

MORPHOSYNTACTIC STRATEGIES FOR EXPRESSING EVIDENTIALITY IN UZBEK AND ENGLISH: A CONTRASTIVE ANALYSIS

Malika Ergasheva

*Head of the Department of Social Sciences
and Foreign Languages, Chirchiq Branch
of Tashkent State Medical University
E-mail: dr.malikaprof@gmail.com*

Abstract. This study presents a contrastive analysis of the morphosyntactic strategies employed for expressing evidentiality in Uzbek and English. Evidentiality—the linguistic encoding of information source and the speaker’s basis for a claim—is realized through fundamentally different mechanisms in these typologically distant languages. Uzbek, as an agglutinative Turkic language, possesses a grammaticalized evidential system marked by dedicated verbal suffixes that obligatorily distinguish direct evidence from hearsay and inference. English, as an analytic Indo-European language, lacks grammaticalized evidentiality but expresses equivalent semantic distinctions through a range of lexical and syntactic strategies including modal verbs, adverbials, complement clause constructions, and reported speech patterns. Drawing on a parallel corpus of 500 sentence pairs extracted from literary, journalistic, and academic texts, this study identifies and classifies the evidential strategies of both languages, maps their functional correspondences, and analyzes the translation challenges that arise from the structural asymmetry between a grammatical and a lexical evidential system. Results demonstrate that Uzbek employs four primary morphological evidential markers, while English distributes equivalent functions across at least twelve distinct lexicosyntactic constructions, creating systematic non-equivalence in translation. The findings contribute to contrastive Turkic–Germanic linguistics and carry practical implications for translation studies, second language pedagogy, and the typological understanding of evidentiality across language families.

Keywords: evidentiality, Uzbek language, English language, contrastive analysis, morphosyntax, grammaticalization, information source, Turkic linguistics, modality

INTRODUCTION

Evidentiality, defined as the linguistic marking of the source of information upon which a statement is based, has emerged as one of the most intensively studied domains of functional grammar over the past four decades. Since Chafe and Nichols’s (1986:vii) seminal volume established evidentiality as a cross-linguistic category of fundamental

importance, research has demonstrated that approximately one quarter of the world's languages possess dedicated grammatical systems for encoding information source, while the remainder express similar semantic distinctions through lexical or syntactic means (Aikhenvald, 2004:1). The typological contrast between languages with grammaticalized evidentiality and those without it constitutes one of the most significant structural differences across language families, with far-reaching implications for translation, language pedagogy, and cognitive linguistics.

Uzbek, a member of the Karluk branch of the Turkic language family, is characterized by a well-developed grammaticalized evidential system. Through a set of dedicated verbal suffixes, Uzbek speakers obligatorily mark whether their statements are based on direct personal evidence, reported information (hearsay), or inference from observable results (Johanson, 2003:273). This system is not optional or stylistic; it is a structural requirement of the grammar, and the omission of evidential marking in contexts where it is expected produces ungrammatical or pragmatically anomalous utterances (Csató & Johanson, 1998:32).

English, by contrast, lacks a dedicated grammatical category of evidentiality. The language does not possess morphological markers that obligatorily encode information source; instead, speakers express evidential meanings through a diverse array of lexical and syntactic strategies, including modal verbs (must, may, might), evidential adverbials (apparently, reportedly, evidently), complement clause constructions (It seems that, It appears that), and reported speech patterns (According to X, X is said to) (Mushin, 2001:68). These strategies are semantically rich but grammatically optional: an English speaker can make an assertion without specifying its evidential basis, an option that is structurally unavailable to a Uzbek speaker.

This fundamental asymmetry between a language with obligatory grammatical evidentiality and a language with optional lexical evidentiality creates significant challenges for translation, second language acquisition, and cross-linguistic communication. Despite the theoretical importance of this contrast, systematic contrastive studies of evidentiality in Uzbek and English remain scarce. The present study addresses this gap through a corpus-based contrastive analysis that maps the evidential strategies of both languages and examines their functional correspondences and translation implications.

The theoretical foundation of this study integrates three complementary frameworks. The first is Aikhenvald's (2004:1) typological classification of evidential systems, which distinguishes between languages with small systems (two-term: direct vs. indirect), medium systems (three-term: direct, reported, inferential), and large systems (four or more distinctions). Within this typology, Uzbek possesses a medium-to-large evidential system, distinguishing at minimum among direct evidence, hearsay (reportive), and inference (Johanson, 2003:275).

