

**Evidentiality and related categories  
in four non-Slavonic languages of the Russian Federation:  
Bashkir, Even, Lezgi and Tatar**

**Teija Greed**

*Academic dissertation to be publicly discussed, by due permission of the Faculty of Arts  
at the University of Helsinki, in Auditorium XII (Main Building),  
on the 2nd of March, 2019, at 12 o'clock.*

General Linguistics  
Department of Languages  
University of Helsinki  
2019

Articles 1 and 4 of Part II have been included in the paperback version with permission from their respective publishers, and Articles 2 and 3 have been published in open access journals.

ISBN 978-951-51-4883-4 (paperback)  
ISBN 978-951-51-4884-1 (PDF)

Printed by Unigrafia  
Helsinki 2019

## **Abstract**

In this thesis I investigate evidentiality and other related functional categories in four non-Slavonic languages spoken in the Russian Federation: a Nakh-Daghestanian language Lezgi, two Turkic languages Bashkir and Tatar, and a Tungusic language Even. The study consists of four peer-reviewed articles and an introductory article. I present common traits of the languages investigated, and explore the distinctive features of each of them, while setting the phenomena in their cross-linguistic context.

I worked with mother-tongue language consultants, studied grammars and language descriptions, and analysed natural texts in printed and electronic form, utilising existing language corpora for two of the languages. A key aspect in the analysis of the data was contextual information in order to accurately interpret the relevant linguistic and pragmatic cues.

I define evidentiality as expression of information source through grammatical means, following and expanding on Aikhenvald (2004), in that I discuss optional evidential expression through grammatical means on a par with obligatory expression. The key evidential semantic parameters established in this study are generally in the area of non-firsthand evidentiality: the so-called propositional evidentials non-witnessed (reported), hearsay, inference and assumption, as well as the illocutionary evidential quotative.

The aim of this study is to consider the following questions: How are evidential meanings conveyed through grammatical means in the four languages; and how does evidentiality interact with other related categories? To answer these questions, I have established the relevant semantic parameters for the analysis of evidentiality in these languages, investigated the interaction of evidentiality with other functional categories and domains, and explored the historical development of the meanings involved in the sphere of evidential semantics. Thus this study employed diverse perspectives: descriptive and typological, on the one hand, and synchronic and diachronic, on the other. The key categories and domains discovered to interact with evidentiality in the four languages are tense-aspect, (epistemic) modality, mirativity, subjectivity, and discourse.

As is cross-linguistically widely attested, also in all the four languages the tense-aspect form carrying the perfect, or resultative, meaning plays an important part in evidential expression. For the four languages, the perfect meaning has evolved to express the evidential meaning of inference. On the path of development of the perfect, or resultative, meaning, Lezgi has remained at the stage of inference, whereas Bashkir, Tatar and the western and central dialects of Even have progressed to express non-witnessed and reported meanings. Evidential meanings were found to interact with the domain of discourse in that in Bashkir, Tatar and Even the originally perfect meaning has evolved through inferential and non-witnessed meanings to express a discourse function: it is used as a special narrative tense, which carries the narrative plot forwards. Such a form becomes a marker of a particular genre.

All of the four languages were found to display mirative meanings, Lezgi mainly with particles, but Bashkir, Tatar and Even also as an extension of the inferential meaning of the erstwhile perfect meaning. In the three languages, in first-person contexts the inferential meaning receives a mirative reading, which development, the first-person effect, is well documented cross-linguistically. In addition, I discovered that Even displays a mirative meaning in second-person contexts. In this “second-person effect” the speaker remains the evidential origo (Sun 2018) of the utterance, as the verb form carrying the evidential meaning signals the speaker’s viewpoint.

The Bashkir illocutionary quotative evidential, typically signalling direct speech, was found to also mark semi-direct speech in conjunction with complement-taking verbs of cognition and perception. Semi-direct speech is a speech report that retains most of the characteristics of direct speech but conveys internal speech, such as thoughts and intentions. Thus it is the cognitive processing of information by the speaker/experiencer that comes to the fore in such expressions: subjectivity interacts with evidentiality, as the speaker employs different linguistic possibilities to convey the degree of her involvement. Subjectivity can also extend further than to the speaker/experiencer, as I discovered in Bashkir: in conjunction with the verb of cognition ‘know’ the quotative codes also the original speaker(s) of the speech report, in addition to the current speaker/experiencer, thus merging the two perspectives. This notion of “multisubjectivity” comes close to what

Evans (2005) calls “multiple perspective”, the encoding of more than one viewpoint simultaneously.

In conclusion, I present a visual display of the interaction of evidentiality and related categories in the four languages. The sketch suggests new avenues for further cross-linguistic exploration of the complex interaction of semantic categories involved in human communication that awaits discovery.

## **Acknowledgements**

What a diverse array of languages there is in the world! I have had the joy of interacting with a great number of people speaking different mother tongues. These people have enabled me to appreciate the fascinating variety of languages, peoples and cultures there is in the world. My initial interest in languages led me to choose a more exotic language for study at high school: Russian. Some years later I completed my studies at the University of Tampere with an MA in Slavonic Philology.

When work took me, my English husband and our two young children to Russia in the early 1990s I had the privilege of studying Tatar at Kazan State University, whose Tatar-language department was then in the capable hands of Professor Flera Safiullina. The distinguished Tatar academic, a director at the Academy of Sciences of Kazan, Professor Mirfatyx Zakiev, and many others passed on the example of the love for their language, culture and people to me and my family. Among them are the distinguished translator, the late Räs Dautov and the poet, the late Firaja Zyjatdinova. I would like to express my special thanks to Docent Gönaz Mugtasimova of Kazan State University, who in my early days of Tatar study tutored me in the language, becoming a dear family friend. More recently, Gönaz Mugtasimova helped in the preparation of the data for my Tatar article as my language consultant. A special thank-you also goes to the editor and translator Nailä Valitova, whose wisdom and knowledge of the intricacies of her mother tongue was of great help in the preparation of my article. I cherish the many hours we spent together discussing a variety of topics in Tatar.

The impetus for continuing my formal studies came from a variety of directions. I am grateful to my colleague and friend Dr Linda Humnick who encouraged me to delve deeper into linguistics, resulting in my application to the University of Helsinki to study general linguistics. I am deeply grateful to Professor Fred Karlsson who encouraged me on this path. My part-time studies in general linguistics commenced in 2007, and after completing a secondary MA, doctoral studies began in 2011.

After Professor Karlsson, I had the privilege of continuing my studies under the supervision of another outstanding linguist, Dr Ekaterina Gruzdeva. I am deeply indebted to her for her willingness to be available whenever I needed help, for her encouragement and advice, for connecting me with other linguists, and for the crucially important comments she made on my articles at the preparation stages. Throughout the years of my studies I have also had the opportunity to learn from, and interact with, many other students of linguistics especially at researcher seminars led by Professor Urho Määttä and Professor Matti Miestamo, as well as to widen my linguistic horizons through participation at workshops and conferences.

For my MA thesis I researched evidentiality in Tatar (Greed 2009), on the basis of which I was honoured to receive an invitation to attend a workshop on evidentiality at James Cook University, Australia in 2012, led by Professor Alexandra Aikhenvald and Professor R. M. W. Dixon. This workshop resulted in a book *The Grammar of Knowledge* (Aikhenvald & Dixon 2014), and contains the first article of this thesis. The participation in the workshop, and the interaction with talented linguists from all over the world gave me added motivation in my research in other non-Slavonic languages of the Russian Federation.

I would like to extend a special thank-you to all my language consultants. In addition to the already mentioned Gölnaz Mugtasimova and Nailä Valitova, who helped with the data of my article on Tatar evidentiality (Greed 2014 (2016)), I am greatly indebted to the distinguished Lezgi writer Arben Kardash, as well as to the translator Ofelija Pirverdijeva for their invaluable help with the data used in my article on evidential coding in Lezgi (Greed 2017). Similarly, I extend thanks to Gölnaz Xužaxmetova and to the writers Gölnara Mugtasimova and Gölsirä Gizzatullina in the role they played in the preparation of my article on the Bashkir quotative (Greed 2018).

With my final article my sincere thanks go first of all to my colleague Esther Burgherr. I had the privilege of assisting her as one of her supervisors for her outstanding MA thesis on the information structure of the Even language. In addition to making some of her own data available to me, Esther put me in touch with Ilana Zaxarova, who became my Even language consultant. And so my thanks go also to Ilana Zaxarova for her contribution in the preparation of the Even language data of my article on the perfect and how it manifests itself in Even (Greed 2019).

In connection with the whole of the thesis, I wish to extend my sincere thanks to my preliminary examiner and opponent Professor Andrej Malchukov for his constructive feedback and encouragement to delve deeper in the matters investigated. I am also deeply grateful to Professor Kasper Boye, my other preliminary examiner, for stimulating my thinking to wider theoretical issues. The preliminary examiners' suggestions and comments enabled me both to clarify my own line of thought and improve the presentation of the topic.

Over the years, I have had the opportunity of interacting with and enjoying the friendship of many people who have had a significant impact on me, my studies and my life's path. Some of them have been encouragers, some have been available for meaningful discussions and have offered invaluable advice, or shared their linguistic findings with me; others have inspired me with the example of their own life and experiences. I can list only but a few of them: Marianne Beerle-Moor, Diana Forker, Margaret Foster, the late Mick Foster, Monika Höhlig, Eric Jones, Gwyn Jones, Erwin Komen, Katja Müller, Geneviève Perrault, Nailya Philippova, Marja Ruotsalainen, Riitta Sauni, Simo Sauni, Zöfär Xamadraximov and Paula Vuorinen. I would also like to thank my wider family, especially my sister Riitta Hutila, for their interest and support.

Finally, I want to express my deepest gratitude to my dear husband Michael, my constant encourager and enthuser. I am amazed at his untiring dedication to his role as my language checker, and for patiently going through all my writings with helpful suggestions for improvement. His contribution to my work has also been significant in the process of research and writing: many a time I would arrive at the dinner table with a linguistic conundrum going round in my mind, and he would help me clarify my thinking, or disentangle me from the pursuit of yet another interesting, but tangential, detail

relating to these fascinating languages. I am grateful for his keenness to think “outside the box” and for helping me step into new avenues of discovery; also for his general technical assistance. My thanks go also to our two remarkable children, Emma-Liisa and Christopher, who have shared together with their parents the joy of learning to speak Tatar and Russian, in addition to their mother’s and father’s tongues, Finnish and English, and who have been a great encouragement to me throughout my years of study.

Last but not least I want to express my gratitude to my Heavenly Father who instilled in me a passion for languages and gave me the perseverance to see this study programme through to completion. I echo the words of the biblical sage who said, “With God are wisdom and strength; he has counsel and understanding” (Job 12:13).



## Contents

|                       |     |
|-----------------------|-----|
| Abstract.....         | i   |
| Acknowledgements..... | iii |

### **PART I: Introductory Article**

|  |    |
|--|----|
| 1. Goal, methods and data of the thesis.....   | 1  |
| 2. Background to evidentiality.....  | 5  |
| 2.1. Evidentiality and communication .....   | 5  |
| 2.2. Evidentiality: history and different viewpoints.....  | 6  |
| 3. Evidentiality in relation to other categories and domains from a theoretical perspective..... | 12 |
| 3.1. Evidential meanings as extensions of tense-aspect.....                                      | 13 |
| 3.2. Evidentiality, mirativity and person.....   | 14 |
| 3.3. Evidentiality, epistemic modality and subjectivity.....                                     | 15 |
| 3.4. Evidentiality and the domain of discourse.....  | 19 |
| 3.5. Evidentiality and scope.....  | 19 |
| 3.6. Evidentiality in semantic space.....  | 20 |
| 4. Abstracts of articles.....  | 21 |
| 5. A brief typological comparison of Tatar, Bashkir, Lezgi and Even.....                         | 27 |
| 6. Evidentiality and other categories and domains in Bashkir, Even, Lezgi and Tatar.....         | 28 |
| 6.1. Evidentiality and tense-aspect.....   | 29 |
| 6.2. Default evidentials and their interaction with other categories.....                        | 35 |
| 6.3. Quotative, subjectivity and the semantic type of verbs.....                                 | 37 |
| 6.3.1. Quotative <i>tip</i> , semi-direct speech, subjectivity and multisubjectivity.....        | 38 |
| 6.3.2. Quotative <i>tip</i> as an adverbial marker.....  | 40 |
| 6.3.3. The main functions of the quotative <i>tip</i> and their distinctive features....         | 41 |

|   |    |
|---|----|
| 6.3.4. Syntactic status of the speech report in semi-direct speech.....   | 42 |
| 6.4. Evidentiality and epistemic modality.....  | 45 |
| 6.5. Evidentiality, mirativity, and person.....   | 48 |
| 6.5.1. First-person effect and mirativity in Even.....  | 49 |
| 6.5.2. Second-person effect and mirativity in Even.....   | 50 |
| 6.6. Evidentiality and discourse.....   | 53 |
| 6.6.1. Discourse genres in Tatar.....   | 54 |
| 6.6.2. Narrative functions of the Even perfect.....   | 55 |
| 6.7. Summary of interaction between evidentiality and related categories and domains in Bashkir, Even, Lezgi and Tatar..... | 58 |
| 7. Conclusion.....  | 60 |
| Abbreviations.....  | 62 |
| References.....   | 64 |

## **PART II: Articles**

### List of Articles

Article 1: The expression of knowledge in Tatar

Article 2: Evidential coding in Lezgi

Article 3: The quotative in Bashkir

Article 4: From perfect to narrative tense: The development of an evidential meaning examined generally and in the Even language

**PART I:**  
**Introductory Article**

## 1. Goal, methods and data of the thesis

This thesis investigates the expression of evidentiality and related categories in four non-Slavonic languages in the Russian Federation and consists of four independent academic articles and an introductory article. Through my work and studies in the Turkic Tatar language I became for the first time aware of the category of evidentiality. This led me to investigate this phenomenon, initially in Tatar, but later also in some other non-Slavonic languages spoken in the Russian Federation. The Nakh-Daghestanian Lezgi drew my attention as a closely related language Aghul had been demonstrated by Majsak and Merdanova (2002) to have evidential opposition of past-tense forms. In Bashkir, a language closely related to Tatar, the frequent use of the quotative marker *tip* especially in connection with non-direct speech proved a rewarding topic. Lastly, Malchukov's (2000) article discussing perfect in the North-Tungusic Even made me recognise the connection between perfect, evidential inference and its relevance in discourse, which gave occasion to delve more deeply into this relationship in the Even language.

