

National Research University Higher  
School of Economics

as a manuscript

Jannigje Helena Verhees

**Evidentiality as part of tense-aspect  
in East Caucasian languages**

Dissertation Summary  
for the purpose of obtaining  
academic degree Doctor of Philosophy in  
Philology and Linguistics HSE

Academic supervisor:  
Michael Daniel, Candidate of Sciences

Moscow 2019

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Summary</b>	<b>7</b>
2.1	Chapter 1 . . . . .	7
2.2	Chapter 2 . . . . .	8
2.3	Chapter 3 . . . . .	13
<b>3</b>	<b>References</b>	<b>18</b>

## 1 Introduction

All East Caucasian languages feature a verb form similar to the English Present Perfect. A prototypical perfect refers to an event completed in the past, and the completion of which has some relevance at the moment of speech, or **CURRENT RELEVANCE** [Comrie 1985: 24-25]. Example 1 from Bagvalal, an East Caucasian language of the Andic group, demonstrates typical current relevance (CR) usage [Tatevosov 2007: 359].

- (1) *iš:i-r            ha:-b-saʔatł:i-r    q'ani-r    q'ani-b-o    ek'wa*  
we.EXCL-ERG this-N-hour-ERG eat-MSD eat-N-CVB AUX.PRS  
'We have [just] dined.'

The sentence in (1) does not simply convey the fact of completion, but emphasizes the completion of the event and its implicit consequences. As described by Tatevosov [2007], it would be appropriate to utter this sentence, for example when speaker X tells speaker Y, that speaker Y's brother should come over for dinner, and speaker Y answers that his brother probably will not come. Speaker Y can use the sentence in (1) as a justification, implying that his brother might still be full from dinner. The same sentence with a general past (in Bagvalal: the Preterite) instead of a perfect would simply convey that the event occurred.

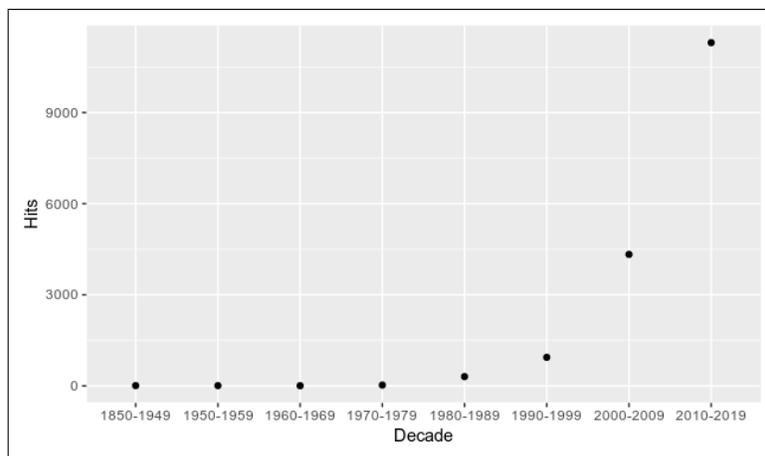
In East Caucasian languages, these types of forms commonly obtain an additional meaning, indicating that the speaker did not witness the event they are talking about, cf. example (2), also from Bagvalal [Tatevosov 2007: 360].

- (2) *ʔali-r    sĩ:    k'wa-b-o    ek'wa*  
Ali-ERG bear kill-N-CVB AUX.PRS  
'Ali [I heard/see,] killed a bear.'

The Perfect in (2) can refer to information from either **INFERENCE** or hearsay, i.e. **REPORTATIVE**. Inference refers to situations where the speaker witnessed a result or consequence from the event, from which they infer that the event took place. Reportative evidentiality indicates that the information is based on a verbal report, without mentioning the exact source of the report, following Aikhenvald [2004]. These two meanings form specific, context-dependent interpretations of the more general meaning of **INDIRECT EVIDENTIALITY**. **EVIDENTIALITY** is defined as the grammatical marking of information source [Aikhenvald 2018: 1]. As a typological category, the concept is relatively new, and especially the last two decades have seen a surge of interest in the topic, which determines the **relevance** of the present investigation. Figure 1 shows

the number of publications for the keyword “evidentiality” in the period 1850-1949, and subsequently for each decade starting from 1950.

Figure 1: Publications for the keyword “evidentiality” per decade



Only 11 publications appeared in the first 100-year period, and none of them dealt with evidentiality as a linguistic phenomenon.<sup>1</sup> This is not surprising, since the term EVIDENTIAL was coined only in 1957 [Jakobson 1957]. A significant rise becomes noticeable starting from the period 1980-1989 (+307), perhaps not coincidentally in the same period when the influential volume *Evidentiality: the linguistic encoding of epistemology* was published [Chafe and Nichols 1986]. This number tripled over the following decade (+941). In the period 2010-2019, a grand total of 11,300 publications were retrieved when the search query was carried out (in April, 2019). Undoubtedly, the number has since grown even further.