The second framework is DeLancey's (2001:369) theory of mirativity and its relationship to evidentiality, which proposes that evidential markers in many languages have extended their functions to encode the speaker's expectations and surprise, a phenomenon directly observable in the Uzbek evidential suffix -ibdi, which conveys both inferential evidence and unexpected information. The third framework is the functional-typological approach to modality-evidentiality interaction (Palmer, 2001:8), which provides the analytical tools for examining how evidential and epistemic modal meanings overlap and interact in languages with different structural resources.

Research on Turkic evidentiality has a substantial history. Johanson (2003:273) provided the most comprehensive typological overview of evidential strategies across Turkic languages, identifying the indirective (also termed the inferential or perfect of surprise) as the core evidential category shared across the family. Csató and Johanson (1998:32) examined the grammaticalization pathways through which Turkic evidential markers have developed from participial and perfect tense forms, demonstrating that the formal similarity of evidential markers across geographically dispersed Turkic languages reflects a shared inheritance rather than contact-induced convergence.

For Uzbek specifically, Sjöberg (1963:145) provided the earliest systematic description of evidential morphology in the language, identifying the past tense distinction between -di (witnessed/direct) and -gan (non-witnessed/indirect) as the primary evidential opposition. More recently, Bodrogligeti (2003:184) and Raun (1969:56) have described additional evidential distinctions encoded by the forms -ibdi (inferential/mirative) and -mish (reportive/hearsay), suggesting that Uzbek possesses a four-term evidential system richer than many closely related Turkic languages.

On the English side, Mushin (2001:68) provided a comprehensive analysis of evidential strategies in English narrative discourse, while Chafe (1986:262) examined the lexical expression of evidentiality in English academic writing. De Haan (1999:84) investigated the relationship between evidentiality and epistemic modality in English, arguing that while the two categories overlap functionally, they are conceptually distinct: evidentiality concerns the source of information, while epistemic modality concerns the speaker's degree of certainty.

METHODS

A purpose-built parallel corpus of 500 sentence pairs was compiled from three genre categories: literary fiction (200 pairs from Uzbek novels and their published English translations), journalistic writing (150 pairs from bilingual news sources), and academic prose (150 pairs from bilingual scholarly publications). The corpus was designed to ensure balanced representation of evidential contexts across genres, as the frequency and type of evidential marking varies significantly between narrative, reportive, and analytical discourse (Mushin, 2001:72). All Uzbek source sentences were selected on the basis of containing at least one identifiable evidential marker; the

corresponding English translations or equivalents were then extracted and analyzed for the strategies employed to convey equivalent evidential meaning.

2.2 Analytical Procedure

Each sentence pair was analyzed in three stages. First, the Uzbek sentence was coded for evidential morphology, identifying the specific evidential marker(s) present and their evidential value (direct, reportive, inferential, or mirative). Second, the English equivalent was coded for the lexicosyntactic strategy employed to convey the corresponding evidential meaning, using a taxonomy developed from Mushin (2001:68) and Chafe (1986:262). Third, the functional correspondence between the Uzbek morphological marker and the English lexicosyntactic strategy was assessed on a three-point scale: full equivalence (evidential meaning preserved), partial equivalence (evidential meaning modified or attenuated), and zero equivalence (evidential meaning lost in translation). Inter-rater reliability for the coding procedure was established with a second trained linguist coding a random 20% sample (Cohen's $\kappa = 0.86$).

RESULTS AND DISCUSSION

The corpus analysis confirmed that Uzbek employs four primary morphological strategies for encoding evidentiality, organized around the fundamental opposition between witnessed (direct) and non-witnessed (indirect) information sources.

Table 1. Uzbek Evidential Morphological Markers

Marker	Ev. Value	Semantic Function	Corpus Freq.	Example
-di	Direct	Speaker witnessed the event	38.4%	U keldi ('He came' [I saw])
-gan	Indirect	Non-witnessed; general knowledge	29.2%	U kelgan ('He came' [I heard])
-ibdi	Inferential	Inference from results; mirative	18.6%	U kelibdi ('He came' [I infer])
-(i)mish	Reportive	Hearsay; reported speech	13.8%	U kelmish ('He came' [they say])

Note. Corpus frequency based on 500 evidentially marked Uzbek sentences.