Furthermore, these four languages are overall less studied in the area of evidentiality and related categories, and represent three different language families. They display and illustrate a range of different evidential and related phenomena, revealing similarities that make them more amenable to comparison, while at the same exhibiting distinctive characteristics, which open up possibilities for an exploration of a variety of different phenomena. The focal points chosen for the four research articles represent less studied topics related to evidentiality. A key motivation for the current study has been to discover, explore and describe connections between different functional categories and domains, to seek to find explanations as to how they relate to each other, and to draw the discoveries together so as to present the interactions in a holistic manner.

Although Tatar is the second most widely spoken language in the Russian Federation and Turkic languages are in general known to have the category of evidentiality (see, for example, Johanson's (2000) discussion on Turkic indirectivity), Tatar has only recently begun to receive attention in this area (see Isxakova et al. 2007 and Tatevosov 2007). Bashkir is also a widely spoken language with over one million

speakers, with evidential meanings mentioned in the official grammar, but otherwise not widely researched. As for Lezgi, a language with ca. 600,000 speakers, it has been studied less than a considerably smaller related language Aghul (see Majsak & Merdanova 2002). Even differs very clearly from the other three languages examined in the thesis, being a language with a much smaller ethnic population of 22,000, only a quarter of whom speak the language. Evidentiality in this language has been researched in more detail both in the works of Malchukov (2000) and Zippel (2012).

The overarching research question of my thesis is the following: How are evidential meanings, that is, expressions of information source, or source of knowledge, conveyed through grammatical means in a selection of non-Slavonic languages spoken in the Russian Federation? As evidentiality, like other linguistic categories, does not function in isolation, this led to another question: How does evidentiality interact with other related categories? For the purpose of answering these questions, my goal is: 1. to establish relevant semantic parameters for the analysis of evidentiality in the languages under study, 2. to explore the interaction of evidentiality with other functional categories and domains, and 3. to investigate the historical development of the meanings involved in the sphere of evidential semantics. The approach employed is both typological and descriptive, and the languages are investigated from both synchronic and diachronic viewpoints.

The emphasis of this study is on verb forms, but other non-lexical<sup>1</sup> means of expression (such as particles/clitics) are also investigated (see, in addition, the discussion on Tatar lexical expression of information source in Greed (2014 (2016): 83–86)). For a form to be investigated, it needs to have an evidential meaning. This meaning does not need to be a default meaning of the marker in question, but it can be a meaning extension of another grammatical category. Through research of individual languages, I explore the more general picture with the goal of proposing where each language fits into the general semantic space of evidentiality, epistemic modality, mirativity, and other related meanings. However, even though my aim is for a holistic portrayal, only Tatar (Article 1), and to a lesser degree Bashkir (Article 3), receive a fuller treatment of evidential values. For the others I do not endeavour to cover all possible markers conveying evidential

---

<sup>1</sup> See further discussion on the distinction lexical–grammatical in Section 2.1.

meanings, but concentrate on aspects of evidential expression which appear to be particularly relevant. Amongst these aspects, I introduce a special semantic dimension of subjectivity, and discuss its relationship with evidentiality. For some of the languages, I also make a proposition about the path of development of the markers involved in the expression of evidential meanings.

The original peer-reviewed articles included in the thesis are the following:

1. Greed, Teija 2014 (2016). The expression of knowledge in Tatar. In: Aikhenvald, Alexandra Y. & R. M. W. Dixon (eds). *The Grammar of Knowledge*. Oxford: Oxford University Press. 69–88.
2. Greed, Teija 2017. Evidential coding in Lezgi. *Languages of the Caucasus*. Volume 2. Available at <https://escholarship.org/uc/item/0w59z5t4>.
3. Greed, Teija 2018. The quotative in Bashkir. *Studia Orientalia Electronica*, Volume 6, 23–55. Available at <https://journal.fi/store/article/view/69685>.
4. Greed, Teija 2019 (in press). From perfect to narrative tense: The development of an evidential meaning examined generally and in the Even language. *Studies in Language* 42:4, 923–966. Will be available via <https://www.benjamins.com/catalog/sl>.

Each article discusses one or more areas of evidential expression in the language, while offering a special perspective to a more general topic that is significant to that particular study. Therefore while Article 1 presents a general picture of the expression of knowledge in Tatar, and concentrates on meanings expressed with grammatical markers, the discussion is also expanded to lexical expression of knowledge. Article 2 on evidential coding in Lezgi differs from the other articles in that the evidential expression of Lezgi is compared with that of a related language Aghul. Article 3 investigates the quotative in Bashkir, and examines the semantic dimension of subjectivity, the concept already introduced in Article 1, and its important role in the pragmatic outworkings of the quotative particle in conjunction with a certain set of verbs. In Article 4 I explore the development of the perfect meaning into a narrative function, and set the investigation of

this process in the Even language in the wider typological context of the evolution of the perfect.

My method of research, common in the preparation of the four articles, was to work with a mother-tongue language consultant, or consultants. For two languages I used available language corpora: for Lezgi, an internet Lezgian Corpus, and for Even, an extensive language corpus that Dr Pakendorf kindly made available to me. In addition, I studied grammars and language descriptions, and analysed natural texts in printed and electronic form. When working with the language consultants virtually, in addition to the consultants' spontaneously created examples, data was also collected through elicitation<sup>2</sup>. While the method of elicitation was more pronounced with the Lezgi data in Article 2, for Tatar (Article 1) and Bashkir (Article 3) most data was from natural texts and conversations. Even was a special case, due to its complex dialect situation (see Pakendorf 2015: 161). The examples I used in Article 4 represent five dialects, which differ from each other not just in phonology and lexicon, but also morphology and syntax. This presented a challenge in the description and analysis of the material. A key aspect for me in the examination of the data was access to contextual – both intratextual, and when possible, extratextual – information, as the situational context is of great importance for the accurate interpretation of the linguistic and pragmatic cues.

This introductory article of the thesis is divided into two main parts: a discussion of the main topics of the four articles from a general cross-linguistic viewpoint, and an empirical exploration of these topics and the results of the study in the four languages researched. Sections 2 and 3 offer the general discussion, and Section 4 presents the abstracts of the four articles. In Sections 5 and 6 I investigate and present the main results of the separate studies. My aim is twofold: to display common traits of the languages studied, and to show the distinctive particulars that are characteristic to each one of them.

The main findings in the four languages are supplemented by graphic displays originally published in the four articles. These are Diagram 1 (in Section 6.1), Diagram 4 (in Section 6.4) and Table 5 (in Section 6.6) for Tatar; Diagram 2 (in Section 6.1) for Even; Diagram 3 (in Section 6.2) for Lezgi and Table 4 (in Section 6.3) for Bashkir. In addition,

---

<sup>2</sup> The method of grammaticality judgments was used in the conjunction with elicitation, but not as a consistent principle.

Figure 1 (in Section 6.1) presents a semantic map bringing together the findings of the functions of the perfect. The discussion is drawn to a close in section 6.7 with a visual display of the interaction between evidentiality and related categories and domains in the four languages, depicted in Figure 2.

I will bring this introductory article to a close with considerations on the contributions this thesis offers to evidential studies, and conclusions drawn from the study of the topics and of the four articles.

## **2. Background to evidentiality**

### **2.1. Evidentiality and communication**

Knowledge cannot be detached from reality, nor is it simply to do with conscious subjects in isolation; it is ‘a social phenomenon, an aspect of the social relations between people’ (Hill & Irvine 1992: 17). The speaker can pass on knowledge using grammatical, lexical and extralinguistic means. Amongst these, evidentiality is a linguistic category to do with expression of information source: the speaker uses grammatical evidential markers to communicate to the addressee the source of the knowledge she<sup>3</sup> is conveying. All languages can express information source lexically, and a number of them convey it through grammatical means. While recognising that the distinction between lexical and grammatical linguistic elements is not necessarily straightforward (see, for example, Boye’s (2018) discussion on the question of the meaning domain vs. grammatical coding in the context of evidentiality), in this study the elements I investigate fall under what Bybee et al (1994: 2) define as grammatical morphemes, or grams. Grams are function words<sup>4</sup> and affixes (Haspelmath 2003: 211), and can be contrasted with content words. They have “some unique grammatical behavior” (Bybee et al 1994: 2), including interactions with other linguistic elements.

Grammatical markers of information source, evidentials, are used for the purpose of effective communication; they are among the linguistic devices and cues the speaker

---

<sup>3</sup> In this introductory paper, the pronoun “she” is used for the speaker, except for in cases where “he” occurs in the source of the example.

<sup>4</sup> Function words also include auxiliaries, but they are for the most part excluded from this study (see, however, Greed 2014 (2016): 84).



can use to signal to the addressee her relationship to the information she wishes to convey. According to Grice (1989: 26–27), successful communication is achieved by following the Cooperative Principle. Sun (2018: 59) relates this Gricean principle to the use of evidentials: “a reliable and informative speaker must provide the addressee with appropriately chosen evidentials”. The speaker needs to take into account the addressee’s knowledge and position, expressing herself in a way appropriate to this and the speech situation.

The participants of a speech situation adhere to the Cooperative Principle when the following maxims are in operation: 1) Quantity – the appropriate amount of information; 2) Quality – the information should not be false, and the speaker should not say what she does not have adequate evidence<sup>5</sup> for; 3) Relation – the message needs to be relevant in the speech situation; 4) Manner – the information needs to be clearly expressed and easily understandable (Grice 1989: 26–27). Of these, maxims 1 to 3 show what the information should be like, and maxim 4, how it should be expressed. Whereas the use of markers to indicate information source would seem, for the most part, to fall under maxim 2, in that an evidential marker specifies the origin of the message and thus contributes to its effective quality, it also relates to the “clear expression” of maxim 4. Sun (2018: 62) sees evidentiality as ultimately having to do with “knowledge packaging and sharing, as the speech-act participants cooperate to achieve effective verbal communication”. The focus of this study is not on the speaker-addressee interaction and the social functions it can convey (see Gipper 2011), but on the viewpoint of the speaker. However, I briefly touch on this promising new interactional angle of investigation, advocated by Gipper and others, in my discussion of (inter)subjectivity and multisubjectivity (see Greed 2018: 27, and Section 6.3.1 of this introductory article).

## **2.2. Evidentiality: history and different viewpoints**

The earliest written records of evidentiality as a linguistic phenomenon go back almost 2,500 years to the 4<sup>th</sup> century BC philologist Pāṇini, who differentiated the Indo-Aryan Sanskrit perfect from the other verb forms by its use for past actions which were not

---

<sup>5</sup> Grice’s use of the term “evidence” appears to combine both the notions of information and its truth value (“proof”), thus relating to epistemic expression. This differs from the term “evidence” used in more recent evidential studies, where it is seen a synonym for “information” and “knowledge”.

witnessed but were reported by the speaker (Cardona 2003: 139). In the 11<sup>th</sup> century AD the lexicographer Maḥmūd al-Kāšyarī mentioned cases when the Old Turkic preterite, usually formed with the suffix *-di*, could also have the suffix *-miš*: the form with *-di* was used when “the action occurred in the presence of the speaker”, whereas *-miš* indicated that “the action occurred in the absence of the speaker” (Dankoff 1982; al-Kāšyarī: 412).

Boas is credited as the originator of the labels “evidence” and “hearsay” in the context of his research on the Northern Wakashan Kwakiutl language in the early 20<sup>th</sup> century (see Jacobsen 1986: 3). In his discussion on how obligatory aspects of grammar vary across languages, Boas (1938: 133) mentions languages where the source of information – “whether seen, heard, or inferred” – is among such aspects. Later, Jakobson (1971: 135) introduces a tentative label “evidential” for a verbal category referring to the source of information about a narrated event. This source can be someone else’s report (Jakobson calls this “quotative” and “hearsay”), a dream (“revelative evidence”), a guess (“presumptive evidence”), and the speaker’s previous experience (“memory evidence”). For Jakobson, this category is separate from mood (Jacobsen 1986: 5).

In the second half of the 20<sup>th</sup> century studies of evidentiality gained momentum, with a variety of definitions and positions, some of the main ones being that evidentiality has to do with the source of information of the message conveyed by the speaker (Aikhenvald 2004), or with the type of justification the speaker has available for making her claim (Anderson 1986: 274). An ever increasing number of languages were documented as expressing evidential meanings through grammatical means, including languages spoken in South America, the Balkans, and the Caucasus; such features were also found in Tibeto-Burman and Turkic languages. Research on evidentiality has also gathered pace in Africa, and there is increasing evidence that this continent, which had been regarded as lacking evidential expression, does indeed use it. Current research focuses on the Sudanic belt and southern Africa (Storch 2018: 612). Güldemann’s (2008) extensive synchronic and diachronic survey in African languages focuses on one particular type of (evidential) marker, the so-called quotative indexes: “linguistic forms signalling the presence of reported speech” (Güldemann 2008: 1). In Eurasia, the main area of interest for the current study, evidentiality is regarded as an areal feature, but the discussion continues as to where the phenomenon is the result of language-internal

transmission, and where it is due to diffusion through language contact, or perhaps the interaction of both language-internal and language-external factors.

Some of the major works discussing evidentiality and related categories in the world's languages are Chafe and Nichols (1986); Willett (1988), offering an early cross-linguistic survey of evidentiality; Johanson and Utas (2000), concentrating on Turkic, Iranian and Uralic languages; Aikhenvald's (2004) monograph on the grammatical category of evidentiality; and a recently published handbook of evidentiality (Aikhenvald 2018). In French linguistics the phenomenon is discussed in the context of *médiativité*, mediativity (Guentchéva 1996; Guentchéva & Landabury 2007), which refers to indirect evidentiality. In the Russian-speaking context, evidentiality is also called *zasvidetel'stvovannost'* 'witnessedness', and the indirect type can be referred to as *zaglaznost'* 'absentiveness'. Evidentiality has been discussed in many recent grammatical language descriptions especially in Caucasian, Finno-Ugric and Turkic languages, and also in a monograph edited by Xrakovskij (2007).

In her seminal monograph on evidentiality, Aikhenvald (2004) defines evidentiality as grammatical expression of information source and establishes six basic semantic parameters employed in languages for expressing evidential meanings. Aikhenvald's (2014: 9) semantic parameters are visual, (non-visual) sensory, inference, assumption, reported, and quotative. Some of the terms, for example, "visual", show reception of information through physical processes, others, such as "inference", involve mental processes. Such terms as "hearsay" and "quotative", concern another person's speech which is relayed by the recipient. Other common terms used for evidential meanings are witnessed vs. non-witnessed, direct and indirect, and non-firsthand and firsthand. Table 1 displays Aikhenvald's (2004, 2014) semantic parameters and shows that the first two parameters fall under "firsthand", whereas the remaining four are subsumed under "non-firsthand" meanings. In addition, the definitions of the parameters are given according to Aikhenvald (2014: 9).