Over the years, a vast body of literature has amassed, describing the specifics of the category in a variety of languages. These data led to the insight that languages spoken in adjacent areas tend to have similar evidential systems. Plungian [2010] distinguishes at least five different **evidential areas**. The East Caucasian languages are spoken in the center of one of these areas, which stretches from the Balkans in Eastern Europe, through the Caucasus, Central Asia and the Volga area into Siberia and East Asia, an area sometimes referred to as the “(Great / Old World) Evidential Belt”. As mentioned in [Chirikba 2003]: “There is a tendency in the Balkans, Caucasus and Central Asia to hold the Turkic languages responsible for the development of the evidential category.”

<sup>1</sup>The term evidentiality is also used in e.g. philosophy and psychology.

The Turkic evidential system is present to some degree in all contemporary languages of the family [Johanson 2018], and is present in the oldest written sources [Erdal 2004]. In addition, the languages of the family are dispersed across the area in question.

The evidential systems found in the Eurasian evidential belt are relatively small, see [Aikhenvald 2004]. They predominantly mark unwitnessed events as opposed to unmarked or witnessed events. At the base of these systems are “perfectoid” forms like the Bagvalal Perfect. The term PERFECTOID was first employed in Plungian [2016], in reference to resultative and perfect-like forms from the area described above. These forms bear some similarity to prototypical perfects, but current relevance is usually not their most frequent or prominent meaning. Since the term perfect is often associated with current relevance, perfectoid is a convenient term to describe forms which are similar to some degree, but display significant differences at the same time.

Another feature specific of perfectoid evidentiality is that many authors prefer not to call it evidentiality at all. A number of alternative terms have been proposed over the years, such as indirective, mediative or non-confirmative. According to the authors, situations where the speaker did not witness the event they are talking about is a frequent context for the usage of these forms, while their MAIN MEANING pertains to a more abstract category, marking the speaker’s **distancing** from the event, or their unwillingness to **vouch** for the truth of the conveyed information. These generalized labels subsume two more categories of knowledge: MIRATIVITY, marking new or surprising information and EPISTEMIC MODALITY, which indicates the speaker’s degree of justification. Both can be considered independent categories, although they often accompany indirect evidential semantics. Some authors prefer to maintain these distinct categories, designating them as semantic EXTENSIONS of indirect evidentiality, following [Aikhenvald 2004].

Existing descriptions of individual languages suggest that the configuration of this complex polysemy varies across languages - the use of an indirect evidential perfectoid may carry a strong epistemic modal connotation in one language, while the use of a similar form in another language might be consistent with an evidential analysis. For some languages, contradictory accounts exist.

**The goal** of the present study is to determine the status of the indirect evidential meaning of perfectoid forms in East Caucasian. In addition, the likelihood of the Turkic origin hypothesis for these particular languages will be assessed.

This goal implies the following three tasks:

- Comparison of the formal features of East Caucasian perfectoids, as well as the distribution of these features across the family and on the map.

- Comparison of the East Caucasian material to the forms represented in neighboring Turkic languages.
- Comparison of the use of perfectoids across different languages

The first and second task require the development of formal criteria to identify and categorize perfectoids. Subsequently, similar forms in the languages of the East Caucasian region can be inventorized and compared. To resolve the third task we propose to conduct a comparative analysis of how perfectoids and other tense forms are used in narrative sequences. NARRATIVE SEQUENCES consist of at least two (finite) clauses [Paducheva 2010], and describe consecutive events. The linear order of the clauses reflects the temporal order in which they occurred. A limited number of forms can be employed for this purpose [Labov and Waletzky 1967], and prototypical perfects are excluded by definition, cf. [Lindstedt 2000]. We expect that speakers of languages in which the evidential meaning is more grammaticalized use the indirect evidential more consistently (= more frequently) to narrate unwitnessed events.

As mentioned in [Forker 2018], no systematic comparative research of evidentiality in East Caucasian languages has been carried out, which determines the **theoretical significance** of the present study.

The main theses put forth in this study include the following:

- The “main meaning” criterion proposed in Aikhenvald [2004] to separate grammatical evidentials from evidential strategies is insufficient to make an informed decision, and leads to inconsistent classifications. The degree of conventionalization of strategies can be defined by the morphosyntactic contexts favoring the evidential reading, as opposed to forms which rely on discursive context for their evidential interpretation.
- Perfectoid forms are attested in all of the contemporary East Caucasian languages. Indirect evidentiality and current relevance semantics are both widespread.
- Among several languages spoken in the southern part of the area, the indirect evidential meaning as a function of a perfectoid is absent, which suggests areal influence from Azerbaijani, a Turkic language in which the corresponding verb form has almost completely lost indirect evidential meaning.
- The presence of an active subject is a relevant diagnostic for the distinction of resultative constructions from resultative perfects, which are more grammaticalized than the former.

