 dream narrations from 7 Turkish speakers, which had a total of 234 inflected verbs, of which 147 referred to events that happened within the dream. They counted both –*DI* and the present progressive form –*I(y)or* to be direct evidentials. In our study, we consider only –*DI* to be direct evidential. With that in mind, they report that in 6 of the 12 narrations there were no past forms whatsoever, 66% of the verbs were in the present tense. In the past tense, they report 23 uses of –*(I)mIş* and 30 uses of –*DI* (out of 170 uses).

Regarding Bulgarian and Turkish in particular, one comparative analysis is Grannes' (1996) volume, which delves into outcomes of language contact and briefly contemplates on the semantic relationship between the Bulgarian reportative and the Turkish

---

4   See Aksu-Koç, Slobin 1986; Arslan 2016; Kornfilt 1997; Lewis 1967; Sezer 2001.

indirect evidential. However, Grannes mainly focused on lexical borrowings in Bulgarian from Turkish.

Friedman (1999), to our knowledge, is the first direct comparative analysis of Bulgarian and Turkish specifically (along with more languages) focused on how genre affects the use of evidentiality. He analyzed a corpus of proverbs, arguing that this genre is interesting for evidentiality since it conveys "objective, universal truths" (135). He found that Bulgarian predominantly uses the reportative (117 out of 159 cases), while in Turkish, both direct and indirect evidentials are used (44 and 32 cases, respectively). He concludes that this reveals substantial differences in the evidential systems in the two languages despite their descriptively similar labels such as direct/confirmative or indirect/non-confirmative and links their use to their historical development.

However, proverbs as a genre are limited by their very nature as they are generic, abstract, and timeless. For that reason, proverbs rarely use a past tense to begin with, as Friedman (1999) confirms; the cases he found were among thousands of proverbs, which is a limitation since in both Bulgarian and Turkish, evidentials are only obligatory in the past (see above). Finally, proverbs are not a productive genre: the speaker merely quotes their content.

Our study is based on data from genres that communicate specific past events, both fictional and real, as described next.

2.4   The Present Study

In this study, we analyzed evidential markers in Bulgarian and Turkish across the following seven narrative genres: (i) general knowledge, (ii) police reports and/or minutes, (iii) Biblical/Quranic texts, (iv) modern fiction, (v) news reports, (vi) fables, and (vii) written recounts of dreams. All texts were extracted from contemporary media sources, for the Biblical/Quranic text we included their modern interpretations in Turkish and Bulgarian, and hence this study is purely synchronic.

3   Methods

From each genre in each language, we extracted 20 or more declarative sentences with finite verbs (see the previous section). An anonymized version of the data collected is publicly available at OSF with the following link: https://osf.io/ueynw/.

3.1   Bulgarian Data

The Bulgarian data collected across the seven categories is summarized below in Table 1 and discussed in more detail below. Special or shifted environments like the contents of letters, quotes, and direct speech were excluded from the data collection since they do not represent the main narrative device but are governed by their own internal rules, which are uniform across genres. They were most relevant in the categories Modern Fiction, Fables, and Bible.

Table 1   Details of text data collected for Bulgarian

Genre	Word count	Number of Sentences
General knowledge	934	52
Reports and minute records	469	32
Bible / Quran	314	19
Modern fiction, short story	224	21
News reports	584	27
Fables / folktale	566	158
Recounts of dreams	645	56

For general knowledge category, we collected six text samples ranging over the following different time periods: (i) Cold war (XX c.),[5] (ii) Roman empire (V c.),[6] (iii) Steve Jobs' biography (1955-2011),[7] and following excerpts taken from the Bulgarian Ministry of Education 9th and 10th grades history textbooks, *History and civilization for 9th grade* and *History and civilizations for 10th grade*:[8] (iv) Political systems in Europe (XX c.), (v) First settlements in Bulgaria (10,000-5,000 BCE), and (vi) Middle Ages (X-XVII c.).

For police reports, we used the official website of the Bulgarian Ministry of Internal Affairs,[9] which reports the daily activity of the

---

5   Accessible in electronic format at https://bg.wikipedia.org/wiki/%D0%A1%D1%8 2%D1%83%D0%B4%D0%B5%D0%BD%D0%B0_%D0%B2%D0%BE%D0%B9%D0%BD%D0%B0.