TABLE 1. Aikhenvald’s (2004, 2014) evidential semantic parameters and their definitions

| Supercategory                                     | Firsthand               |  | Non-firsthand                                   |  |   |   |
|---|-------------------------|--|---|--|---|---|
| Aikhenvald’s (2004) parameter                     | <i>Visual</i>           | <i>Sensory</i>                                     | <i>Inference</i>                                | <i>Assumption</i>  | <i>Reported</i>   | <i>Quotative</i>  |
| Definition: Type of evidence (Aikhenvald 2014: 9) | Acquired through seeing | Acquired through hearing (and smell, taste, touch) | Based on visible or tangible evidence or result | Based on evidence other than visible results: logical reasoning, general knowledge | Reported information with no reference to the original source | Reported information with an overt reference to the quoted source |

In recent linguistic literature questions have been raised as to whether the quotative can be included amongst evidentials. Boye (2018: 264–265) focuses attention on the difference between the reportative (reported) and the quotative and points out that whereas the reportative has scope over propositions (“meaning units that can be said to have a truth value”), the scope of the quotative is over illocutions<sup>6</sup>. This point of view is followed in my study, and the quotative is defined as an illocutionary quotative (see Table 3, and a fuller discussion in Section 3.5).

Willett (1988: 54) suggests a three-way division of evidentials into those which convey sensory, reported and inferential evidence. Similarly, Plungian (2001), initially distinguishing between direct and indirect evidence, also proposes a three-way distinction of evidential values: direct evidence, reflected evidence and mediated evidence, the latter two being a subdivision of indirect evidence. He also introduces the term “personal evidence”, which occurs with both direct evidence and reflected evidence of the indirect type, as shown in Table 2.

---

<sup>6</sup> Due to its illocutionary scope, Boye (2012: 32) regards quotative as non-evidential and non-epistemic. The investigation of the controversial status of the quotative falls outside the scope of this thesis.

TABLE 2. Different types of evidential oppositions (from Plungian 2001: 353)

|                          |   |  |
|--------------------------|---|--|
| <b>Direct evidence</b>   | <b>Indirect evidence</b>                                    |  |
|                          | <b>Reflected evidence</b><br>(= Inference and Presumptives) | <b>Mediated evidence</b><br>(= Quotatives) |
| <b>Personal evidence</b> |   |  |

Plungian’s direct evidence coincides with Aikhenvald’s visual and sensory parameters, indirect reflected evidence with inference and assumption, and mediated evidence with Aikhenvald’s reported. While Aikhenvald lists the six evidential parameters as being on a par with each other, and helps the researcher to straightforwardly pinpoint the different evidential values, Plungian displays these in a manner of subdivisions, links and connections, showing how they relate to one another. This is helpful when investigating how different factors, such as cognition, general knowledge, and the distance of the speaker from the original knowledge, to name but a few, play their part in a speech situation.

In his description of evidentiality in Turkic languages, Johanson’s (2003: 274) view on evidentiality is centred around the concept of “indirectivity”, which coincides with the notions of hearsay and inference used in other theories. The focus is not so much on the source of information, but on the way “a conscious subject” receives knowledge of a narrative event. Such knowledge may be received through reported speech or hearsay (“reportive use”), inferred on the basis of reflection, or as a logical conclusion (“inferential use”). In Johanson’s third type of possible reception, “perceptive use”, it is possible to receive knowledge through direct sensory perception, and this use also includes “indirect perception” based on traces of an event. Thus, the first type comes close to Aikhenvald’s “reported” use, the second one overlaps with “assumption”, and the third one is close to Aikhenvald’s “inference”. However, as inference can also have a sensory component, it overlaps with the semantic parameter of “sensory”.

For Aikhenvald, the point of departure in investigating evidential phenomena is to regard evidentiality as a grammatical category. However, in practice, as she defines evidentiality through the semantic notion of information source (see Aikhenvald 2004: 1),

it is viewed from a functional angle. The functional viewpoint is more pronounced in Cornillie’s (2009: 45) approach. He sees the functional domain of evidentiality as being present in most languages, which opens up the study of evidential meanings from languages with obligatory grammatical evidentiality to languages where evidentiality is “an optional linguistic category”. Evidentiality can, thus, be researched both from the point of view of the forms that express grammatical meanings, and, at the other end, from a semantic viewpoint. The latter viewpoint can include the investigation of lexical expression of evidential meanings.

This brief overview outlines some of the main approaches to evidentiality, which have been influential in my own study. Table 3 is my presentation of these views in an integrated way.

TABLE 3. An integrated display of evidential parameters based on Plungian (2001), Aikhenvald (2004) and Johanson (2003)

| <b>Direct</b><br>(information source/<br>reception of<br>information) |                | <b>Indirect</b><br>(information source/<br>reception of information) |   |                                    |                                     |
|---|----------------|--|---|------------------------------------|-------------------------------------|
|   |                | <b>Cognitive/Reflected</b>   |   | <b>Mediated</b>                    |                                     |
| <i>Visual</i>   | <i>Sensory</i> | <i>Inference</i><br>(conclusion from tangible evidence)              | <i>Assumption</i><br>(conclusion from reasoning, general knowledge) | <i>Reported</i><br>(propositional) | <i>Quotative</i><br>(illocutionary) |

As already mentioned, I base my understanding of the category of evidentiality on Aikhenvald (2004), and make use of her six parameters of evidential meanings (marked in italics in Table 3), as they are straightforward and amenable to comparison. While information source is regarded as the key meaning involved in evidentiality, Johanson’s perspective of viewing the evidential process from the angle of a conscious subject, and their reception of information, is also helpful. Both these angles play a role in distinguishing between the indirect and direct evidential meanings, as indicated in Table

3. Plungian, for his part, makes an explicit distinction between evidential types connected with human cognitive processes and those outside these processes, visible in the table in the division into “reflected” and “mediated” types.

Cornillie’s approach of investigating all possible evidential meanings, including lexical ones, and independent of whether evidentiality is an obligatory linguistic category, is reflected in this study in that a number of markers which are investigated do not have evidential meanings as their default meanings but as extensions of other (grammatical) meanings. In Aikhenvald’s (2004) approach such markers are not evidentials proper, but are explained as instances of evidential strategies (Aikhenvald 2004: 105).

As stated, the linguistic elements investigated in this thesis are limited to grammatical ones. The majority of markers examined are affixes and particles, including some more complex constructions. The viewpoint of my study is functional, and the investigation concentrates on the inherent, pragmatic and contextual meanings of the grams.

### **3. Evidentiality in relation to other categories and domains from a theoretical perspective**

Evidential meanings can be expressed across languages through a variety of grammatical means. Among these are affixes, particles or clitics, and special verb forms (Aikhenvald 2004: 67). Evidentials can interact with other categories both on the level of form, that is, in their morphosyntactic expression, and on the level of meaning (see Forker (2018a: 66) for possible types of interaction). The primary type of interaction observed in this study is on the level of meaning. In the following, I will give an overview of the main categories and domains that are relevant for the study of the four languages of the thesis. These are tense-aspect (3.1), mirativity and person (3.2), epistemic modality and subjectivity (3.3), discourse (3.4), and scope (3.5). At the end of this section I will outline the concept of semantic space (3.6) and the methodology of semantic maps, as this will help us in defining how the four languages we are investigating are located in semantic space. This methodology is put into practice in two semantic maps (Figures 1 and 2) illustrating the findings of this study.

### 3.1. Evidential meanings as extensions of tense-aspect

As verbs “encode information about events and situations” (Forker 2018a: 65), it is not surprising that evidentiality is among the categories frequently expressed with verbs across languages. Amongst verbs, past tenses are typologically attested as being more prone to expressing evidential meanings than other tenses. In a number of languages the expression of evidential meanings is fused with tense, and therefore one grammatical marker carries both these meanings. This is the case in Tatar, where the categorical past tense can express a witnessed meaning, and the perfect<sup>7</sup> form conveys a non-witnessed meaning, as in examples (1) and (2).

- (1) Rišat qajt-ty.  
Rišat return-PST.3SG  
Rišat returned. (Greed 2014 (2016): 72)
- (2) Rišat qajt-qan.  
Rišat return-PRF.3SG  
Rišat returned/has returned. (Greed 2014 (2016): 72)

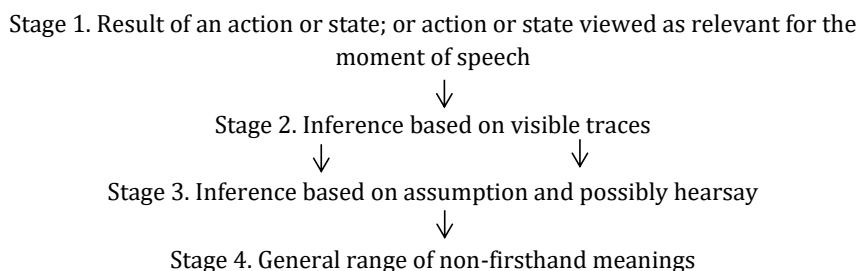
Example (2) is a case of the perfect form carrying an evidential meaning. It is a cross-linguistically common feature for the perfect form to acquire meanings of indirect evidentiality, as discussed in Greed (Article 4). A tense-aspect form perfect expresses a past action which has continuing relevance for the present (Nedjalkov & Jaxontov 1988: 15). One of the main evolutionary paths attested for the resultative meaning is the progression to the evidential meaning of inference (see Bybee et al. 1994: 104–105), and on to other non-firsthand evidential meanings, or the development through the perfect meaning to perfectives or simple pasts. In the path of development outlined by Aikhenvald (2004: 116), as shown in Scheme 1, the perfect meaning acquires the meaning of inference, both meanings expressing the speaker’s viewing an event after it has happened, from its discernible traces. The inferential meaning can then develop into the meaning of assumption, as has happened with the Aghul past resultative general factual form, described by Majsak and Merdanova as expressing “weakened” inference (2002: 7–

---

<sup>7</sup> In my original article (Greed 2014 (2016)), this form was called “resultative (perfect) past”. In this introductory article some terms and glosses of examples have been modified to better fit the general discussion.



8; see Greed 2017: 12). The inferential meaning can also take a different path and advance to the meaning of hearsay (reported), as will be shown below in the discussion on Even. The final stage of the development is a general range of non-firsthand meanings.



SCHEME 1. Evidential extensions for perfects and resultatives (modified from Aikhenvald 2004: 116)

Whilst the evolution of the languages of this study is in general similar to those outlined above, the discussion below will present a more precise scheme for them.

### **3.2. Evidentiality, mirativity and person**

One of the categories attested in conjunction with evidentiality is mirativity. As the category has been found to interact with person in a number of languages, especially in inferential contexts, it is discussed here together with person. Mirativity “refers to the marking of a proposition as representing information which is new to the speaker” (DeLancey 2001: 369). Besides newness of information, mirativity can express unexpectedness and lack of consciousness, and the meaning of surprise may be a further interpretation that a mirative marker can receive. Forker (2018a: 83) sees mirativity as a “side-effect” of evidentiality, and suggests that it is an independent category from a semantic point of view, although cross-linguistically only occasionally grammaticalised. Overall, this approach fits the languages investigated in this thesis, since examples of a mirative meaning occurring as an extension of an evidential meaning (see section 6.4), as well as an example of a mirative meaning expressed independently of evidential meanings

were found. In addition, Lezgi has at least one particle with a default mirative meaning (see section 6.5).

Mirativity and person can interact in certain contexts of evidential expression in a special way. This happens in cases where inference combines with first person. As the speaker is usually aware of her own experiences and actions, the inferential meaning is not possible in first-person contexts. However, if an inference does occur, this meaning receives a mirative interpretation, indicating that the speaker is observing the (results of an) event after it took place, and that she was not conscious of it happening. Such an action was unintentional and involuntary (Curnow 2003: 42). Example (3) from Aghul is from a first-person context, where the speaker realises that he has mixed up the seeds he had been sowing. He says,

- (3)    zun    sul    uzu-naa,    χabar=ra    a-dawa-j!  
      I     rye    sow-PRF    message-ADD   be\_in-NEG-CVB  
      (Apparently) I have sown rye, without noticing it myself! (Majsak & Merdanova  
      2016: 399)

The perfect form with *-naa* expressing inference receives an additional mirative interpretation of unexpectedness and surprise (Majsak & Merdanova 2016: 399).

The first-person effect in conjunction with inference is not the only interaction that was attested with the markers investigated in the articles of this thesis. A verbal particle *ikān* in Tatar with the default meaning of “speaker-oriented subjectivity” (Greed 2014 (2016): 80), acquires both a mirative and an evidential extension also in third-person contexts. In addition, one of the main discoveries of Article 4 was the occurrence of the mirative interpretation of inference not just in first-person but also in second-person contexts, as discussed in section 6.5.1 of this introductory article. These discoveries indicate that the interaction of evidentiality, mirativity and person is intricate and complex, exhibiting features not found in the general discussion thus far.

### **3.3. Evidentiality, epistemic modality and subjectivity**

When discussing the relationship of evidentiality with other categories in the semantic space, a key category is epistemic modality, as both categories relate to knowledge and

belief. Subjectivity, for its part, is to do with the consciousness or perspective of the speaker, and is frequently discussed as part of modality. Some linguists, especially in the earlier days of evidential studies, have argued that evidentiality is part of epistemic modality (see, for example, Palmer 1986), whereas Chafe and Nichols (1986: vii) see evidentiality as being about both the source and reliability of knowledge, and incorporate such meanings under epistemology. According to Boye (2012: 1) the three basic positions with regard to how evidentiality and epistemic modality relate to one another are as follows: 1) the categories are separate, although potentially related; 2) the categories are distinct but overlap; and 3) one of the categories comprises the other. Boye (2012: 2) himself adopts the view that these two categories are distinct, and proposes that together they should be subsumed under a category which he calls epistemicity.

Boye (2012: 15–18) finds support for this position in philosophy, where ‘knowledge’ is defined by the belief, truth, and justification conditions. If a subject is to know a proposition, the conditions a) that she believes it, b) that the proposition is true, and c) that the belief is appropriately justified need to be in place (Williams 2001: 16, cited by Boye 2012: 16). While the belief and justification conditions are to do with the actual reasoning by the subject, the truth condition links knowledge with truth, thereby connecting the reasoning with the reality of the world. Thus the notion of truth falls outside of epistemic notions. The notion of belief is linked with certainty, necessity and possibility, that is, notions usually described within the context of epistemic modality. The notion of justification connects with evidence (Boye 2012: 17).