- Narrative use is an indicator of a certain level of grammaticalization of the evidential meaning, and in languages where the meaning is more grammaticalized, speakers use it more consistently to render unwitnessed narratives, as compared to speakers of languages where the semantics of the perfectoid are more complex.

The **practical significance** of the present study is due to the development of descriptive and analytical frameworks that can be applied for other purposes. The overview in Chapter 2 provides a starting point for more detailed, typologically informed descriptions of similar verb forms in East Caucasian idioms which currently lack such descriptions. Additionally, the systematization of these characteristics might prove helpful in the creation of language teaching materials. The framework for the analysis of narrative texts presented in Chapter 3 can be applied in similar inquiries on languages belonging to other families. The **data** used for the study include full reference grammars, grammar sketches, and specialized articles, as well as fieldwork data collected by the author and texts published in grammars. Fieldwork was carried out predominantly with speakers of Andi dialects. As a result, a bias towards these data may be noticeable, in the sense that many examples and specific features of Andi are referred to throughout the study.

The results of the study were **presented** in several talks at international conferences:

- Tense, Aspect, Modality and Evidentiality (17-18 November 2016, Diderot University Paris)  
**“Evidentiality and other usages of the perfect in Avar and Andi (Nakh-Daghestanian)”**
- Electronic writing systems of the peoples of the Russian Federation (16-17 March 2017, Syktyvkar)  
**“Evidentiality and the perfect in Rutul (data from the village Kina)” in Russian**
- SLE (10-13 September 2017, University of Zürich)  
**“Evidentiality and the perfect in the Rikwani and Zilo dialects of Andi (East Caucasian)”**
- 14th Conference on typology and grammar for young researchers (23-24 November 2017, St.-Petersburg)  
**“Evidentiality in the Rikwani dialect of Andi”**

- 16th Conference of the International Pragmatics Association (9-14 June 2019, Hong Kong) “Narrative use: a measurable feature of evidentiality as a meaning of the perfect”

The dissertation consists of three chapters, an introduction and a conclusion. The summary in the next section provides a short overview of the content of the main chapters. Interactive versions of the maps discussed in Chapter 2 are available at Github. Data visualizations were created in R/R Studio [R Core Team 2018], using the following packages: **tidyverse** for data manipulation [Wickham 2017], **lingtypology** for maps [Moroz 2017], and **ggpubr** for lollipop charts [Kassambara 2018].

## 2 Summary

### 2.1 Chapter 1

The first chapter is dedicated to theoretical and typological approaches to evidentiality as a linguistic category. A central problem in this field is how to define the category. Alexandra Aikhenvald is a famous proponent of the grammar-only view, according to which only grammatical expressions should be considered evidentials. Other means of expressing the same meanings, such as lexical items (i.e. “they say”, “apparently”), or EVIDENTIAL STRATEGIES, forms that can express an evidential meaning in context, but have a different meaning as their main function [Aikhenvald 2004], [Aikhenvald 2018]. In the typology of Aikhenvald [2004], several perfectoid forms from the evidential belt are branded as evidential strategies, while it remains unclear what distinguishes these forms from similar cases in other languages of the area, which are classified as grammatical evidentials. Evidential perfectoids are usually not entirely grammaticalized (otherwise they would be evidential pasts), but they are notably different from evidential strategies which rely entirely on their discursive context, such as Germanic modal verbs. Depending on the language, specific morphosyntactic contexts can be identified which favor or even require an evidential reading.

Therefore, we propose to distinguish four different types of evidential expressions:

- Forms with an invariant evidential meaning + marking evidentiality is obligatory (grammatical evidentials)

- Forms with an invariant evidential meaning + marking evidentiality is not obligatory (grammatical evidentials which are not obligatory, such as certain particles or clitics, and lexical items)<sup>2</sup>
- Forms which have an evidential meaning in specific morphosyntactic contexts
- Forms which have an evidential meaning in specific discursive contexts

In contrast to the approach of Aikhenvald, most researchers view evidentiality as a semantic-functional domain, the values of which can be realized with different means. Nonetheless, the failure to distinguish between functionally different entities leads to a lot of confusion, misappropriation of concepts and, as Aikhenvald rightfully points out, it “obscures the status of evidentiality in languages which do have it as a grammatical category quite distinct from modality, mood or tense” [Aikhenvald 2003: 19]. Chapter 1 also discusses the organization of evidentiality’s semantic domain into specific values. Typological accounts depart from the meanings expressed by specific grammatical forms: meanings expressed with a dedicated morpheme constitute distinct values, which are grouped under macro-labels (e.g. “indirect” covering inference and hearsay) based on their co-occurrence in other forms. Evidentials can also be viewed as deictic or indexical markers, since they determine the directness of access of a deictic centre (usually the speaker). This idea goes back to Jakobson [1957], but was not elaborated until recently, see [Bergqvist 2018]. The analytical advantage of deictic or indexical approaches is that they can capture the communicative content of evidentials beyond their semantics, and take into account the perspective of the addressee and other factors of the speech situation.