6   Accessible in electronic format at https://bg.wikipedia.org/wiki/%D0%A0%D0% B8%D0%BC%D1%81%D0%BA%D0%B0_%D0%B8%D0%BC%D0%BF%D0%B5%D1%80%D0%B8%D1%8F.

7   Accessible in electronic format at https://bg.wikipedia.org/wiki/%D0%A1%D1%8 2%D0%B8%D0%B2_%D0%94%D0%B6%D0%BE%D0%B1%D1%81.

8   Prosveta Sofia AD, authors Mihail Gruev, Boris Stoyanov, Kiril Slavchev, Yavor Siderov, ISBNs: 978-954-01-4420-7 and 978-954-01-4278-4, available in electronic format at https://www.e-uchebnik.bg/product/istoriya-i-civilizacii-za-9-klas-2#&gid=1&pid=7 and https://www.e-uchebnik.bg/product/istoriya-i-civilizacii-za-10-klas.

9   Accessible in electronic format at https://www.mvr.bg/sdvr/%D0%98%D0%BD%D1%8 4%D0%BE%D1%80%D0%BC%D0%B0%D1%86%D0%B8%D0%BE%D0%BD%D0%B5%D0%BD-%D1%86%D 0%B5%D0%BD%D1%82%D1%8A%D1%80.

police, e.g. criminal actions like robberies, attacks, or accidents such as road accidents, fires, and missing people. The sample contained 469 words in 27 sentences, of which there were 32 finite verbs marking past events and 4 marking real present (e.g. "The investigation *is* still ongoing"). The Bible sample was taken from the Modern Bulgarian translation.[10] We sampled various parts of both the Old and the New Testament, documented individually in the data repository. For modern fiction category, we collected data from various novels in the category Modern Bulgarian novel on the free public library.[11] We sampled the most recent novels, published from 2007 to 2021. We chose works originally written in Bulgarian in order to eliminate any effects of translation. The list of all novels – both excluded and analyzed – is found in the supplementary material in the data repository.

News data were collected from three leading and reputable news websites.[12] Due to the nature of the evidential system described in the introduction, we focused on news about past events and excluded news about future events, such as about upcoming elections, planned events etc. We analyzed 6 news items with a total of 31 sentences. We only analyzed the main text of the news, noting that titles and subtitles form their own sub-genres in journalistic writing, which would need a larger sample to be analyzed separately; this goes beyond the scope of the present paper. We have included these in the raw data for completeness, but not in the analysis. For fables and/or folk tales, we analyzed data from five traditional Bulgarian fables collected and edited by Nikolay Raynov (adapted from Zahari Stoyanov 2005. ISBN 954-739-618-8), with a total of 158 sentences. For dreams, we collected data from a Facebook discussion group dedicated to interpreting people's dreams.[13] Authors post describing their dream and others comment their interpretation. We collected 56 sentences from 10 dream recounts by different authors.

3.2   Turkish Data

For Turkish data, text excerpts are from diverse sources described below, for a total of 189 sentences (2694 words), broken down in Table 2. General knowledge data were excerpted from three sources for a

---

10   Accessible in electronic format at `https://www.bible.com/`.

11   Accessible in electronic format at `chitanka.info`.

12   Accessible in electronic format at `https://novini.bg/`, `https://news.bg/`, and `https://www.vesti.bg/`.

13   Accessible in electronic format at `https://www.facebook.com/groups/sanovnika`.

total of 30 sentences: the Wikipedia entries for Nikia Riots,[14] and for the Halley Comet,[15] and a passage regarding the Renaissance period from the history school textbook published by the Turkish National Ministry for high school grade 11 (*Ortaöğretim Tarih 11 ders kitabı* ISBN 978-975-11-4963-3: MEB Yayınları).

Table 2 Overview of the composition of the Turkish data

Genre	Word count	Number of Sentences
General knowledge / history	472	30
Reports and minute records	429	20
Bible / Quran	473	40
Modern fiction, short story	337	25
News reports	423	23
Fables / folktale	264	22
Recounts of dreams	296	29

Police and official reports data were from the publicly available template reporting forms available online used for police reports for traffic and workplace accidents[16] and meeting minute reports from the Bursa Uludağ University Automotive Working Group meetings available online.[17] This data included a total of 20 sentences.