The categories of epistemicity, and its subcategories epistemic modality and evidentiality are for Boye theoretical descriptive categories used to make generalisations over language-specific linguistic meanings (Boye 2012: 3). While keeping with the general meaning assigned to the basic evidential meaning, that is, information source, Boye prefers the term “epistemic justification”, and for epistemic modality his term is “epistemic support”. The term describing the supercategory epistemicity, which contains both evidentiality and epistemic modality, is a generalisation – and a combination – of epistemic justification and epistemic support, called “justificatory support” (Boye 2012: 18–19).

Boye's proposal to link the categories of evidentiality and epistemic modality by incorporating them under the category of epistemicity is in general in line with my own experience of the evidential phenomena which I have encountered and researched in a number of languages. It is appropriate to regard evidentiality and epistemic modality as subcategories of a more general category, which all pertain to different aspects connected with knowledge in a speech situation: the source of knowledge, as expressed by the speaker, and the attitude or evaluation of this knowledge that the speaker communicates. However, since the term "epistemic justification" proposed for information source would appear to imply truth-related commitment by the speaker to the proposition conveyed (as Boye himself comments (2018: 264)), I use the established terms: the basic evidential meaning conveys information source, and the meanings expressing the speaker's evaluation of, or attitude towards the message are called epistemic or modal (see Aikhenvald 2014: 9; 2004: 65).

A newer aspect in the study of evidentiality is the notion of subjectivity, usually discussed in the context of modality. Narrog (2012: 41) sees three basic positions in understanding the notion: 1) Subjectivity can be understood as referring to 'speaker involvement' or 'speaker commitment' in contrast with objectivity. This view is followed by the majority of scholars. 2) Subjectivity can also be seen in terms of evidentiality, and in this case it is contrasted with intersubjectivity, the former having to do with knowledge relating only to the speaker, and the latter to knowledge shared by the speaker and the hearer (Narrog 2012: 41). Nuyts (2001: 393), a proponent of this view, states that in the case of a subjective expression, the speaker "assumes strictly personal responsibility for the epistemic qualification". 3) In the framework of cognitive linguistics (for example, Langacker 1990) subjectivity is understood as "the degree to which the speaker (conceptualizer) is expressed only implicitly as opposed to putting him- or herself on stage" (Narrog 2012: 41). Accordingly, for Mushin (2001: 1) linguistic subjectivity is "the interpretation of linguistic expressions in terms of some cognizant observer, thinker, emoter, and, of course, speaker". The speaker/experiencer can use a variety of linguistic expressions to convey different viewpoints and manipulate them to convey varying degrees of subjectivity.

Mushin (2001: 2) links subjectivity with the expressive function of language: the speaker, as well as the addressee, has the ability to express aspects of their subjective self in a particular speech situation. Even a statement of a fact sounding “objective” is an expression of the speaker’s desire to inform. In this study subjectivity is seen as the expression of the subjective self. The speaker or experiencer is involved in the cognitive processing of the information, and she makes use of different linguistic possibilities to convey this involvement, which can be of different degrees.

The key area where subjectivity becomes relevant in this thesis is in the context of complementation, semi-direct speech and the quotative, as discussed in Bashkir (Greed 2018: section 5; section 6.3.1 of this introductory article) and in Tatar (Greed 2014 (2016): 76–77). A subjective interpretation arises from the interaction of the quotative and the complement-taking verbs. Verbs which take a complement are from a limited semantic set, such as verbs of perception and cognition. Common to these verbs is that they “relate to the nature of the human mind, [and] the ways in which information is coded and communicated” (Dixon 2006: 2). When the Bashkir quotative *tip* marks semi-direct speech as a complement clause, the interpretation conveys different levels of subjectivity, depending on the verb in the matrix clause. In the case of Bashkir, subjectivity was discovered to extend further than to the speaker/experiencer in the context of the verb of cognition ‘know’, and the verb of perception ‘hear’: the marker *tip* coded also the originator(s) of the speech report, thus merging the two perspectives. For this phenomenon, I introduced the term “multisubjectivity”, which refers to the involvement of other subjective selves in addition to the speaker/experiencer in the processing of information, and will be discussed in section 6.3.1.

In my study I view subjectivity as an independent semantic dimension which can overlap (at least) with evidential and modal meanings. It is essentially speaker-oriented: if something is marked as subjective, it shows that what is conveyed is connected with the consciousness or perspective of the speaker. This understanding has some overlap with Nuyts’ (2001) definition of subjectivity, especially in the aspect of the speaker taking responsibility for the information conveyed. Subjectivity does not pertain to assessing the truth value of the information conveyed, or its epistemic evaluation. However, such meanings can arise contextually in conjunction with subjective meanings.

### 3.4. Evidentiality and the domain of discourse

In addition to being used in everyday communication to convey information source, evidentials can cross over to a different domain and acquire new functions that may be a reflection of its earlier meanings (evidential or otherwise), or they can even be assigned uses that seem to be contrary to their earlier meanings. One such domain is discourse<sup>8</sup>. Evidential markers may develop special functions in different discourse genres. For example, the Tatar past-tense form with *-DY*, which can receive the interpretation of witnessed, has developed a use in the narrative genre whereby it signals realistic narration and is used for foreground events that move the story line forward, that is, it has a plot-advancing function.

On the other hand, the evidential counterpart of *-DY*, the perfect with *-GAN*, has come to be used as a “genre token” (see Aikhenvald 2004: 310) of traditional narrative, including fairy stories and folk tales. While in terms of evidentiality the form still displays its earlier general meaning of non-witnessed/reported, with regard to discourse it has switched from the backgrounding function deriving from its perfect nature to a very different plot-advancing function. Such a use of perfect in the narrative function is widely attested in the context of traditional narratives. What is of interest is the fact that a form characteristically non-narrative, as it is detached from other past facts, has lost its perfect nature and is used as a plot-advancing narrative form (see Lindstedt 2000: 365–366), a function similar to that of the *-DY* form in realistic narrative.

Once a marker has become a genre token, it becomes a determining factor for a piece of discourse, and can override other conventions of the language (see Aikhenvald 2004: 331). Within this chosen framework other verb forms may come to be used in a new function, not characteristic to them.

### 3.5. Evidentiality and scope

In addition to discourse, another domain to consider in the context of evidentiality is semantic scope, that is, how widely the evidential meaning of a grammatical form affects the stretch of language with which it occurs (see Crystal 2007: 407). The scope of an

---

<sup>8</sup> I define discourse as a stretch of communication, organised and following (linguistic and literary) conventions characteristic to the type or genre in question (see Dooley & Levinsohn 2001: 3–4). A discourse can be, for example, a story or a conversation.

evidential marker can be, for example, a word or a clause. In actual language use, what have evidential semantic parameters (visual, sensory, inference, assumption, hearsay and quotative) scope over? Boye (2018: 264–265) points out that with evidentials like the reportative (hearsay), the scope is over a proposition, by which is meant a unit of meaning conveying a speaker’s representation of a situation in the world, which also has a truth value. For the quotative evidential, the scope is over an illocution, that is, over an utterance made by a speaker which in itself constitutes an act performed by her, such as promising (see Crystal 2007: 226). In the context of this thesis an illocution relates to a speech report, an utterance by an original speaker which is reported verbatim by the current speaker. As the scope of the quotative marker is over an illocution, in Boye’s view the quotative falls outside of evidential meanings. I acknowledge this difference, but since the quotative marks information source – that someone else’s speech is relayed – in this thesis I include it in the discussion, calling it an “illocutionary” evidential (see section 6.3), while the others are “propositional” evidentials (see also Speas’s (2018) discussion on proposition-level and illocutionary-level evidentials).

### **3.6. Evidentiality in semantic space**

We have looked at how evidentiality relates to and interacts with other categories and domains. This leads to the question: Where in the universal semantic space of grammatical meanings do evidential (and related) meanings fit in?

When discussing the place of evidentiality in the universal semantic space, Plungian’s (2001) point of departure is that grammatical values of different languages need to be comparable for typological study to be possible. Languages are thought to share semantic substance, and some of this substance can become grammaticalised. Universal semantic space consists of “semantic elements that are grammaticalized ... at least in one natural language” (Plungian 2001: 351). For Plungian (2001: 350), semantic space includes both the “established inventory of cross-linguistic values” and “a structure to account for their interrelations”, a semantic map. This offers a cross-linguistic view to evidentiality, whereas within the context of each individual language the semantic domain of evidentiality looks different, both in terms of values and distribution.

Georgakopoulos and Polis (2018: 1) describe a semantic map as “a way to visually represent the interrelationships between meanings in languages”<sup>9</sup>. The two basic types of semantic maps are classical maps and proximity maps. A classical semantic map is typically displayed as a graph, and it is a “geometrical representation of functions in ‘conceptual/semantic space’ that are linked by connecting lines and thus constitute a network” (Haspelmath 2003: 213). In addition to the obvious benefit of visualisation, if compared with, for example, just listing all the functions of a gram, a semantic map places the set of functions of a gram as “a coherent chunk of universal network” (Haspelmath 2003: 213).

Of the four languages of this thesis the basic evidential values/functions and their markers were established for Tatar (see Section 6.1), and Lezgi (Section 6.2), and the relevant Lezgi parameters and their grammatical markers are displayed in Section 6.2. For Bashkir, I presented a brief overview of Bashkir evidentiality (see Greed 2018: Section 2) in order to prepare the ground for a thorough investigation of one of the evidential parameters, the illocutionary quotative. For the quotative, different functions were established, and their interrelations with other semantic domains were investigated. For Even (Greed 2019), the discussion centred on the path of development of the grammatical meaning of the perfect form. A semantic map comparing the perfect marker in the languages under consideration in this thesis, including Aghul, is provided in section 6.1, Figure 1.<sup>10</sup> The discoveries of the study will be drawn together and displayed visually in a modified semantic map in Figure 2 of Section 6.7.

#### **4. Abstracts of articles**

This section presents the abstracts of the four articles of the thesis. These will be followed by a fuller discussion of the topics I outlined in the first part of the introductory chapter.

---

<sup>9</sup> See also the introduction (Cysouw et al. 2010), and the contributions to *Linguistic Discovery 8:1*, 2010 for approaches and methodology for semantic maps.

<sup>10</sup> I am grateful to Andrej Malchukov for suggesting the inclusion of more extensive coverage of semantic maps in the discussion.



Each topic will be presented in its relevant contexts within the study which I conducted in the four languages of this thesis.

Article 1. Greed, Teija 2014 (2016). The expression of knowledge in Tatar. In: Aikhenvald, Alexandra Y. & R. M. W. Dixon (eds). *The Grammar of Knowledge*. Oxford: Oxford University Press. 69–88.

The article investigates how knowledge is expressed in (Volga) Tatar, a Kipchak-Bulgar Turkic language spoken by the second largest people of the Russian Federation. The discussion has two viewpoints: that of evidentiality, and that of epistemic modality. The former pertains to the expression of information source by the speaker through grammatical means, and the latter to the conveying of attitudes towards the message. In addition to grammatical expression of information source, ways of lexical expression are also investigated.

Evidential meanings are expressed grammatically in Tatar with past-tense verb suffixes and grammaticalised particles. The past-tense verb form perfect (resultative) with *-GAN* can acquire an evidential non-firsthand meaning, and while the categorical past with *-DY* is in general evidentially neutral, it can receive a firsthand interpretation. Thus the two forms can form an evidential opposition of non-firsthand–firsthand. Other semantic parameters of evidentiality relevant to Tatar are inferred, assumed, reported/hearsay and quotative. Of these, firsthand (neutral) stands alone, whereas the other parameters are subsumed under the general parameter of non-firsthand.

The grammaticalised particles conveying evidential meanings are *dip*, *di* and *ikän*. For the markers *dip* and *di*, their default meaning is an evidential one, which makes them “evidentials proper” (see Aikhenvald 2004: 105). Evidentiality was found to interact with mirativity and modality. With the past-tense form in *-GAN*, the mirative meaning of ‘unprepared mind’ occurs as an extension of the inferential meaning and arises only in first-person contexts. The interaction between evidentiality, modality and mirativity is at its clearest with the multifunctional particle *ikän*, which was discovered to have a subjective modal meaning as its default meaning. This meaning can receive both a mirative extension, and an evidential extension of assumption in declarative clauses, while interrogative clauses convey a speculative meaning. In addition to *ikän*, the

quotative *dip* conveys the speaker's subjective viewpoint. No clear connection was found between non-firsthand expression of evidentiality and the evaluation of the truth value of the information conveyed.

Tatar evidentials play an important part in literary genres: the verb suffixes *-GAn* and *-DY* and the particle *di* function as markers of genre. The Tatar past-tense form *-GAn* follows the cross-linguistically established path of development of evidential meanings: the evidential meanings of inference and reported have developed from the resultative meaning. As for the quotative particle *dip*, originating from the converb form of the speech verb *dijü* 'say', I suggest that it has through grammaticalisation first become a marker of direct speech, and then developed to also mark indirect speech, and extended further to functioning as a clausal complementiser and clausal connector, signalling the meaning of goal and reason.

For lexical expression of information source, Tatar uses verbs of perception and cognition and various more extensive expressions. The speaker can convey her attitude towards the information on the scale of impossible–unlikely–uncertain–possible–likely–certain, or her evaluation of its reliability with the help of full explanatory clauses, as well as with modal verbs, particles, clitics and predicative words.

Article 2. Greed, Teija 2017. Evidential coding in Lezgi. *Languages of the Caucasus*. Volume 2. Available at <https://escholarship.org/uc/item/0w59z5t4>.

Evidentiality is “grammatical marking of how we know something” (Aikhenvald 2014: 3). As evidentiality is a well-known feature in many Nakh-Daghestanian languages, spoken in the Caucasus region of Russia and Azerbaijan, this paper investigates the expression of evidential meanings in Lezgi, a language which has received less attention in this area. As a closely related language Aghul has been shown to express evidential meanings through verb forms (Majsak & Merdanova 2002), these findings were the motivation behind the current investigation: does Lezgi express evidentiality in a comparable way to Aghul? In addition, the paper sought to discover where Lezgi is located within the semantic space of evidential and other related meanings.

Regarding indirect evidentiality, the Lezgi Perfect was found to display the meaning of inference, thus largely coinciding with the inferential use of the Aghul Perfect.

In addition, a verb construction involving a nominalised predicative and an equative particle conveys the meaning of inference accompanied by epistemic uncertainty. The Lezgi Aorist is in general evidentially neutral, but can acquire readings of direct, witnessed information source, accompanied by the modal meaning of reliability. However, the opposition displayed in Aghul between witnessed and reported meanings does not find reflection in the Lezgi past tense forms.