A special section of chapter 1 is dedicated to smaller evidential systems and the typology of perfect forms. This section describes how resultative or perfect forms develop evidential and other non-current relevance meanings through the conventionalization of conversational implicatures that are to some degree available in all languages with resultatives or perfects. One of the crucial points discussed in this section is that the inferential meaning appears first. It is optionally followed by the development of a reportative meaning, due to which narrative use becomes possible.

## 2.2 Chapter 2

The second chapter presents an overview of evidentiality in the East Caucasian languages, with a special focus on perfectoids. The first section of the chapter sketches

---

<sup>2</sup>This group can be further divided into lexical and non-lexical items if necessary.

the historical and sociolinguistic context of these languages, followed by a short description of their general grammatical profile. A special section is dedicated to the expression of evidentiality. The following means of marking evidentiality are attested in the East Caucasian languages:

- Perfectoids (indirect evidentiality)
- General pasts (direct evidentiality)
- Specialized auxiliary verbs (various)
- Particles (reportative and inferential)
- Cliticized lexical verbs (reportative and inferentials)
- Forms expressing spatial and personal deixis (various)

In the section on perfectoids, we defined perfectoids as forms which:

- consist of a non-finite verb form with perfective or past meaning and a present tense auxiliary, or can be traced back to such a structure diachronically
- express any of the meanings typologically associated with perfects: resultative, current relevance, indirect evidentiality

Suffixal forms of unknown origin which met the semantic criterion were evaluated separately. They were attested only in Udi, Zakatal Avar and Akhvakh (Akhakhdere dialect).

The sample consists of 59 perfectoids and one defective form with unclear properties (in Bats), from a total of 42 surveyed idioms, including 13 additional dialects besides the 29 languages of the family. Most forms consist of a converb and a present auxiliary (or historically derive from such a structure). Participial forms are attested but rather rare. Resultative and current relevance meanings are attested across the family, while indirect evidentiality is absent among some Lezgian languages and Khinalugh, which are spoken in the southern part of the area.

East Caucasian perfectoids find their origin in resultative constructions, and many of the forms surveyed retain the resultative function. Resultative in the narrow sense refers to a current state as the result of a past event. It should be distinguished from resultative perfect, which is a meaning from the current relevance family. Resultatives are monovalent constructions which designate a state that overlaps with the moment

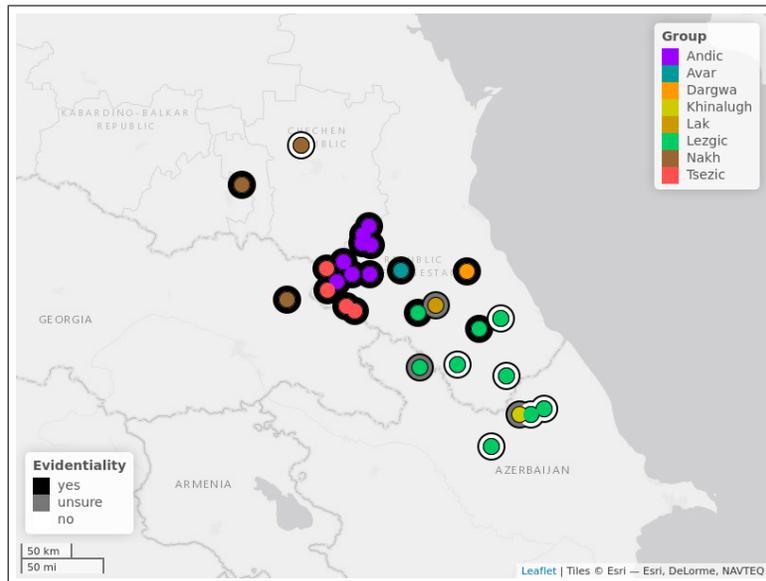
of speech. As a result, their use is limited mostly to intransitive change-of-state verbs. Resultative perfects do not designate an ongoing state but simply a result. In addition, they retain the argument structure of the lexical verb and allow the presence of an active subject. Introducing an active (ergative) subject to a resultative construction forces its reinterpretation as a resultative perfect or an indirect evidential. Other current relevance meanings besides resultative perfect, such as experiential and recent past, are marginally attested. Continuative so far has been attested only in Avar, but this finding requires corroboration.