For religious texts, we extracted 20 sentences from The Quran's Chapter 8/75, *Surah Al-Anfal* 'The Spoils of War' (Diyanet İşleri Başkanlığı, 2011) and 20 sentences from The New Testament Bible's Chapter Luke 8/1–3 (New International Version, 2011).

For modern fiction, we extracted 25 sentences from Sait Faik Abasiyanik's short story *Karanfiller ve Domates Suyu* 'Carnations and Tomato Juice' (Abasiyanik 2021). For news reports, we extracted 23 sentences from two newspaper articles published in Oksijen, a national weekly news magazine in Turkey, including news about global cities index and Istanbul[18] and New York Times report on the aftermath of the 6th February Earthquakes.[19] For folklore tales,

---

14 Accessible in electronic format at https://tr.wikipedia.org/wiki/I._Justinianus#Nika_Ayaklanmas%C4%B1.

15 Accessible in electronic format at https://tr.wikipedia.org/wiki/Halley_kuyruklu_y%C4%B1ld%C4%B1z%C4%B1.

16 Accessible in electronic format at https://onikisubat.meb.gov.tr/www/tutanaklar/icerik/677.

17 Accessible in electronic format at https://uludag.edu.tr/ocg/tutanaklar-66745.

18 Accessible in electronic format at https://gazeteoksijen.com/turkiye/istanbul-kuresel-sehirler-endeksinde-25inci-siraya-yukseldi-194365.

19 Accessible in electronic format at https://gazeteoksijen.com/new-york-times/nyt-depremden-sonrasini-yazdi-bir-ailenin-mucadelesi-202095.

we took 22 sentences from two tales about Nasraddin Hodja, a folk character in Turkish and broader neighboring cultures.[20] And finally, written recounts of dreams comprised of 29 sentences from different sources including social media posts and discussion blogs including the threads under the *KizlarSoruyor* online anonymous blog (n.d.).

3.3   Procedure and Data Analysis

After extracting the data as described above, the authors annotated the evidential markers on every finite/matrix verb. We tallied the total number of direct and indirect evidential forms, as well as other relevant forms, such as the assumption and epistemic morpheme -DIR in Turkish and the historical present in Bulgarian, per genre, and calculated their percentage within the total number of finite verbs inflected for one of the evidential forms. The authors held meetings to compare annotation and data coding, so as to be able to follow a comparable strategy. For statistical analysis, we used Chi-square and Fisher's Exact tests to compare condition differences across each genre in each language.

4   Results

4.1   Results from Bulgarian Data

Table 3 presents a summary of the findings in Bulgarian in absolute counts and percentage.

In three of the seven categories, the results showed that one single form was categorically used. For Modern fiction and the Bible, this is the direct evidential, for Historical facts, it was the historical present, i.e. in that category, no evidential forms were used whatsoever. In Fables, the most dominant marker was reportative. An exact binomial test determined whether the proportion of REP occurrences (86 out of 90 tokens) was significantly greater than chance level ($p =.50$). The results indicated a statistically significant preference for REP over other categories, $p <.001$, 95% CIs [0.90, 1.00], with a sample proportion of 0.96.

---

20 Accessible in electronic format at https://www.milliyet.com.tr/cocuk/eglence/nasrettin-hoca-fikralari-komik-fikralar-denildiginde-akla-gelen-nasreddin-hoca-fikralari-6214835.

Table 3 Summary of results for Bulgarian

	Bible	Modern fiction	Fables	Dreams	Wikipedia/ history textbook	Police report	News texts
words	314	224	566	645	934	469	584
sentences	19	21	158	56	52	27	31
Total number of evidentially marked forms	46	39	90	95	76	32	38
Direct	46 (100%)	39 (100%)	0	84 (88.4%)	0	1 (3.1%)	0
Reportative	0	0	86 (95.5%)	0	0	2 (6.2%)	11 (28.9%)
Indirect	0	0	0	0	0	12 (37.5%)	12 (31.5%)
Hist.pres.	0	0	4 (4.4%)	11 (12.2%)	76 (100%)	17 (53.1%)	15 (39.5%)

For Police reports, a Chi-square test revealed that the four categories were not equally distributed across the compiled text, $\chi^2(3, N = 32) = 31.50$, $p <.001$. The majority of tokens were HIST.PRES (n = 17) and IND (n = 12), while DIR (n = 1) and REP (n = 2) were rare. A Fisher's Exact Test revealed that the counts of IND and HIST.PRES conditions were not statistically different (p =.75, Odds Ratio = 1.40, 95% CIs [0.33, 6.64]). This suggests that police reports used both forms equally likely.