On the widely-attested cross-linguistic path of development for the resultative/perfect meaning (see Bybee et al. 1994: 104–105) Lezgi has moved to the stage of inference, whereas Aghul has progressed further to express also the non-witnessed meaning. Since the two languages are closely related, the reasons for these differences of evolution were considered. Both languages have been in long-standing contact with Turkic through the Azerbaijani language, so if the development of evidential meanings is due to diffusion, it may be that Lezgi, being one of the larger languages of the area, has a more conservative approach towards borrowing. An interesting example of clear borrowing from another Turkic language Turkish, discovered by Marianne Beerle-Moor (2016 p. c.) is the expression of non-witnessed and witnessed meanings with past-tense forms in the Kirne variant of Lezgi spoken in Turkey. It appears that in the Kirne dialect the evidential function has further extended to cover non-past forms, a development which has been attested less frequently cross-linguistically than with past forms (see Forker 2018a: 66; Aikhenvald 2004: 266).

Article 3. Greed, Teija 2018. The quotative in Bashkir. *Studia Orientalia Electronica*, Volume 6, 23–55. Available at <https://journal.fi/store/article/view/69685>.

Evidentiality is a widely researched category in contemporary linguistics. However, the illocutionary evidential quotative, which codes a speech report with an explicit reference to the quoted source, has received less attention than other markers with evidential meanings. This article investigates the quotative particle *tip* in Bashkir, a Kipchak-Bulgar Turkic language spoken in the Russian Federation. In its default quotative meaning, *tip* signals direct speech, and functions as a syntactic complementiser. The quotative function of *tip* was found to have extended to code not just spoken utterances but also thoughts and experiences in the context of semi-direct speech, a term introduced for Bashkir in this

study. A separate function of *tip* is its use as a marker of logical relation, conveying the meaning of intention or purpose.

In addition to teasing out the different functions of the marker *tip*, this paper offers a theoretical viewpoint to the interconnections of different categories that come together in the functions of *tip*. Semi-direct speech can occur with a certain set of complement-taking verbs which are mainly to do with cognition and perception. When combining with these verbs the meaning that the particle *tip* conveys was discovered to be linked with a semantic dimension of subjectivity. Subjectivity pertains to the cognitive processing and expressing of information by the speaker/experiencer. The marker *tip* was investigated in conjunction with ten complement-taking verbs, and the degree and strength of subjectivity was found to range from neutral expression with the verb of perception *ujlau* 'think' to strong subjectivity with the verb *qurqyu* 'be afraid'. With the verbs *beleü* 'know' and *išeteü* 'hear' *tip* was discovered to convey a multisubjective meaning: in addition to signalling what the experiential subject has heard or found out, the marker also codes the involvement of some other subject, the original source. Thus the quotative *tip* gives voice to multiple speakers and merges them.

Article 4. Greed, Teija 2019 (in press). From perfect to narrative tense: The development of an evidential meaning examined generally and in the Even language. *Studies in Language* 42:4, 923–966. Will be available via <https://www.benjamins.com/catalog/sl>.

From the earliest attested description of evidential meanings of verb forms, the perfect has been one of the central forms. This use is widely attested in Eurasia, including in the North-Tungusic language Even. This paper discusses the development of the perfect meaning into evidential meanings, and further into coding a narrative tense function, first in general, and then concentrating on Even, an endangered language whose speakers live scattered in north-eastern Siberia. The three main Even dialects and their subdialects form a dialect chain with up to 20 different variants. In the area of evidentiality the dialects have experienced varying degrees of contact influence from Sakha, with the western and central dialects being in direct contact with it, whereas the eastern dialects, including the literary Ola dialect, fall outside this influence.

The verb form carrying evidential meanings in Even is the perfect form, which originates from a perfect participle and needs copula support in first- and second-person contexts. Malchukov (2000) has shown that the meanings of the perfect form are dialect-dependent: the eastern dialects do not convey evidential meanings, whereas moving towards the west, the number and use of evidential meanings increases. However, the semantics of the Even perfect is somewhat ambiguous, as isolated occurrences of the perfect meaning can be found in the eastern dialect. In addition, in Levin's (1936) description of the eastern Ola dialect, as presented in this paper, the finite perfect forms convey both evidential and mirative meanings.

The western dialect has developed the most elaborate variety of evidential meanings in Even. In addition to the inferential meaning of the Even perfect discussed by Malchukov (2000), the non-witnessed meaning of the perfect form was attested in the current study. In western and central Even the non-witnessed meaning has developed into a genre token, a form signalling a narrative function, and used in discourse types referring to events detached from the moment of speech. The resulting meanings of non-witnessed and reported do not convey any epistemic overtones to do with reliability, but the stance of the speaker is neutral and objective. A genre token chosen by a narrator can override other conventions of the language (Aikhenvald 2004: 331–332), which was shown to be the case in Even, where the chosen narrative form with -čÄ provides the framework for a story. The form used as a genre token loses its perfect nature and codes a plot-advancing function, and consequently other verb forms acquire usages subject to the main narrative form.

In Even the occurrence of the mirative meaning in conjunction with the evidential meaning of inference is connected with person. Inference receives the mirative reading in first-person contexts, which case is cross-linguistically fairly frequent. However, in this study the mirative meaning was found to extend to cover inferences also in second-person contexts. In both the first- and second-person contexts the speaker's experience and viewpoint is signalled by the form. Even appears to have acquired evidential meanings expressed with its perfect form through contact with the Turkic Sakha. However, as Malchukov (2000) suggests, this diachronic development has not happened spontaneously, or in isolation. Language-internal factors, such as the structure of the

language, other existing verb forms and their functions, and the overall linguistic situation within Even itself have played their part in the emergence and evolution of meanings.

## **5. A brief typological comparison of Tatar, Bashkir, Lezgi and Even**

The four languages studied for this thesis are all spoken in Eurasia, where evidentiality is regarded as an areal feature, some of the key locations of evidential expression being the Caucasus, the Balkans, and the locations where the typologically and areally Altaic languages (see Janhunen 2005: 37) are spoken, including the Turkic and Tungusic languages. Of the languages investigated (Volga) Tatar and Bashkir are closely related Kipchak-Bulgar Turkic languages, which are spoken mainly in Central Russia between the Volga River and the Ural Mountains. Tatars are the second largest people of Russia after Russians, and out of the ethnic population of 5.3 million, 4.3 million speak Tatar as their mother tongue. Bashkir is also one of the larger languages of Russia, with 1.2 million speakers out of an ethnic population of 1.6 million (Simons & Fennig 2018). Even is an endangered North-Tungusic language with up to twenty dialects whose speakers live in a vast area in north-eastern Russia, from the north-west of the Republic of Sakha to the Kamchatka Peninsula of the Russian Far East. Out of the 22,300 Evens, approximately 5,000 still speak their mother tongue (Russian Census 2010). Lezgi is a Nakh-Daghestanian language, spoken by 402,000 people out of 474,000 ethnic Lezgis in southern Daghestan in the Russian Federation. In addition, there are 350,000 Lezgis in Azerbaijan of whom 193,000 people speak Lezgi (Simons & Fennig 2018).

Typologically the four languages display a number of similar features in their phonology, morphosyntax and information structure. All four languages are morphologically agglutinative and use suffixes to denote grammatical categories. In syntax, they make extensive use of non-finite verb forms, such as converbs and participles, for clausal subordination. The neutral word order is AOV/SV, but for pragmatic purposes each language allows for word order variation. For example, the topicalised element of a proposition tends to occur sentence-initially, whereas the default position for a focal element is immediately before the verb (see, for example, Burgherr

2017 for Even information structure, and Desnitsky et al. 2002: 88–89, for Tatar). In phonology, the languages have vowel harmony (see, for example, Haspelmath 1993: 48–50 for Lezgi), and also display consonant assimilation to varying degrees. For expression of grammatical relations, Tatar, Bashkir and Even are nominative-accusative languages, while Lezgi differs from the others with its ergative-absolutive pattern.

The expression of evidentiality in the four languages is in general terms fairly similar. Within the context of verbs, each of them conveys evidential meanings as an extension of the tense-aspectual form perfect. The quotative is formed in a similar manner in all the languages, being a form of the speech verb ‘say’. However, the process of grammaticalisation of the quotative differs in the languages, with Bashkir and Tatar using a fully grammaticalised marker, whereas in Even and Lezgi the process is still ongoing, with different dialects displaying different stages of this development (see, for example, Donet 2014: 118 for the Quba dialect of Lezgi).

Nevertheless, each language displays its own idiosyncrasies, which fact is reflected in the different choices of topics for the articles, emphasising a particular key feature of the evidential expression in each language. In the following sections I will discuss some of the central topics researched in the context of the four languages with the aim of presenting the findings within the more general framework of the first part of this introductory article.

## **6. Evidentiality and other categories and domains in Bashkir, Even, Lezgi and Tatar**

Evidentiality interacts with other semantic categories in a variety of ways. The categories discussed in this section are tense-aspect (Section 6.1), epistemic modality (6.4), mirativity and person (6.5), the last two both interacting and being somewhat distinct in the languages researched. In addition, markers with a default evidential meaning are presented in Section 6.2. In Section 6.3, the semantic dimension of subjectivity is introduced, which became especially relevant in the study of the Bashkir quotative. I conclude the section with a discussion on evidentiality and discourse (Section 6.6), as this topic is central in two of the languages investigated, namely in Tatar and Even.

## 6.1. Evidentiality and tense-aspect

In the languages studied in this thesis, verbal expression of evidential meanings is the main way for conveying information source through grammatical means. This section discusses verbal expression of evidentiality with tense-aspect forms, whereas other evidential markers originating from verb forms are discussed in Section 6.4.

In the four languages verbal expression of evidentiality in finite contexts is, for the most part, to do with the interaction with the category of tense-aspect<sup>11</sup>, where one grammatical marker carries both the tense-aspectual and the evidential meaning, thus fusing them. Past tenses are reported as being more prone to expressing evidential meanings than other tenses (see Forker 2018a: 66). This is the case with all the four languages, as the evidential meanings are extensions of the default perfect (or resultative) meaning of the perfect form. In all the four languages the perfect form acquires a further evidential meaning of inference. In example (4) from Lezgi the speaker has seen the remains of a cake on the floor and draws an inference using the perfect form:

- (4) Kic'-i            č̣i    tort            t'u-nwa!  
dog-ERG        our    cake(ABS)    eat- PRF(INF)  
The dog has/must have eaten our cake! (AK<sup>12</sup> 2:23; Greed 2017: 17)

For Even, the expansion of the perfect meaning is attested in the western and central dialects, whereas in the eastern dialect the form conveys mainly only the resultative meaning, with the occasional occurrences of the perfect meaning, whereas the expression of evidential meanings appears to be absent (compare, however, the discussion on Levin 1936 in Section 6.3.). This differing evolution can be explained by contact with the Turkic Sakha (see Malchukov 2000), which has developed an evidential meaning of inference from the perfect (resultative) meaning of the participle form called past resultative I (SG

---

<sup>11</sup> The term “tense-aspect” is used in this study to refer to clear temporal forms, such as past, present, and future, and also to perfect, which is neither a clear aspect, nor a clear tense, expressing “a relation between two time-points, on the one hand, the time of the state resulting from a prior situation, and, on the other, the time of that prior situation” (Comrie 1976: 6, 52–53). The “real” aspects, which are concerned with “the internal temporal constituency of the situation” (Comrie 1976: 5), do not necessarily occur in the same slot with tense: in Even they take the slot preceding the tense; in Tatar and Bashkir, aspectual meanings are expressed with auxiliaries forming a complex verb with the lexical verb in a converb form, while in Lezgi the forms combine tense and aspect.

<sup>12</sup> Initials after examples refer to the language consultants.



1982: 310–311). Western and central Even dialects have copied this meaning to a similar form existing in the language.

For all the four languages the form carrying the perfect meaning originates from a nonfinite verb form. While in Tatar, Bashkir and Even, this form is a past<sup>13</sup> participle (-*GAn* for Tatar and Bashkir, and -*čA* for Even), for Lezgi the perfect form is diachronically a combination of an aorist converb *-na* and a copula *awa* ‘be in’ (Haspelmath 1994: 275–276), resulting in the form *-nwa* (see example (4) above).

In terms of the current use of the form, in Tatar, Bashkir and all three dialects of Even the past participle is used in its default function as an adnominal modifier, including the use as a predicate of a relative clause. It also occurs as a finite verb form conveying the perfect meaning, with further evidential meaning extensions in Tatar and Bashkir, and in western and central Even. The two Turkic languages differ from Even in the finite use of the perfect form, in that in Even the form needs copula support in first- and second-person contexts due to morphosyntactic reasons, whereas in Bashkir and Tatar no copula is used. Lezgi differs from the other three, as it has its own separate form functioning as a past participle which is different from the converb-based perfect.

The perfect form of Tatar, Bashkir and the western and central dialects of Even can form an evidential opposition with another basic past-tense form, this being the common form used to refer to past events. For Tatar, the form is the categorical past tense with *-DY*, for Bashkir it is the definite past with *-DY*, and for Even, mainly the past tense with *-RI*. Examples (5) and (6) show the evidential opposition of the Tatar witnessed (with *-DY*) and the non-witnessed (with *-GAn*) forms.

- (5) Bügen Mäskäüdä jaŋgyr jau-gan.  
today Moscow:LOC rain pour-PRF.3(NWIT)  
Today it rained in Moscow. (Greed 2014 (2016): 81)
- (6) Bügen Mäskäüdä jaŋgyr jau-dy.  
today Moscow:LOC rain pour-PST.3(WIT)  
Today it rained in Moscow. (Greed 2014 (2016): 81)

---

<sup>13</sup> The Even form is called perfect participle, but for comparison purposes I refer to it as past participle.

Example (5) is spoken by someone who did not witness the rain, but heard about it, for example, on the phone, whereas example (6) is said by a person who witnessed the rain. Diagram 1 displays the evidential semantic parameters and the grammatical forms carrying these meanings in Tatar, as they were established in Article 1 (Greed 2014 (2016): 84). As is seen from the Diagram, in Tatar the reported/hearsay marker *di* conveys two different meanings, which are due to differences of register and genre (see further discussion in 6.6.1): reported *di<sub>SP</sub>* refers to the reported meaning in spoken register and hearsay *di<sub>LEG</sub>* to the hearsay function of the narrative genre of folk tales.