The evidential reading of perfectoids is influenced by different factors in different languages. The most prominent factor is discursive context: an unwitnessed narrative context may override a default resultative interpretation. Morphosyntactic parameters include: **lexical semantics** of the verb, non-change-of-state verbs in some languages are evidential by default; **aspect**: in some Dargwa varieties, evidential interpretations are specific of perfectoids based on imperfective verbs; **case marking**: the Lak perfectoid is interpreted as evidential when the active subject is in the ergative (instead of in the absolutive, in which case its meaning is current relevance); the presence of an **active/ergative** subject: the introduction of an ergative subject to a construction with a perfectoid in Rikvani Andi triggers an indirect evidential interpretation.

In languages where the perfectoid has an indirect evidential meaning, an auxiliary in the form of the perfectoid is often used to derive the indirect counterparts of periphrastic past tenses. While the perfectoid itself remains polysemic, the forms headed by a perfectoid auxiliary have an invariant evidential meaning.



Figure 3: Distribution of evidential perfectoids in East Caucasian



If indirect evidentiality as a meaning of a perfectoid verb form appeared in East Caucasian as the result of language contact with local Turkic languages, this constitutes a form of **SELECTIVE COPYING** in the terminology of [Johanson 2002], whereby a specific meaning/function of a similar verb form was copied. This implies that the donor form (at least diachronically) should be a perfectoid, with a resultative/current relevance meaning. The Kumyk and Azerbaijani forms meet this requirement. However, due to a lack of diachronic language data, we cannot pin-point the period in time when the indirect evidential usage appeared in any of the East Caucasian languages concerned, and subsequently attempt to relate this to the sociolinguistic situation of the time. Azerbaijani seems to have mostly lost the indirect evidential feature and curiously, it is absent among several languages which have been in close contact with this language. It should be pointed out that the indirect evidential meaning is nonetheless present among two languages from the Azerbaijani zone of influence: Tsakhur and Agul. Cross-linguistically, the development of an indirect evidential meaning from a perfect-like form is quite natural, and can result from language-internal processes. Due to the lack of historical data for East Caucasian (on language use and language contact), it is difficult to defend the Turkic hypothesis, because it is impossible to connect any specific East Caucasian form to its possible donor from Kumyk, either by formal similarities of the forms, or diachronic correlations between the appearance of the eviden-

tial meaning and the nature of language contact. In addition, due to the complex multilingual situation in the area, a scenario whereby one or two East Caucasian languages adopted the feature from e.g. Kumyk and subsequently dispersed it further to minor languages, cannot be ruled out. According to [Johanson 2006], the Turkic languages in the Caucasus merely promoted “latent tendencies” toward indirect evidentiality. In this light, one could argue that the influence of Azerbaijani perhaps promoted current relevance, and simultaneously suppressed indirect evidential readings.

At least in the case of the East Caucasian area, the use of evidential perfects in narratives could tentatively be considered as a possible vehicle for dissemination of the feature. Turkic peoples in Daghestan played an important role in the spread of Islam and literature and culture from the Islamic world. The Islamization of the East Caucasian peoples spread from the south towards the north. As mentioned in [Adžiev 1991], this pattern is reflected in the geographical distribution of folklore motives. The famous North Caucasian Nart sagas are well-attested in the Nakh languages, to some degree also in languages spoken in northwestern and central Daghestan, including Kumyk, while they are absent in Lezgian languages and in Azerbaijani. This could have something to do with a cultural and literary convergence in the respective areas (north vs. south). In addition, narrative use of perfects can be viewed as a kind of minimally counterintuitive concept (MCI) - a concept from cognitive anthropology which is used to account for the spread of religious ideas. MCIs are concepts that adhere to some of the expectations they entail (which makes them intuitive), and simultaneously violate certain other expectations (which makes them memorable) Boyer [1994]. An often cited example of an MCI is a flying carpet. A flying carpet is essentially a carpet like any other, aside from the fact that it can fly, which contradicts its conceptualization as an inanimate object with no built-in mechanism for motion. MCIs have been shown to spread more easily than intuitive concepts on one hand, and maximally counterintuitive concepts on the other, see e.g. Barrett and Nyhof [2001]. If a speaker of a language with a prototypical perfect is confronted with an indirect evidential perfect from another language, their expectations are partially met by the retained resultative/current relevance functions, while (indirect evidential) narrative use is counterintuitive (it contradicts the semantics of current relevance), which makes it highly salient and thus promotes copying.