For News, the observed counts were: DIR (0), REP (11), IND (12), and HIST.PRES (15), totaling 38 observations. A Chi-square test was statistically significant ($\chi^2(3, N = 38) = 13.58$, $p =.004$), indicating that the categories were not equally distributed. Specifically, DIR had no occurrences, whereas REP, IND, and HIST.PRES were observed more frequently. Posthoc pairwise comparisons using the Fisher's exact tests with Bonferroni correction revealed that DIR differed significantly from all other conditions of REP, IND, and HIST.PRES frequencies (adjusted ps =.003,.001, and <.001, respectively). There were no significant differences between REP, IND and HIST.PRES (all adjusted $p = 1.00$).

4.2 Results from Turkish data

The frequency and percent occurrences of evidential forms found in the Turkish data is given Table 4. Based on the data extracted from general knowledge (historical and scientific texts), we found high use of the assumption/epistemic marker *–mIs+DIR* 17 (48.6%) and direct evidential *–DI* 11 (31.4%) comparative to indirect evidential *–mIs* and habitual *–DIR* markers. A chi-square goodness-of-fit test was significant $\chi^2(3, 35) = 15.17$, $p =.001$, indicating that occurrences of evidential markers were not equally distributed. We further

compared the counts between the assumption/epistemic and direct evidential markers (17 vs. 11). A Fisher's Exact Test revealed no significant difference (p = 0.22, Odds Ratio = 0.49, 95% CIs = [0.16, 1.43]). Therefore, we found that the general knowledge genre showed dominant uses of assumption/epistemic and direct evidential markers without a difference between the two.

In police/minute reports, generic –*DIR* 2 (13.3%) and –*mIs+DIR* 13 (86.6%) markers emerged only, and no other evidential forms were observed. A chi-square showed that those observed frequencies differed significantly from an equal distribution ($\chi^2(1)$ = 8.07, p = 0.004), indicating that generic assumed/epistemic form is used more often.

Table 4 Frequency and percent evidential forms appraised in different text genres in the Turkish data

	General knowledge	Police / minute reports	Religious texts	Modern fiction	News reports	Folktales	Dreams
Nr. Sentences	30	20	40	25	23	22	29
Total finite verbs	37	22	63	40	29	45	42
Finite verbs with an evidential form	35	15	44	16	25	30	39
Direct evidential -DI	**11 (31.4%)**	0 (0%)	**27 (61.3%)**	**13 (81.2%)**	**25 (100%)**	2 (6.6%)	**35 (89.7%)**
Indirect Evidential -mIs	2 (5.7%)	0 (0%)	2 (4.5%)	0 (0%)	0 (0%)	**28 (93.3%)**	4 (10.3%)
Habitual, Generic -DIR	5 (12.2%)	2 (13.3%)	15 (34.1%)	3 (18.7%)	0 (0%)	0 (0%)	0 (0%)
Assumption, Epistemic -mIs-DIR	17 (48.6%)	13 (86.6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

In religious texts (Bible and Quran), direct evidential –*DI* 27 (61.3%) and habitual/generic –*DIR* form 15 (34.1%) were the most dominant markers. The use of indirect evidentials was merely minimal and restricted in the context where the information was quoted. A chi-square test showed a significant difference in frequencies, $\chi^2(3, 44)$ = 43.09, p <.001. A Fisher's Exact Test showed that difference between the evidential and the generic form was statistically significant (p =.015, Odds Ratio = 3.19, 95% CIs = [1.22, 8.72]). Thus, the direct evidential form was the most dominant in religious contexts. We should note that we counted a number of non-evidential forms, mostly observed under quotes and finite complements, as well, those included aorist (n=7), optative (n=3), imperative (n=6), progressive present (n=3).