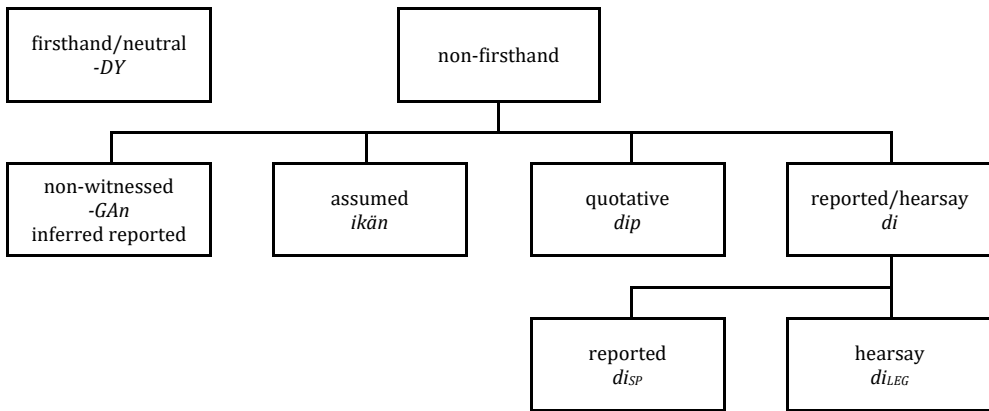


DIAGRAM 1. Tatar evidentiality: Semantic parameters and grammatical forms (from Greed 2014 (2016): 84)

For all the four languages, the form receiving the witnessed interpretation is in general neutral, and this meaning occurs only in contexts that support this reading. The Lezgi aorist with *-na* can also be interpreted as witnessed in certain contexts, but no evidential opposition with the perfect with *-nwa* was attested in this study.

Cross-linguistically one of the main paths of development attested for the perfect (resultative) meaning is the progression to the evidential meaning of inference (see Bybee et al. 1994: 104–105). This meaning may advance further to express other non-firsthand evidential meanings (see Aikhenvald 2004: 116). With regard to the four languages, Tatar and Bashkir (see Greed 2018) have developed a further non-witnessed or reported

meaning from inference, as have the western and central dialects of Even. The Lezgi perfect has remained at the inference stage on the evolutionary path of the perfect (Greed 2017: 19–20), in contrast with a closely related language Aghul, which displays the non-witnessed meaning in narrative contexts (Majsak & Merdanova 2002: 6–7). The eastern dialect of Even differs from the others in that it has overall remained on the resultative stage of development. Diagram 2 is my proposal of the path of development of the perfect form in the main three dialects of Even.

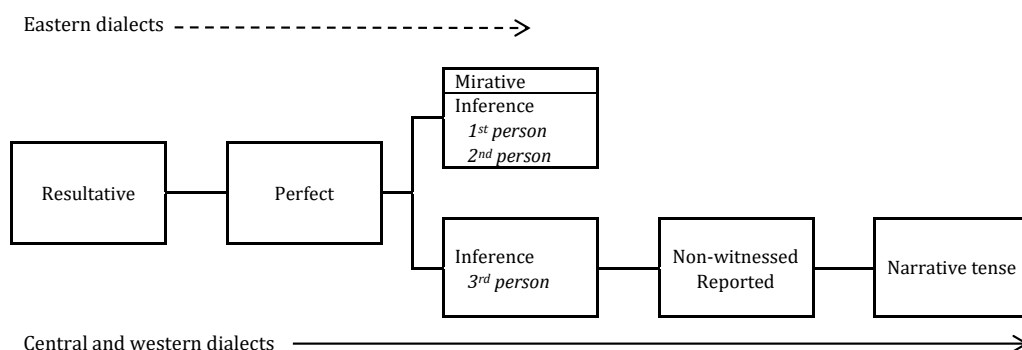


DIAGRAM 2. Path of development of the Even perfect (modified from Greed 2019)

The Even perfect/resultative follows partly the cross-linguistically attested semantic path suggested by Aikhenvald (2004: 116; see Scheme 1 in section 3.1). According to Aikhenvald’s Scheme the meaning of inference can develop from either the resultative or the perfect. For Even, the path appears to begin from resultative and move to perfect, as the eastern dialect conveys almost exclusively only the resultative function (see Malchukov 2000: 455; cf. also Bybee & al. 1994: 105). This meaning evolves further to cover the evidential meaning of inference from visible traces, and on to non-witnessed/reported. In the central and especially western dialects, therefore, the perfect has acquired the inferential meaning, which has evolved to express the meaning of non-witnessed and reported meanings. Unlike in Aikhenvald’s Scheme, where the fourth stage of development is the expression of a more general range of non-firsthand meanings, the

final stage of the Even perfect is an expression of a discourse function: the form is used as a narrative tense and becomes a genre token (see Section 6.6.2).

As we have seen, evidential meanings occur in these four languages only in connection with past tense forms. An exception to this is Kirne, a variant of Lezgi spoken in Turkey. As a result of borrowing from Turkish the perfect form with *-nwa* expresses a non-witnessed, and the aorist *-na* a witnessed meaning (Marianne Beerle-Moor 2016 p. c.). What is of special interest is that in addition to past-tense forms, the evidential function has extended to cover also non-past forms, with the imperfective formed with *-zwa* conveying the non-witnessed, and the habitual *-da* the witnessed meaning – a development thus far rarely attested typologically (Greed 2017: 21–22).

I will conclude the discussion of tense-aspect with a semantic map presenting a cross-linguistic comparison of the languages under investigation. As we have seen, each of the five languages of this thesis utilises the perfect form for expressing evidential meanings. Figure 1 offers a semantic map with cross-linguistic comparison of the perfect gram<sup>14</sup> in these languages.

In Figure 1 the different functions are labelled by the nodes of the network (for example, the node “perfect”). The closeness of functions is indicated by closeness in space and a connecting line (Haspelmath 2003: 216), as is the case, for example, with “resultative” and “perfect” in this map. The criterion for connecting two functions is if in at least one language these meanings are expressed by the same linguistic item (Georgapoulos & Polis 2018: 2).

---

<sup>14</sup> See Anderson’s (1985, 1986) pioneering studies which offer some of the first applications of semantic mapping methodology. The first article discusses and displays cross-linguistic semantic maps of the perfect in a number of languages, and the latter centres around evidentiality, featuring maps of “evidential space” in some ten languages, presenting a general map of “mental space” for evidential meanings (Anderson 1986: 284).

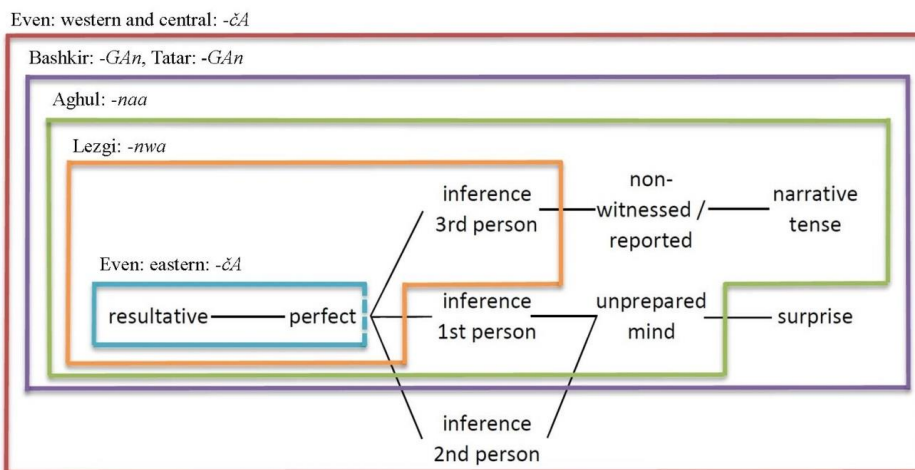


FIGURE 1. Semantic map of the perfect domain; a cross-linguistic comparison between Bashkir, Even eastern, Even western and central dialects, Lezgi, Tatar, and Aghul

In the semantic map, the languages or dialects occupying their own distinctive area in the perfect domain are indicated with different colours. The core area common to all the languages are the resultative and perfect functions. The Even language is divided into two: western and central Even cover the widest area in this semantic space, while eastern Even occupies the smallest area. Contrastingly, the two closely related Turkic languages Bashkir and Tatar take up an identical part of this domain. The other two closely related languages, Aghul and Lezgi, have different distributions of functions, as Aghul has evolved further to express the non-witnessed and narrative tense functions, as well as first-person inference, in addition to the resultative and perfect meanings, and third-person inference common to them both. The distinctive feature of western and central Even is the occurrence of the cross-linguistically rare function of inference in second person. The dotted line of the right-hand edge of the perfect function of eastern Even shows the porous border of this function, as some occurrences of the perfect meaning have been reported (Zippel 2012), and there were cases of inference attested in the 1930s (Levin 1936), as discussed in this thesis.

## 6.2. Default evidentials and their interaction with other categories

In addition to expressing evidential meanings with tense-aspectual forms, usually with verbal affixes, other verb forms are also involved in expressing evidential meanings. In the four languages investigated such markers express information source as their default meaning. An exception is the Tatar particle *ikän* which has a modal/subjective basic meaning (see Section 6.4). Even though evidential markers with a default evidential meaning are semantically and grammatically more independent than markers which have an evidential meaning as an extension of another category, they also interact with other categories and domains, as is shown in this study.

The markers with a default evidential meaning mentioned in this study express the (propositional) evidential meaning of hearsay and the (illocutionary) meaning of quotative (see discussion on illocutionary evidentials in Section 3.5). The Tatar particle *di(lär)* ‘it is said/they say’ and the Lezgi clitic *-lda* ‘it is said’ convey a meaning of hearsay. In Tatar the hearsay marker is restricted mainly to the spoken register, and it has also developed a literary function, being used in a genre close to oral speech, that is, in fairy stories and folk tales (see Section 6.6). A similar clitic to the Lezgi one is also attested in Aghul, the language Lezgi is compared with in Greed (2017). This marker is *-baj*, a grammaticalised form of the speech verb *abas* ‘say, tell’ (Ganenkov et al. 2009:7).

In Diagram 3 I present the semantic parameters relevant in Lezgi and the grammatical markers that code these meanings. In addition to the markers with a default evidential meaning, the diagram also displays the tense-aspectual markers that have developed an evidential meaning.

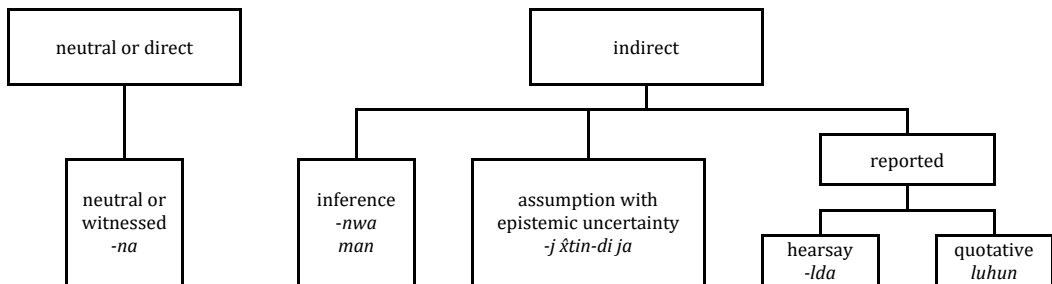


DIAGRAM 3. Semantic parameters of evidentiality and their grammatical markers in Lezgi (from Greed 2017: 27)

In Lezgi the so-called *luhun* strategy (Haspelmath: 1993: 367) occurs in contexts of direct speech when the verb of speech is other than ‘say’. In such a situation a converb form of the speech verb, for example, the aorist converb *lahana* ‘saying’, accompanies the finite verb, functioning as a quotative. The quotative markers of Tatar (*dip*) and Bashkir (*tip*) are fully grammaticalised, whereas in Lezgi a grammaticalisation process has not been completed (Greed 2017: 24). However, in the Quba dialect of Lezgi spoken in Azerbaijan, in oral stories the quotative marker *lahana*, being frequently repeated, is often reduced to *lana* (see Donet 2014: 118). In addition, in some dialects of Even the speech verb displays a quotative-type function. The direct speech is accompanied by the speech verb *göön-* ‘say’ in its converb form, and this is followed by another speech verb in a finite form. In example (7) from the literary eastern dialect of Even the present-tense converb *göniken* is followed by the finite speech verb *kunirin* ‘shouted.3SG’, which carries the tense and person marking of the whole predicate construction.

- (7) – Momi-wu            nʼan    abdu-wu                    hinmač  
 boat-ACC:POSS.1SG    and    belongings-ACC:POSS.1SG    quickly  
  
 emu-li            minŋi-w            d’ermi-nri!  
 bring-IMP.2SG    mine-ACC            steal:NF-2SG  
  
 – gön-iken        kuni-ri-n.  
 say-CVB.PRS.SG    shout-PST-3SG

– Bring quickly my boat and my belongings, you stole what is mine! shouted (the fox). (Robbek & Robbek 2009; cited by Burgherr 2017: 43; eastern: literary dialect; Greed, 2019)

Section 6.3 presents a more in-depth study of the Bashkir quotative on the basis of the study in Greed (2018).

### 6.3. Quotative, subjectivity and the semantic type of verbs

As we saw (Section 3.5) the quotative differs from other basic evidentials in that its scope is over illocutions (Boye 2012: 32), and therefore it is included in the discussion as an illocutionary evidential marker. In the study of the Bashkir quotative the semantic dimension of subjectivity turned out to be particularly useful, together with the semantic type of the verbs occurring with the quotative, as these were shown to be the key to the interpretation of the combination of the quotative marker and the verb. A similar development was also found to take place with the Tatar quotative *dip* (see Greed 2014 (2016): 75–78).

In its default quotative meaning, the Bashkir quotative *tip* marks a speech report (indicated in examples with square brackets) as direct speech, and functions as a syntactic complementiser. In example (8) the speech report is followed by *tip* and the speech verb *öndäšte* ‘addressed’, which carries the tense, person and number marking.

- (8) Ul arala öjđän ber äđäm  
 that interval:LOC house:ABL one human.being  
 syqty. Qarlyqqan tauyş:  
 come.out:DEF.PST become.hoarse:PST.PTCP voice  
 [– Nindäj izge bändälär kilde?]  
 what.kind good human.being:PL come:DEF.PST  
 – tip öndäšte.  
 QUOT address:DEF.PST.3

Meanwhile a man had come out of the house. A hoarse voice addressed (them), “What good people have arrived?” (Musin 1987: 53; Greed 2018: 31)



### 6.3.1. Quotative *tip*, semi-direct speech, subjectivity and multisubjectivity

From the use with direct speech *tip* has extended to code not just spoken utterances, but “internal speech”, that is, the content of thoughts and experiences of the subject of the main clause. This use takes place in the context of semi-direct speech<sup>15</sup> (see Aikhenvald 2008: 383 for the term), a term useful for Bashkir in this study. Semi-direct speech occurs in Bashkir in conjunction with a restricted set of complement-taking, mainly non-speech verbs of cognition and perception. The use of the quotative in this function is common also in Tatar (see Greed 2014 (2016): 76), and occurs in Lezgi in conjunction with verbs of thinking and of emotional expression (Haspelmath 1993: 368).