Ev. = Evidential.

The direct evidential marker -di was the most frequent in the corpus (38.4%), reflecting the communicative prominence of first-hand experience as an information source. The indirect marker -gan (29.2%) was particularly prevalent in narrative and academic discourse, where speakers frequently reference events outside their direct experience. The inferential/mirative marker -ibdi (18.6%) appeared most frequently in literary and conversational genres, often conveying the speaker's surprise or discovery. The reportive marker -(i)mish (13.8%) was concentrated in journalistic and folklore texts, consistent with its hearsay function (Johanson, 2003:278).

Table 2. English Lexicosyntactic Evidential Strategies

Strategy	Ev. Function	Corpus Freq.	Example
Modal verbs	Inferential	24.8%	He must have come
Evidential adverbs	Various	19.6%	Apparently / reportedly
Complement clauses	Inferential	16.2%	It seems/appears that...
Reported speech	Reportive	15.4%	According to X; X said
Sensory verbs	Direct	10.8%	I saw / heard / noticed
Zero marking	Direct (default)	13.2%	He came (no ev. marker)

Note. Corpus frequency based on 500 English equivalents of Uzbek evidentially marked sentences.

The analysis revealed that English distributes evidential functions across at least six major strategy types, with modal verbs being the most frequent (24.8%). Notably, 13.2% of the English equivalents employed zero evidential marking—bare assertions with no overt indication of information source—a strategy that is structurally unavailable in Uzbek, where every past-tense assertion obligatorily encodes its evidential basis. This finding quantifies the well-known typological observation that English treats evidential specification as pragmatically optional, while Uzbek treats it as grammatically obligatory (Aikhenvald, 2004:6).

The analysis of functional correspondence between Uzbek evidential markers and their English equivalents revealed significant asymmetries. Full functional

equivalence—where the English translation preserved the evidential meaning of the Uzbek original—was achieved in only 47.6% of sentence pairs. Partial equivalence, where evidential meaning was modified or attenuated, was observed in 33.8% of cases, while evidential meaning was entirely lost in 18.6% of translations. The loss was most pronounced with the mirative function of *-ibdi* (34.2% zero equivalence), for which English possesses no conventionalized single-word equivalent, and translators frequently resorted to extended paraphrases or simply omitted the mirative nuance (DeLancey, 2001:375).

The direct evidential *-di* presented the fewest translation difficulties, as English unmarked assertions default to a direct evidence interpretation—the assumption that the speaker has first-hand knowledge unless otherwise indicated. The reportive *-(i)mish* mapped relatively straightforwardly onto English reported speech constructions and the adverb *reportedly*. The indirect *-gan* and inferential *-ibdi*, however, created the greatest translation challenges, as their meanings distributed across multiple English constructions (modal verbs, complement clauses, and adverbials), with no single English form capturing the full semantic range of either marker (Csató & Johanson, 1998:38).

The systematic non-equivalence identified in this study has direct implications for Uzbek–English language pedagogy. Uzbek-speaking learners of English must learn to express obligatory grammatical distinctions through optional lexical means, a transition that requires not merely vocabulary expansion but a fundamental reconceptualization of the relationship between information source and linguistic form. Research on the acquisition of evidential systems by second language learners suggests that this reconceptualization is particularly challenging when the learner's L1 and L2 differ in whether evidentiality is grammaticalized or lexicalized, as the structural habits formed in the L1 create strong transfer effects that persist even at advanced proficiency levels (Aikhenvald, 2004:382).

Conversely, English-speaking learners of Uzbek must develop an automatic sensitivity to evidential distinctions that their native language allows them to leave unspecified. The challenge is not merely learning new morphological forms but acquiring the underlying cognitive habit of attending to information source as a mandatory dimension of every past-tense assertion. Pedagogical materials for Uzbek–English language instruction should explicitly address the evidentiality asymmetry, providing systematic training in the recognition and production of evidential marking across both languages (Palmer, 2001:35).

Specifically, contrastive exercises that present the same propositional content with different evidential framings—for example, requiring learners to translate a single Uzbek sentence with *-di*, *-gan*, *-ibdi*, and *-(i)mish* into four distinct English renderings—can develop metalinguistic awareness of the evidential domain and improve both

translation accuracy and spontaneous production. Such exercises align with the broader pedagogical principle that explicit attention to typological differences between L1 and L2 accelerates acquisition and reduces fossilization of transfer errors (De Haan, 1999:95).