### 2.3 Chapter 3

Chapter 3 introduces narrative use as a measurable feature. In the first stage of the research, elicited data were collected with speakers of Rutul, Andi and Avar. The results of these surveys were discussed in [Verhees 2018]. However, elicited data on ev-

identicality is subject to a number of factors that influence the speaker's answer, and which are difficult to control for. Evidentiality is a strongly context-dependent factor, and speakers of evidential languages are known not to produce evidential forms they would use in a natural dialogue, when they find themselves in an unnatural situation such as elicitation by a researcher Aikhenvald [2004]. Another problem can be **lack of concentration**, which prevents the speaker from fully understanding the intended context and imagining the most appropriate utterance. As a result, elicitation may produce very divergent results with individual speakers, making any claim based on a single or a few surveys difficult to prove or falsify. In addition, the method does not allow for the distinction of more and less conventionalized meanings - a meaning is either present in the speaker's mind, or not. The only survey questions which seemed to produce more or less consistent results consisted of a short narrative from different perspectives: from personal experience, hearsay from someone familiar, framed as a fairy tale, etc.

Two extended versions of this narrative, one witnessed and one unwitnessed, were subsequently collected with 13 speakers of different ages and genders, from different Andi villages (5 speakers from Rikvani, 6 speakers from Zilo and 1 speaker from the village Rushukha and 1 from Muni). At least Rikvani, Zilo and Muni have distinct dialects, of which Muni is considered the most divergent. The Perfect was significantly more frequent in the unwitnessed context, although alternative strategies, for example with an evidentially neutral Aorist, were also attested. Interestingly, even speakers for whom the evidential semantics were not salient, used the Perfect as an indirect evidential narrative tense.

An analysis of unelicited narrative texts in Bagvalal (Andic) and Tsakhur (Lezgian) showed differences in the frequency of Perfect forms used in narratives, which correlates with their function. Bagvalal has an evidential system typical for East Caucasian languages. The Perfect has resultative, current relevance, indirect evidential and occasionally mirative meaning [Tatevosov 2001]. It is paradigmatically opposed to a less marked, general past tense (the Preterite). Periphrastic past tenses are formed with a Preterite auxiliary, and the Perfect auxiliary derives their indirect evidential counterparts.

Table 1: Basic past tenses of ‘stand up’ in Bagvalal

	preterite hec’i	perfect hec’i-b-o ek’wa
pluperfect	hec’i-b-o b-uk’a	hec’i-b-o b-uk’a-b-o ek’wa
imperfect	hec’ira:χ b-uk’a	hec’ira:χ b-uk’a-b-o ek’wa
past habitual	hec’iro:-b b-uk’a	hec’iro:-b b-uk’a-b-o ek’wa

In Tsakhur, tenses with the copula *wo-d* ‘be’ in its non-attributivized form (NAF) mark the speaker’s **distancing** from an event, which often coincides with unwitnessed events, but this is not its main meaning [Maisak and Tatevosov 2007]. This is quite similar to Lazard’s definition of **MEDIATIVE**, a sort of abstract macro-category comprising all the contextual interpretation of indirect forms [Lazard 1999]. Analogous forms with an attributivized copula (AF) are neutral. Though the NAF/AF opposition is characteristic of the entire verbal paradigm, the evidential-like meaning is specific to forms with *wo-d*, which include the Perfect, the Durative (present) and the Prospective (future).

Table 2: Basic past tenses of ‘open’ in Tsakhur

	NAF	AF
Aorist	aqi	aqi-na
Perfect	<b>aqi wo-d</b>	<b>aqi wo-d-un</b>
Pluperfect	aqi ixa	aqi ixa-na
Imperfect	aqi ixa	aqi ixa-na

The data consist of spoken texts recorded during fieldwork in the 1990s-2000s, which were published as transcribed and annotated texts in Kibrik et al. [2001] on Bagvalal, and Kibrik [1999] on Tsakhur. The sample contains texts belonging to different genres and produced by different speakers. From each text, narrative sequences consisting of at least two consecutive finite clauses were collected. Annotation consisted of identifying the first sentence of a sequence and the main line. Dependent clauses (which include a large amount of converbial clauses and participial relative clauses) are all marked “dependent” and not counted further. Reported speech is marked as **quote** and also excluded. The reported speech strategy in both languages is mostly direct, and creates an embedded context where the regular narrative restrictions on use of forms do not apply. Quotes are frequent and contain a larger variety of verb forms than the main line, including non declarative forms like imperatives. Metadata on texts include:

the speaker (anonymized with a code) and their gender, topic of the text, genre (fairy tale, local legend, anecdote, personal anecdote, personal recollection) and perspective (indirect, direct, dream, unclear). Table 3 summarizes the content of the sample.

Table 3: Content of text samples

	Bagvalal		Tsakhur	
	M	F	M	F
Speakers	7	2	6	1
Texts	18	6	11	1
Sentences	406	60	317	16
Verb forms	1220	149	864	41
Sequences	43	6	13	4

Since narrative texts tend to be fairly homogenous in terms of the forms used, and speakers pick a particular narrative strategy to render the main line of the story, we expect a small number of forms to be highly frequent. Figure 4 shows that this was indeed the case: a limited number of forms is highly frequent. The long tail of various forms which are used only once comes from the reported speech context. If we limit the analysis to main clauses, the variety of high frequency forms is further diminished, as shown in Table 4.