In modern fiction, the direct evidential 13 (81.2%) was used more often than the other forms, including generic DIR 3 (18.7%), evidenced by a chi-square test ($\chi^2(1)$ = 6.25, p = 0.01). In fiction, several uses

of the aorist (n=14) and some present progressive forms (n=3) were found as serial verb constructions under a final direct evidential matrix verb.

In news reports the only form in the sample was the direct evidential. Folktale data showed the reverse, exclusively using indirect evidential form 28 of 30 finite verbs analysed (93.3%). The direct evidential forms found here (n=2) were in fact used on quoted complements.

In dreams, direct evidential –DI was used in 35 out of 39 analyzed finite verbs (89.7%); the other 4 were the indirect evidential. The frequency of these observations was significantly different from being equally distributed ($\chi^2(1) = 24.64$, $p < 0.001$). The indirect evidential –(I)mIş was used in context of mirativity or lack of consciousness. This is in contrast with Bosnali and Kahraman (2015), who found that about half of the dream narrations they collected were narrated entirely in the present tense (see above).

5　Discussion

Our aim was to investigate how Bulgarian and Turkish use evidentiality across seven text genres: general knowledge (textbooks, Wikipedia), police reports (and minutes), Biblical/Quranic texts, modern fiction, news reports, fables, and recounts of dreams. We found genre-determined evidential use in most categories. Table 5 summarizes the most dominant form that emerged per genre.

Table 5　Dominant evidential form emerged across different genres

Genre	Bulgarian	Turkish
Bible and/or Quran	Direct	Direct
Modern Fiction	Direct	Direct
Fables, Folktales	Reportative	Indirect
Dreams	Direct	Direct
General knowledge	Historical Present	Assumption / Direct
Police reports	Indirect / Historical Present	Assumption / Generic
News	Indirect / Reportative / Historical Present	Direct

In both languages, the direct evidential was predominantly used in the modern fiction, religious, and dream categories, while the indirect (reportative in Bulgarian) was used in traditional fables. It is important to note that fables and religious texts are somewhat archaic, they do not necessarily point towards synchronous use of evidential forms but were important in understanding how direct

evidential in religious contexts, and indirect evidential in fables are traditionally used. The two languages differed strikingly in the News category, where Turkish used the direct evidential form virtually exclusively, while Bulgarian showed a relatively even distribution among all forms except the direct, namely: reportative, indirect, and historical present. This might be because epistemic uncertainty associates with Turkish indirect evidential (see Arslan 2020) possibly more strongly than in Bulgarian, and as a result, direct evidential is preferred to maintain an authority that the information provided is certain or truthful, while in Bulgarian this might not be necessary.

In the general knowledge category, Bulgarian uses solely historical present and does not use any evidentials whatsoever, while Turkish uses assumption/epistemic *–mIş+DIR* (48.6%) but also the direct evidential (31.4%). The two languages also differed in the reports category, where Bulgarian used both the indirect evidential (37.5%) and the historical present (53.1%), while Turkish used the epistemic form *–mIş+DIR*.

Several patterns emerge from these results. One is the strict ban of the direct evidential in Bulgarian in realistic genres like historical/science facts and news, while the Turkish direct evidential was used as the main form in news and about a third of the historical category. This suggests a deeper semantic difference between the two direct markers despite their similar use in prototypical conversational scenarios.

The second pattern is a competition of non-evidential forms with evidential ones, such as the Bulgarian historical present form and the Turkish generic assumption marker *–mIş+DIR*): both languages used these forms in genres that by design represent information as factual and neutral, such as textbooks/general knowledge and police reports (though notably not in news). However, a limitation must be acknowledged here, namely the close relationship between evidential markers and modal meanings in both languages, particularly with respect to mirativity and epistemic meanings. In both the languages, indirect evidential forms can be used in mirative contexts or used to indicate uncertain information; however, based on written genres, analyzing such contexts pose a limitation. In certain genres, such as news reports, especially in Turkish, information seems to have been manipulated in a way that direct evidentials were exclusively used to remove any forms of epistemic uncertainty, an observation that fits in with Makartsev (2014). This observation merits more detailed exploration in future work.