In example 9 the quotative *tip* functions as a syntactic complementiser linking the speech verb to the non-speech verb *beleü* ‘know’. The speech report complement clause consists of the two words *bäläkäj ašnaqsy* ‘little chef’, and shows the thought or view of “many”.

- (9) 3 jäšlek           Güzäl M.-ny           küptär  
3 year.old       Güzäl M.-ACC many:PL  
[bäläkäj       ašnaqsy]       tip   belä.  
little           chef           QUOT know:PRS

Many know the three-year-old Güzäl M. as the little chef. (Modified from Juldaš 2017; Greed 2018: 37)

The introduction of the semantic dimension of subjectivity (see Section 3.3) helps in the understanding of this function of the quotative. Subjectivity concerns the speaker’s or experiencer’s cognitive processing and expressing of information. As such a “subjective self” is involved in the cognitive processing of knowledge, she can manipulate different linguistic possibilities at her disposal. The marker *tip* brings a subjective element to speech reports conveying thoughts or experiences. Depending on the complement-taking verb, this subjectivity is of varying types and degrees, and the pragmatic context plays a role in the interpretation. The different degrees of subjectivity range from a neutral meaning with a verb of cognition *ujlau* ‘think’ to an expression of a strong subjective involvement by the speaker with the verb *qurqyu* ‘fear’.

<sup>15</sup> In African linguistics a similar phenomenon is also called “combined speech” (see Güldemann 2008: 9).

When combining with the verb of perception *išeteü* ‘hear’, I discovered that *tip* brings a new aspect into the interpretation: in addition to coding what the subject has heard, it signals that some other subject, the original source, was also involved. This is a case of multisubjectivity, a notion I introduced in this thesis (see Section 3.3), which comes close to Evans’ “multiple perspective”: languages can encode more than one perspective simultaneously, be it the viewpoints of conversational participants, or different “reference points in temporal, spatial, social, attentional or epistemic space” (Evans 2005: 93). In a case of multisubjectivity, a marker does not just code what the subject has heard and is reporting, but also signals the involvement of some other subject, the original source. Thus, with the Bashkir *tip*, the quotative gives voice to multiple speakers and merges them, as is the case in example (10).

- (10) [Ul qajt-qan] tip išet-te-m.  
 3SG return-PRF QUOT hear-DEF.PST-1SG  
 I heard that he/she has come back. (GX 12; Greed 2018: 44)

The speech report *Ul qajtqan* ‘he returned’ could be seen both as the actual words that a third party had said about someone having come back, and also as the ensuing thought in the speaker’s mind. While the grammatical form of the main verb *išettem* ‘I heard’ clearly refers to the first-person speaker, the marker *tip* becomes semantically detached from it, no longer conveying only the speaker’s cognition but also referring to some unspecified people and their speech. The viewpoints of the speaker and a third party merge, and the marker *tip* also acquires the evidential meaning of reported. In such a situation, the current speaker is reproducing the original speaker’s words, and at the same time “owning” them, while simultaneously showing with *tip* that the thought expressed was not her own original thought but someone else’s statement.

The combination of *tip* with the verb *küreü* ‘see’ brings in visual ambiguity and epistemic uncertainty, for example in the context of a dream scene. When the speaker uses *tip* and *küreü* together, she indicates that the situation described is not necessarily real, but subjectively experienced, be it dreams (example 11) or other mental images.

- (11) [[jyɣanan jete hyjyr kilep syqty],  
 river:ABL seven cow come:CVB come\_out:DEF.PST  
 tip kürðem.  
 QUOT see:DEF.PST:1SG

I saw as if seven cows were coming out of the river. (GX-a 1:10; Greed 2018: 45)

Of the evidential markers (propositional and illocutionary) discussed in this thesis, while they all fall within Grice's maxims (see Section 2.1), the maxim of quality is the one that is emphasised. The use of the quotative marker *tip* in conjunction with semi-direct speech in Bashkir (and similarly in Tatar, see Greed 2014 (2016): 76–77) is a particularly interesting case in that the marker signals different degrees of subjectivity, levels of involvement by the speaker, and also the involvement of third parties in the case of the verb of perception *išeteü* 'hear', thus yielding a multisubjective meaning. By using the marker *tip*, the speaker qualifies her message appropriately.

### 6.3.2. Quotative *tip* as an adverbial marker

A separate function of the quotative *tip* is its use as a marker of logical relation in the context of a speech report. In this use *tip* functions as an adverbial marker and forms a subordinate clause together with the speech report, conveying the meaning of purpose or intention.<sup>16</sup> Whilst in the complementiser function with complement-taking verbs the quotative *tip* and the verb following it can form a semantically "merged" meaning, in the adverbial marker function *tip* syntactically marks the speech report, forming a subordinate clause together with it. However, *tip* and the verb in the matrix clause are not semantically connected but independent; therefore the marker *tip* can occur with any verb in the main clause, except verbs which require a complement.

In this function of *tip*, as with the complementiser function, a key semantic aspect of *tip* is the expression of subjectivity. This stands in clear contrast to the partially synonymous, but neutral, construction formed with the postposition *ösön*. The use of the

---

<sup>16</sup> This type of quotative function occurs in a number of Turkic languages: from the largest one, Turkish, where the gerund of the speech verb *diye* fulfils this role (Lewis 2000: 176), to some that have just a small number of speakers, for example, the Northern Altaic Chalkan, where the connective *dep* expresses purpose, in the way equivalent to the Russian connective *čtoby* (Marjana Sumačakova, p. c.).

markers *tip* and *ösön* also differs in terms of register, *tip* occurring more widely in spoken discourse. Examples 12 and 13 contrast the marker *tip* with the postposition *ösön*.

- (12) [Irtä tor-a-m] tip,  
 early get\_up-PRS-1SGQUOT  
 ul bögön irtä jatty.  
 3SG today early lie\_down:DEF.PST

He/She went to bed early today (in order) to get up early (literally: ...saying I will get up early). (GM 1:3; Greed 2018: 48)

- (13) Irtä tor-or ösön  
 early get\_up-FUT.PTCP POSTP  
 ul bögön irtä jatty.  
 3SG today early lie\_down:DEF.PST

He/She went to bed early today to get up early. (GM 1:4; Greed 2018: 48)

Personal deixis of *tip* and *ösön* display different characteristics in these contexts. Speech reports by definition present the speaker's uttered words verbatim, or the content of her inner speech, reflecting thus the perspective of the original speaker. As the marker *tip* connects with speech reports, the personal form retains the form of the original speech. With *ösön* a change in personal deixis takes place if the subject of the subordinate clause is in the first or second person: as with indirect speech, the person reference shifts to the third person. In examples (12) and (13) the main difference between the two constructions is that when using *tip* the speaker indicates a decision that has been made, whereas *ösön* is more neutral, showing the purpose for going to bed early. The quotative *tip* signals a subjective intention through an internal thought.

### **6.3.3. The main functions of the quotative *tip* and their distinctive features**

As we have seen, the Bashkir quotative *tip* has two basic syntactic functions: complementiser and marker of adverbial subordination. Both types occur with speech reports, but whilst the complementiser *tip* is an indicator of reported speech, the adverbial marker *tip* assigns the meaning of purpose or intention to the subordinate clause which it follows. The complementiser *tip* can carry additional nuances of meaning.

These depend on the type of verb with which it combines. Table 4 presents the main distinguishing features of the functions of *tip*. A feature, occurring with them all, is the presence of subjectivity.

TABLE 4. The Bashkir quotative *tip* in conjunction with speech reports (from Greed 2018: 51)

|   | <i>tip</i> <sub>Direct Speech</sub>   | <i>tip</i> <sub>Semi-Direct Speech</sub>   | <i>tip</i> <sub>Logical Relation</sub>  |
|---|---|--|---|
| <b>Syntactic function</b>                 | complementiser  | complementiser   | marker of adverbial subordination   |
| <b>Semantic function</b>                  | quotative <sub>1</sub> : signals reported speech with overt reference to source | quotative <sub>2</sub> : signals inner thought/ experience with overt reference to experiencer       | purpose, intention, aim   |
| <b>Form of predicate in speech report</b> | finite  | finite   | finite: indicative, especially in future tense; conditional; imperative; infinitive                                   |
| <b>Verb types in the matrix clause</b>    | complement-taking speech verbs  | complement-taking verbs  | any non-complement-taking verb  |
| <b>Subjectivity</b>                       | perspective of the original speaker   | varying degrees and types  | present   |
| <b>Other evidential meanings</b>          | –   | reported   | –   |
| <b>Multisubjectivity</b>                  | no  | for some verbs   | not relevant  |
| <b>Other characteristics</b>              | can be recast as indirect speech with no change in meaning                      | can be recast as indirect speech with a change in meaning from subjective to non-subjective/ neutral | similar in meaning and usage to the postposition <i>ösön</i> in the context of speech reports, but different register |

#### 6.3.4. Syntactic status of the speech report in semi-direct speech

As we saw, the speech report in the Bashkir construction with semi-direct speech matches the “original” thought or internal speech in form. However, when investigating these constructions, I frequently encountered cases where the constituent which would have been the subject of the “equivalent” direct speech took the form of an accusative-form

object in the context of semi-direct speech, at the same time becoming a syntactic constituent of the matrix clause.

Example 14 (discussed earlier in Section 6.3.1) contains a speech report complement clause which is syntactically linked to the non-speech verb *beleü* 'know' by the complementiser *tip*. In addition to the complement clause *bäläkäj ašnaqsy* 'little chef' being a complement of the matrix clause verb, the verb also has a noun-phrase accusative object *Güzäl M.-ny*.

- (14) 3 jäschlek      Güzäl M.-ny küptär  
 3 year.old      Güzäl M.-ACC many:PL  
 [bäläkäj      ašnaqsy]      tip      belä.  
 little      chef      QUOT      know:PRS

Many know the three-year-old Güzäl M. as the little chef. (Modified from Juldaš 2017; Greed 2018: 37)

If this is compared with the equivalent construction containing direct speech and a speech verb, the “original” comment that the speakers might utter about little Güzäl would be *3 jäschlek Güzäl M. – bäläkäj ašnaqsy* ‘The three-year-old Güzäl M. is a little chef’. The sentence containing this direct speech is:

- (15) Küptär:      “[3 jäschlek      Güzäl M. –  
 many: PL      3 year.old      Güzäl M.(NOM)  
 bäläkäj      ašnaqsy]”, – tip      äjtä.  
 little      chef      QUOT      say:PRS

Many say, “The three-year-old Güzäl M. is a little chef”. (GM-a T:31; Greed 2018: 38)

These two examples raise two questions: How can we account for the difference between the syntactic status of the argument *3 jäschlek Güzäl M.* ‘the three-year-old Güzäl M.’ in the sentence with direct speech and the sentence where the speech report is embedded in the matrix clause? And how do the object and the speech-report clause of example (14) relate to one another syntactically and/or semantically?

The question of the different syntactic statuses of one argument exemplified by (14) and (15) can be explored with the help of Langacker's (1995) discussion on the phenomenon "subject-to-object raising"<sup>17</sup>. The term originates from the argument-raising rules of classical transformational analysis (Langacker 1995: 3). In subject-to-object raising, the subject of the subordinate clause functions simultaneously as the object of the main clause<sup>18</sup> (Langacker 1995: 35).

Interestingly, the characteristic predicates occurring with subject-to-object raising are verbs expressing "some degree of commitment to the reality of a situation", or "verbs of communication serving to inform the addressee about a situation or to effect it by virtue of the speech act named", such as 'like', 'believe'; 'think', 'feel', 'know'; 'say', 'acknowledge' (Langacker 1995: 48). This coincides with the discussion on the types of verbs with which the Bashkir *tip* occurs, as discussed earlier in 6.3.1 (see a fuller discussion of the verbs occurring in such constructions in Greed 2018: 40–48).

In example (15) the direct speech clause is a complement of the speech verb *äjtä* 'says', and *tip* signals a syntactic link between them. In example (14) the argument *3 jäšlek Güzäl M.* '3-year-old Güzäl M.', has been raised to be an explicit direct object within the matrix clause. Thus it can be seen as syntactically fulfilling the role of a direct object within the matrix clause while at the same time being an implicit subject of the subordinate clause.

We note, however, that instead of being in the pragmatically unmarked position where the object follows the subject, the object is fronted to a sentence-initial position. For Bashkir, this indicates that the object constituent is the topic of this utterance. Thus the realisation of the constituent *3 jäšlek Güzäl M.* '3-year-old Güzäl M.' in example (14) as an explicit direct object of the matrix clause can also be explained with information structure: the subject argument of direct speech becomes the object of the matrix verb in an equivalent utterance with semi-direct speech and falls outside the speech report, if this argument is an established topic.

---

<sup>17</sup> I am grateful to Kasper Boye for drawing my attention to this fact.

<sup>18</sup> In Langacker's example *I expect Don to leave* subject-to-object raising is explained in cognitive linguistic terms in the following way: "*Don* functions simultaneously as the landmark of *expect* and the trajectory of *leave*" (Langacker 1995: 34–35).

As for the syntactic relation between the object and the speech-report clause, I found two possible explanations. Firstly, following Schultze-Berndt and Himmelmann (2004) the Bashkir speech report predicate could be viewed as a “depictive secondary predicate”. In their discussion on the semantics and syntax of constructions containing a second predicative element in addition to the main predicate, Schultze-Berndt and Himmelmann (2004: 65) make a distinction between predicative complements, such as subject and object complements, and depictive secondary predicates. The main difference between them is that while predicative complements are obligatory as part of the argument frame of the predicate, depictive secondary predicates are optional. The Bashkir speech report predicate displays most of the seven criteria listed by Schultze-Berndt and Himmelmann (2004: 77–78) which a predicate construction with a depictive secondary predicate should fulfil. The key criteria are the following: the speech report predicate has a predicative relation to a participant of the main predicate, its controller, and is therefore obligatorily controlled; and the controller is expressed separately as an argument of the depictive predicate.

Secondly, the Bashkir secondary predicate could be interpreted as an object complement, and thus an obligatory part of the argument frame of the main predicate. There is no syntactic marking to show such a dependency; semantically, however, the speech report is associated with both the object (its controller) and the verb of the matrix clause. The speech report cannot be removed, unless the quotative *tip* is also taken out.