Figure 4: Forms in Bagvalal and Tsakhur narratives (all clause types)

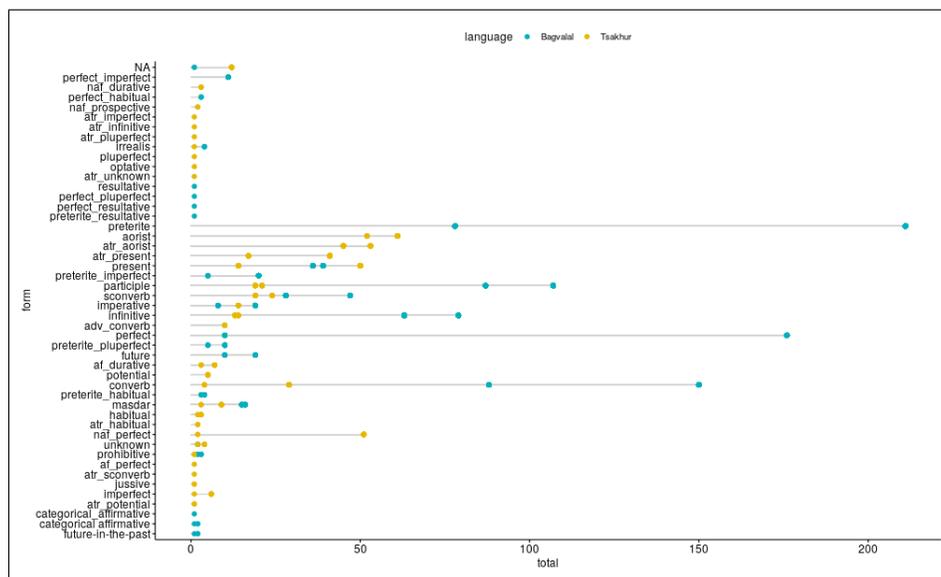


Table 4: Use of pasts and perfects in Bagvalal and Tsakhur narrative sequences (main clauses)

	Bagvalal		Tsakhur	
	Direct	Indirect	Direct	Indirect
Perfect	6	165	2	50
Past	181	63	79	86
Other	57	40	49	14

In both languages, the Perfect is significantly more frequent in the indirect contexts. In Bagvalal, it is more frequent than the general past in this context, while in Tsakhur the general past is more frequent in both the direct and indirect contexts. In Bagvalal, the correlation between perspective and forms used is statistically significant, regardless of the asymmetry between the Perfect and the Preterite ( $\chi^2 = 199.99$ ,  $df = 1$ ,  $p\text{-value} < 2.2e-16$ ). For Tsakhur, the Chi-square test also showed a statistically significant difference ( $\chi^2 = 30.915$ ,  $df = 1$ ,  $p\text{-value} = 2.696e-08$ ). This does not concur with the idea that the Tsakhur Perfect is “less evidential”. Further research with a larger sample of languages and texts is necessary to verify these tendencies.

### 3 References

- Adžiev, A. M. (1991). Vvedenie [introduction]. In G. Gamzatov and U. Dalgatov (Eds.), *Tradicionnyj fol'klor narodov Dagestana* [*The traditional folklore of the peoples of Daghestan*], pp. 5–18. Moscow: Nauka.
- Aikhenvald, A. Y. (2003). Evidentiality in typological perspective. In A. Y. Aikhenvald and R. M. Dixon (Eds.), *Studies in Evidentiality*, pp. 1–32. Amsterdam/Philadelphia: John Benjamins.
- Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. (2018). *The Oxford Handbook of Evidentiality*. Oxford: Oxford University Press.
- Barrett, J. L. and M. A. Nyhof (2001). Spreading non-natural concepts: The role of intuitive conceptual structures in memory and transmission of cultural materials. *Journal of cognition and culture* 1(1), 69–100.
- Bergqvist, H. (2018). Evidentiality as stance: event types and speaker roles. In A. Foolen, H. de Hoop, and G. Mulder (Eds.), *Evidence for evidentiality*, pp. 19–44. Amsterdam/Philadelphia: John Benjamins.
- Boyer, P. (1994). *The naturalness of religious ideas*. Oakland: University of California Press.
- Chafe, W. L. and J. Nichols (1986). *Evidentiality: The linguistic coding of epistemology*. Norwood, NJ: Ablex.
- Chirikba, V. (2003). Evidential category and evidential strategy in abkhaz. In A. Y. Aikhenvald and R. M. Dixon (Eds.), *Studies in Evidentiality*, pp. 243–290. Amsterdam/Philadelphia: John Benjamins.
- Comrie, B. (2000/1985). *Tense*. Cambridge: Cambridge university press.
- Erdal, M. (2004). *A grammar of Old Turkic*. Leiden: Brill.
- Forker, D. (2018). Evidentiality in nakh-daghestanian languages. In A. Y. Aikhenvald (Ed.), *The Oxford Handbook of Evidentiality*, pp. 490–509. Oxford: Oxford University Press.