6 Conclusions

This study shows that evidentials are restricted to different narrative modes with strikingly little optionality. We find that Bulgarian and Turkish pattern identically in four of the seven environments. The biggest differences were (i) in the use of the direct evidential in news and general knowledge in Turkish that was not attested at all in Bulgarian; and (ii) the lack of the indirect in Turkish in the report category compared to Bulgarian, where they were used on a par with the historical present.

To conclude, our findings highlight the need for more nuanced accounts of the relationship between evidentiality, epistemic modality, and discourse genre, as well as the competition between evidential and other narrative forms, such as the historical present and the assumption marker.

Acknowledgments

Bibliography

Abasıyanık, S.F. (2021). *Karanfiller ve domates suyu*. İstanbul: İş Bankası Kültür Yayınları.

Aikhenvald, A.Y. (2004). *Evidentiality*. Oxford, UK: Oxford University Press.

Aksu-Koç, A.; Slobin, D.I. (1986). "A Psychological Account of the Development and Use of Evidentials in Turkish". Wallace, C.; Nichols, J. (eds), *Evidentiality. The Linguistic Coding of Epistemiology*, 159-67. Norwood, NJ: Ablex Publishing.

Aksu-Koç, A.; Arslan, S. (2024). "Tense, Aspect, Modality and Evidentiality". Johanson, L. (eds), *The Encyclopedia of Turkic Languages and Linguistics Online, Section Psycholinguistics and Turkish*, 1-12. Leiden: Brill.

Alexander, R.; Zhobov, V. (2009). "New Conclusions on the Conclusive". *Journal of Slavic Linguistics*, 17(1-2), 61-85.

Arregui, A.; Rivero, M.L.; Salanova, A.P. (2017). "Aspect and Tense in Evidentials". Arregui, A.; Rivero, M.L.; Salanova, A.P. (eds), *Modality Across Syntactic Categories*, 211-34. Oxford: Oxford University Press.

Arslan, S. (2020). "When the Owner of Information is Unsure: Epistemic Uncertainty Influences Evidentiality Processing in Turkish". *Lingua*, 247, 102989.

Arslan, S. (2016). *Neurolinguistic and Psycholinguistic Investigations on Evidentiality in Turkish*. Groningen: University of Groningen.

Bosnalı, S.; Kahraman, C. (2015). "Moods and Modality in the Turkish and Japanese Dream Narration". *LILA´15 / Linguistics and Language Conference Proceedings*, 169-85. İstanbul: DAKAM.

Chafe, W.L.; Nichols, J. (1986). *Evidentiality: The Linguistic Coding of Epistemology.* Norwood, NJ: Ablex Publishing.

Fielder, G.E. (1995). "Narrative Perspective and the Bulgarian l-Participle". *The Slavic and East European Journal*, 39(4), 585-600.

Fitneva, S.A. (2001). "Epistemic Marking and Reliability Judgments: Evidence from Bulgarian". *Journal of Pragmatics*, 33(3), 401-20.

Friedman, V.A. (1999). "Proverbial Evidentiality: On the Gnomic Uses of the Category of Status in Languages of the Balkans and the Caucasus". *Mediterranean Language Review*, 11, 135-55.

Friedman, V.A. (2003). *Turkish in Macedonia and Beyond: Studies in Contact, Typology, and Other Phenomena in the Balkans and the Caucasus*. Wiesbaden: Harrassowitz Verlag.

Friedman, V.A. (2004). "Evidentiality, Modality, and Narrative in Macedonian and Other Balkan Languages". Friedman, V.A. (ed), *Macedonian Studies: Papers from the Fifth International Macedonian–North American Conference on Macedonian Studies*, 85-102. Columbus, OH: The Ohio State University.

Friedman, V.A. (2008). "The Typology of Balkan Evidentiality and Areal Linguistics". Tomić, O.M. (ed), *Balkan Syntax and Semantics*, 101-34. Amsterdam: John Benjamins.

Givón, T. (1982). "Logic vs. Pragmatics, with Human Language as the Referee: Toward an Empirically Viable Epistemology". *Journal of Pragmatics*, 6(2), 81-133.