CONCLUSION

This contrastive analysis has demonstrated that evidentiality constitutes a domain of fundamental structural divergence between Uzbek and English. Uzbek encodes evidential distinctions through a four-term morphological system that obligatorily marks every past-tense assertion for its information source: direct witness (-di), non-witnessed knowledge (-gan), inference from results (-ibdi), and hearsay report (-(i)mish). English expresses functionally comparable meanings through an open-ended set of lexicosyntactic strategies—modal verbs, adverbials, complement clauses, reported speech constructions, and sensory verb phrases—none of which is grammatically obligatory.

The corpus-based analysis revealed that this structural asymmetry creates systematic translation challenges, with full evidential equivalence achieved in fewer than half of the analyzed sentence pairs. The mirative function of the Uzbek inferential marker -ibdi proves particularly resistant to English translation, as English lacks a conventionalized grammatical means of encoding the speaker's surprise at newly discovered information. These findings contribute to the typological understanding of evidentiality as a cross-linguistic category and provide empirical evidence for the degree of semantic loss that occurs when evidential meanings are transferred between a grammaticalizing and a lexicalizing language.

The practical implications of this study extend to several applied domains. In translation studies, awareness of the four-to-twelve mapping between Uzbek morphological markers and English lexicosyntactic strategies can improve translation quality and reduce information loss, particularly in literary translation where evidential nuances carry significant narrative and characterization functions. The finding that mirative meaning is most vulnerable to translation loss highlights the need for translators to develop explicit strategies for rendering this semantically rich but cross-linguistically rare category.

In second language pedagogy, the contrastive framework developed in this study provides a principled basis for designing instructional materials that address the evidentiality gap between Uzbek and English. Explicit instruction in cross-linguistic evidential differences, supported by contrastive translation exercises, can accelerate the acquisition of target-language evidential conventions and reduce transfer-induced errors that otherwise persist at advanced proficiency levels.

In computational linguistics and natural language processing, the detailed mapping of evidential correspondences between Uzbek and English presented in this

study can inform the development of machine translation systems capable of handling this semantically rich and typologically variable domain of human language. The identification of systematic patterns of equivalence and non-equivalence provides training data for computational models that aim to preserve evidential meaning across languages with fundamentally different encoding strategies. Future research should extend this contrastive framework to other Turkic–Indo-European language pairs and investigate the acquisition of cross-linguistic evidential competence through longitudinal learner corpus studies.

REFERENCES

1. Aikhenvald, A. Y. (2004). Evidentiality. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198299783.001.0001>
2. Bodrogligeti, A. J. E. (2003). An academic reference grammar of modern literary Uzbek. Lincom Europa.
3. Chafe, W. (1986). Evidentiality in English conversation and academic writing. In W. Chafe & J. Nichols (Eds.), *Evidentiality: The linguistic coding of epistemology* (pp. 261–272). Ablex.
4. Chafe, W., & Nichols, J. (Eds.). (1986). *Evidentiality: The linguistic coding of epistemology*. Ablex.
5. Csató, É. Á., & Johanson, L. (1998). Turkish. In L. Johanson & É. Á. Csató (Eds.), *The Turkic languages* (pp. 30–42). Routledge.
6. De Haan, F. (1999). Evidentiality and epistemic modality: Setting boundaries. *Southwest Journal of Linguistics*, 18(1), 83–101.
7. DeLancey, S. (2001). The mirative and evidentiality. *Journal of Pragmatics*, 33(3), 369–382. [https://doi.org/10.1016/S0378-2166\(01\)80001-1](https://doi.org/10.1016/S0378-2166(01)80001-1)
8. Johanson, L. (2003). Evidentiality in Turkic. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), *Studies in evidentiality* (pp. 273–290). John Benjamins. <https://doi.org/10.1075/tsl.54.15joh>
9. Mushin, I. (2001). Evidentiality and epistemological stance: Narrative retelling. John Benjamins. <https://doi.org/10.1075/pbns.87>
10. Palmer, F. R. (2001). *Mood and modality* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139167178>
11. Raun, A. (1969). *Basic course in Uzbek*. Indiana University Publications.
12. Sjoberg, A. F. (1963). *Uzbek structural grammar*. Indiana University Publications.