- Jakobson, R. (1957). *Shifters, verbal categories and the Russian verb*. Cambridge, MA: Harvard University, Department of Slavic Languages and Literature. Reprinted in L. R. Waugh and M. Monville-Burston (Eds.) 1984. *Russian and Slavic grammar studies 1931-1981*. Berlin: Walter de Gruyter, 386-392.
- Johanson, L. (2002). Contact-induced change in a code-copying framework. In M. C. Jones and E. Esch (Eds.), *Language change*, pp. 285–314. Berlin/New York: Mouton de Gruyter.
- Johanson, L. (2006). On the roles of turkic in the caucasus area. In Y. Matras, A. McMahon, and N. Vincent (Eds.), *Linguistic areas. Convergence in historical and typological perspective*, pp. 160–181. New York: Palgrave MacMillan.
- Johanson, L. (2018). Turkic indirectivity. In A. Y. Aikhenvald (Ed.), *The Oxford Handbook of Evidentiality*, pp. 510–524. Oxford: Oxford University Press.
- Kassambara, A. (2018). *ggpubr: 'ggplot2' Based Publication Ready Plots*. R package version 0.2.
- Kibrik, A. E. (1999). *Élementy kazurskogo jazyka v tipologičeskom osveščanii [Elements of Tsakhur grammar in a typological perspective]*. Moscow: Nasledie.
- Kibrik, A. E., K. I. Kazenin, E. A. Ljutikova, and S. G. Tatevosov (2001). *Bagvalinskij jazyk. Grammatika, teksty, slovari [The Bagvalal language. Grammar, texts, dictionary]*. Moscow: Nasledie.
- Labov, W. and J. Waletzky (1967). Narrative analysis: oral versions of personal experience. In J. Helm (Ed.), *Essays on the verbal and visual arts. Proceedings of the 1966 Annual Spring Meeting of the American Ethnological Society*, pp. 12–44. Seattle/London: University of Washington Press.
- Lazard, G. (1999). Mirativity, evidentiality, mediativity, or other? *Linguistic Typology* (3), 91–109.
- Lindstedt, J. (2000). The perfect-aspectual, temporal and evidential. In E. Dahl (Ed.), *Tense and aspect in the languages of Europe*, pp. 365–384. Berlin/New York: Mouton de Gruyter.
- Maisak, T. and S. Tatevosov (2007). Beyond evidentiality and mirativity. evidence from tsakhur. In Z. Guentchéva and J. Landabaru (Eds.), *L'Énonciation médiatisée II*, pp. 377–406. Louvain/Paris/Dudley MA: Peeters.

- Moroz, G. (2017). *lingtypology: easy mapping for Linguistic Typology*.
- Paducheva, E. V. (2010). *Semantičeskie issledovanija: semantika vremeni i vida v russkom jazyke; semantika narrativa, vtoroe izdanie* [Semantic studies: the semantics of tense and aspect in Russian; the semantics of narrative, second edition]. Moscow: Jazyki slavjanskoj kul'tury.
- Plungian, V. (2010). Types of verbal evidentiality marking: an overview. In G. Diewald and E. Smirnova (Eds.), *Linguistic realization of evidentiality in European languages*, pp. 15–58. Berlin/New York: De Gruyter Mouton.
- Plungian, V. (2016). K tipologii perfekta v jazykax mira. predislovije [towards a typology of the perfect in the world's languages: Introduction]. In T. Maisak, V. Plungian, and K. Semënova (Eds.), *Tipologija Perfekta* [Typology of the perfect], pp. 7–38. St. Petersburg: Nauka.
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Tatevosov, S. (2001). From resultatives to evidentials: Multiple uses of the perfect in nakh-daghestanian languages. *Journal of Pragmatics* 33(3), 443–464.
- Tatevosov, S. G. (2007). Évidencial'nost' i admirativ v bagvalinskov jazyke [evidentiality and admirative in bagvalal]. In V. S. Xrakovskij (Ed.), *Évidencial'nost' v jazykax Evropy i Azii: sbornik statej pamjati Natal'i Andreevny Kozincevoj* [Evidentiality in the languages of Europe and Asia: papers in honor of Natalia Andreevna Kozintseva], pp. 351–397. St.-Petersburg: Nauka.
- Verhees, S. (2018). The perfect in avar and andi. cross-linguistic variation among two closely-related east caucasian languages. pp. 261–279.
- Wickham, H. (2017). *tidyverse: Easily Install and Load the 'Tidyverse'*. R package version 1.2.1.