Grannes, A. (1996). *Turco-Bulgarica: Articles in English and French Concerning Turkish Influence on Bulgarian*. Wiesbaden: Otto Harrassowitz Verlag.

Izvorski, R. (1997). "The Present Perfect as an Epistemic Modal". Lawson, A.; Cho, E. (eds), *Proceedings of SALT 7*, 222-39. Ithaca, NY: CLC Publications.

Johanson, L.; Utas, B. (2000). *Evidentials: Turkic, Iranian and Neighbouring Languages*. Berlin: Walter de Gruyter.

Johanson, L. (2003). "Evidentiality in Turkic". Aikhenvald, A.Y.; Dixon, R.M.W. (eds), *Studies in Evidentiality*, 273-90. Amsterdam: John Benjamins.

Koev, T. (2017). "Evidentiality, Learning Events and Spatiotemporal Distance: The View from Bulgarian". *Journal of Semantics*, 34(1), 1-41.

Kornfilt, J. (1997). *Turkish*. London: Routledge.

Korotkova, N. (2016). *Heterogeneity and Uniformity in the Evidential Domain* [Doctoral Dissertation]. Los Angeles, CA: University of California.

Korotkova, N. (2020). "The Subjective Heart of Evidentiality". Paper presented at *GLOW 43: Semantic Universals in the Modal and Attitudinal Domain*, Humboldt University & ZAS, Berlin, Germany.

Lewis, G.L. (1967). *Turkish Grammar*. Oxford: Clarendon Press.

Makartsev, M.M. (2014). *Эвиденциальность в пространстве балканского текста* (Evidentiality in the Balkan text space). Moscow: Russian Academy of Sciences.

Palmer, F.R. (2001). *Mood and modality*. Cambridge: Cambridge University Press.

Plungian, V.A. (2001). "The Place of Evidentiality within the Universal Grammatical Space". *Journal of Pragmatics*, 33(3), 349-57.

Şener, N. (2011). *Semantics and Pragmatics of Evidentials in Turkish* [Doctoral Dissertation]. Storrs, CT: University of Connecticut.

Sezer, E. (2001). "Finite Inflection in Turkish". Erguvanlı-Taylan, E. (ed.), *The Verb in Turkish*, 1-46. Amsterdam: John Benjamins.

Simeonova, V. (2020). *The Syntax and semantics of light attitudes* [Doctoral Dissertation]. Ottawa: University of Ottawa.

Simeonova, V. (2021). "Challenging Evidential Non-Challengeability". Paper presented at SemDial 25, University of Tübingen, Germany.

Slobin, D.I.; Aksu, A.A. (1982). "Tense, Aspect and Modality in the Use of the Turkish Evidential". P.J. Hopper (ed.), *Tense-Aspect: Between semantics and pragmatics,* 185-200. Amsterdam: John Benjamins.

Smirnova, A. (2011). *Evidentiality and Mood: Grammatical Expressions of Epistemic Modality in Bulgarian* [Doctoral Dissertation]. Columbus, OH: The Ohio State University.

Smirnova, A. (2013a). "Evidentiality in Bulgarian: Temporality, epistemic modality, and information source". *Journal of Semantics*, 30(4), 479-532.

Smirnova, A. (2013b). "The Meaning of the Bulgarian and Turkish evidentials". *Съпоставително езикознание / Сопоставительное языкознание*, 2-3, 205-23.

Smirnova, A. (2021). "Evidentiality in abductive reasoning: Experimental support for a modal analysis of evidentials". *Journal of Semantics*, 38(4), 531-70.

Sonnenhauser, B. (2013). "'Evidentiality' and point of view in Bulgarian." *Съпоставително езикознание / Сопоставительное языкознание*, 2-3, 110-30.

Tarpomanova, E. (2021). "Semantic readings of the indirect evidentials in Bulgarian: knowledge vs. reliability." *Papers of BAS. Humanities and Social Sciences*, 8(2), 125-33.

Willett, T. (1988). "A Cross-Linguistic Survey of the Grammaticization of Evidentiality". *Studies in Language*, 12(1), 51-97.

Yavaş, F. (1980). *On the Meaning of the Tense and Aspect Markers in Turkish* [Doctoral Dissertation]. Lawrence, KS: University of Kansas.