Even though the syntactic structure of this construction does not fully coincide with either of the two concepts discussed above, semantically the secondary predicate has depictive characteristics, as it expresses “a physical state or condition, or a role, function or life stage” (Schultze-Berndt & Himmelmann 2004: 65).

#### **6.4. Evidentiality and epistemic modality**

In the four languages studied the verbal expression of evidentiality with past-tense verb forms is in general distinct from epistemic evaluations by the speaker with regard to the certainty or reliability of the information conveyed (see Section 3.3 for discussion). In Tatar the perfect form would usually need to have contextual support for the epistemic



interpretation of uncertainty or doubt, for example, the inclusion of the reported particle *imeš* ‘it is said’ (see Greed 2014 (2016): 74, 86).

With Lezgi, based on the results of the study, I suggest that even though the aorist is often interpreted as conveying reliable information, the perfect form does not convey the opposite interpretation of doubt or uncertainty with regard to the reliability of the utterance. This coincides with Forker’s (2018b: 493) statement about evidentials not combining with (epistemic) modality in Nakh-Daghestanian languages. Outside of expression of evidential meanings with past-tense forms, Lezgi has, however, a verbal construction, which combines the evidential meaning of assumption with epistemic uncertainty, as shown in example (16). The construction in question consists of a nominalised participle combined with the equative particle *xtin*, and a copula.

- (16) Ali    ata-nwa-j    xtin-di    ja.  
Ali    come-PRF-PTCP like-NMLZ    COP  
It looks like Ali has come. (Literally: ‘Like Ali having come is.’) (AK 2:4; Greed 2017: 22)

The Tatar grammaticalised particle *ikän*, originating from a past participle of the verb ‘be’, was described in Article 1 (Greed 2014 (2016): 80–81) as being connected with modality, evidentiality and mirativity. While *ikän* is multifunctional, the meaning present in all its uses is the speaker’s viewpoint and epistemic stance. In the article I defined this meaning component as “speaker-oriented subjectivity”, and as I saw subjectivity as a subcategory of modality, the basic meaning was seen to be modal. In Article 3 (Greed 2018) I discuss the Bashkir quotative, where the notion of subjectivity also plays a major part. There I regard subjectivity as a semantic dimension separate from modality (see Section 6.3 for discussion). Further investigation is needed to determine the relationship between modality and subjectivity, but in the following description of the particle *ikän* I keep to the terminology of Article 1.

The Tatar particle *ikän* can combine with different tense-aspect forms, expressing the non-firsthand meaning of assumption. When accompanying the perfect form, *ikän* strengthens the non-firsthand meaning of the verb, but when it combines with present and future tenses where the verb by itself does not carry a non-firsthand meaning, the

particle conveys the meaning of assumption. In example (17) *ikän* accompanies a verb with the categorical future tense suffix *-AčAq* which signals a certain event in the future.

- (17) Alsu jaz-ačaq ikän  
 Alsu write-CAT.FUT.3 NFIRSTH/ASSUM  
 Alsu will certainly write, it appears. (Zakiev 1992: 191; Greed 2014 (2016): 79)

The particle shows that the speaker draws a conclusion, making an assumption on the basis of what she has heard from some other source.

The particle can also receive an inferential reading and be accompanied by a mirative meaning, as in example (18), where the speaker, when entering, guesses from the voices she hears that they have guests. Here, the particle *ikän* contains both the mirative meanings of unexpectedness and surprise.

- (18) Èj, bezgä qunaqlar kil-gän ikän!  
 oh 1PL:DAT guest:3PL come-PRF(INF) ASSUM(MIR)  
 Oh, guests have come to us! (Greed 2014 (2016): 80)

Mirative markers can also be manipulated, for example, in jokes (Aikhenvald 2014). In Tatar, *ikän* can be employed for politeness, to soften a statement of bare facts. Example (19) shows yet another use of the particle. The example is the headline of a news item.

- (19) Universiadada tatar tele kiräk ikän...  
 university\_games:LOC Tatar language:POSS.3 is\_needed ASSUM(MIR)  
 The Tatar language is apparently needed in the University Games... (Zakirova 2012; Greed 2014 (2016): 81)

In the text itself the writer queries whether the information stated in the heading will be realised, and uses *ikän* to convey irony.

Diagram 4, which shows the different meanings and meaning extensions of *ikän*, concludes the discussion of the Tatar particle. The basic meaning of the particle *ikän* is subjective, conveying that information is received “by a conscious subject” (see Johanson 2003: 274). This meaning can receive various mirative extensions. The interpretation of

assumption is also possible in declarative contexts, and in interrogative speech situations, *ikän* can acquire a speculative nuance.

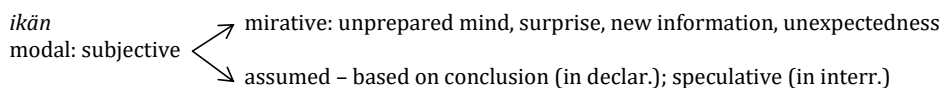


DIAGRAM 4. Modal *ikän* in declarative and interrogative clauses with extensions of mirativity and assumption/conclusion (from Greed 2014 (2016): 81)

In the languages of this study evidentiality is the primary category when compared with epistemic modality: markers coding evidentiality may receive an epistemic extension, but the opposite does not usually take place. An exception is the Tatar particle *ikän* with a modal/subjective meaning, for which the evidential assumption can become an extension in certain contexts.

### 6.5. Evidentiality, mirativity, and person

In this study, mirative meanings were encountered in all the four languages investigated. Mirative meanings occur both with tense-aspect verb forms, and with particles (see Greed 2017: 25 for the Lezgi inferential particle *man*). With verb forms these meanings occur almost exclusively as extensions of the evidential meaning of inference, which has developed from a tense-aspect meaning. An exception is a construction in Tatar which has the form of third-person interrogative imperative, as shown in example (20). It can receive a mirative interpretation, but there is no evidential meaning involved.

(20) Uzgan            žyelyšqa        qatnašmagan            ide,  
       pass:PRF.PTCP meeting:DAT participate:NEG:PRF.PTCP be:DEF.PST  
 bügenge        žyelyšqa        berenče        bulyp  
       today:ADJZ meeting:DAT first be(come):CVB  
 kil-mä-sen-me!  
       come-NEG-IMP.3-Q

Even though not participating in the earlier meeting, did he not then come first to today's meeting! (TG 1993: 143; Greed 2014 (2016): 80, footnote 16)

As for particles, a particle can express a mirative meaning as its default meaning, or this meaning can be an extension of another meaning. Lezgi has a number of particles used for mirative and related meanings. The particle *ha* has a default mirative meaning. It occurs in sentence-final position and has scope over the whole utterance. Example (21) is from a story where little Zemfira has received a bagful of live fish from her father. She shows them to her mother who is ready to prepare them for a meal. Zemfira expresses her alarm:

- (21) Čan\_alaj-bur ja ha!  
 living-PL COP PTCL(MIR)  
 They are alive! (MD1 5:1; Greed 2017: 25)

In Bashkir (Greed 2018: 29, footnote 7) and Tatar (see Section 6.4), mirative meanings also occur in conjunction with the modal (subjective) particle *ikän*.

When the default meaning of a marker is mirative, the speaker chooses it to increase the appropriate quality of her message. However, when the interpretation of a marker is dependent on other linguistic factors, for example, the grammatical person, other aspects in addition to the speaker's choice play a major part in the appropriate interpretation of the message. This interaction is clearly visible with the meaning of inference in first-person contexts. First-person effect (see Section 3.2) occurs in Tatar, Bashkir and Even, and also in Aghul (Majsak & Merdanova 2002: 5), but it was not attested in Lezgi. In the following, I will first discuss first-person effect in Even (Section 6.5.1), and then proceed to discuss another related phenomenon, "second-person effect", encountered in Even (Section 6.5.2).

### **6.5.1. First-person effect and mirativity in Even**

In the main three Even dialects, the meaning of inference occurs in the western and central dialects, but not in the eastern dialect. Malchukov (2000: 443) suggests that in western and central dialects the inferential meaning may be the main meaning of the perfect form since there are limitations to its use which are to do with the first person. The meaning Malchukov assigns for the inference in first-person contexts is lack of

control by the agent. Example (22) from the western dialect of Sebyan-Kyul has first-person effect.

- (22) Tarıt bi: hukle-hen-če bi-he-m uručun  
 then 1SG sleep-MOM-PRF.PTCP be-NF-1SG PTCL(INF)  
 dalše e-he-m öjdö:-r.  
 further.R NEG-NF-1SG remember.S-NEG.CVB

Then I must have fallen asleep, I don't remember anything more.  
 (IVK\_memories\_023; western: Sebjan-Küöl; Greed, 2019)

The verb form *huklehenče bisem*, consisting of a perfect participle and the non-future tense in first person singular, receives both an evidential interpretation of inference and a mirative interpretation, which in this case conveys lack of consciousness and is further strengthened by the inferential particle *uručun*. While it is possible to also express the default perfect meaning in first-person contexts, such an interpretation requires a “strong context” (Malchukov 2000: 444), for example, the addition of an adverb *ukal* ‘already’. Thus, in example (23) with a first person plural form, the default interpretation combines inference and mirativity, showing the lack of control of the speakers, and the unexpected nature of the event.

- (23) Egd'en möden evgid-du-n nulge-sen-če-l bi-se-p.  
 big flood this.side-DAT-3SG nomadize-MOM-PRF-PL be-NF-1PL  
 (It turned out that) we have nomadized just before the big flood. (Malchukov  
 2000: 444)

### 6.5.2. Second-person effect and mirativity in Even

Over 80 years ago Levin (1936) published a small Even-Russian dictionary, which utilised material from the eastern literary Ola dialect. The dictionary is accompanied by a brief grammatical essay on the language. Levin (1936: 205) describes the Even participle as being able to occur in any grammatical function in a sentence, including the (finite) predicate function. What is remarkable is that the three examples he presents contain a finite form with the perfect participle *-čA* and clearly display the meaning of inference, which use has been shown to be absent from the current eastern dialects. These examples

indicate that at least in the 1930s the functions of the finite perfect form were broader in eastern dialects than just the resultative use. In addition, as the Ola dialect is spoken in the Magadan area, it is outside the direct influence of the Sakha language, the language which has been suggested as the main cause for the appearance of the inferential meaning in the western and central dialects. Levin's (1936) examples are in the third (24), first (25), and second (26) persons.

- (24) Kike nukečee-m ma-ča.  
 Kike moose-ACC kill-PRF.PTCP  
 Kike has killed a moose. (Levin 1936: 206; eastern Ola dialect; Greed, 2019)

In example (24) the finite verb *mača* 'has killed' expresses the perfect proper meaning. With first and second persons, a copula *bi-* 'be' is added for syntactic reasons to mark tense, person and number. In examples (25) with first person plural and (26) with second person singular, the perfect form receives the interpretation of inference, with an additional meaning component of mirativity.

- (25) Mut nakat-u ma-ča-l bi-se-p.  
 we bear-ACC kill-PRF.PTCP-PLbe-NF-1PL  
 (It turns out that) we have killed a bear. (Levin 1936: 206; eastern Ola dialect; Greed, 2019)
- (26) Hi huliča-m ma-ča bi-se-nri.  
 you fox-ACC kill-PRF.PTCP be-NF-2SG  
 (It turns out that) you have killed a fox. (Levin 1936: 206; eastern Ola dialect; Greed, 2019)

Levin (1936: 206) comments, "in its predicative role in 1<sup>st</sup> and 2<sup>nd</sup> persons singular and plural [the perfect participle] ... has the meaning 'it turns out to be'<sup>19</sup>". Thus example (25) has the first-person effect with a mirative meaning, as the speakers came to recognise the outcome of their action only after it had happened, through inference from visible results. Example (26) with the second-person form *mača bisenri* 'you(SG) have apparently killed' presents a very interesting case, as the mirative interpretation occurs also in the context

---

<sup>19</sup> In Russian: "оказывается" (*okazyvaetsja*).

of a second-person reference. This “second-person effect” differs from first-person effect in that it is not the second-person agent of the utterance who does the inferring, but the speaker. The speaker remains the “evidential origo” (see Sun 2018: 62) of the information conveyed: the form signals the speaker’s experience and viewpoint.

When Levin’s examples were presented to the language consultant from Sebyan-Kyul speaking a western dialect, her views about the meanings and differences between examples (24), (25) and (26) coincided with Levin’s explanation. The mirative meaning is one of lack of awareness, unexpectedness, and surprise. In example (24) the event is factually stated, without any epistemic evaluation; in addition, the speaker was not an eyewitness of the event. Thus the perfect form displays both an inferential and non-witnessed meaning. An equivalent utterance to example (24), but with a witnessed meaning would be the following:

- (27) Kike nukečee-m ma-ri-n.  
 Kike moose-ACC kill-PST-3SG  
 Kike killed a moose. (IZ: 1; western, Lamunkhin: Sebyan-Kyul; Greed, 2019)

Mirative meanings are less frequently attested in second-person than in first-person contexts. However, a second-person effect appears to be possible in the Finno-Ugric language Komi. When discussing the mirative meaning in an indirect (non-firsthand) evidential context, Leinonen and Vilkuna (2000: 501) illustrate this with example (28), where in the second sentence there is a second-person subject and the verb is in the second past tense, the “unwitnessed past” (Leinonen & Vilkuna 2000: 495), which can also express the meanings of resultative and perfect.

- (28) Mis’a, gačkö, te munin n’in.  
 I\_think perhaps 2SG go:PST1:2SG  
 A tani na völömyd.  
 but here still be:PST2:2SG  
 I thought you had already gone, but here you still are. (Leinonen & Vilkuna 2000: 501; Greed, 2019)

According to Leinonen and Vilkkuna (2000: 501), in this case indirectivity extends to express “sudden revelations or reinterpretations of one’s own experience or action”. The second-person verb form coding the speaker’s inference and the accompanying surprise functions here in a manner similar to the second-person perfect form of the Even example (26).

When discussing the different functions the Even perfect has in the different dialects, Malchukov (2000: 443–444) appears to suggest that in the western and central dialects, in first-person contexts the default meaning of the perfect form is that of inference accompanied by a mirative meaning. The basic perfect interpretation, without a mirative nuance, is also possible, but this reading needs contextual support.

As far as third-person contexts are concerned, in evidential contexts the perfect form does not inherently contain a mirative meaning. However, as we have seen, the mirative meaning is part of the default reading of the perfect in second-person inferential contexts in western and central Even. Therefore, I propose that the meaning of evidential inference is the default, unmarked meaning of not just the Even perfect in first-person contexts, but also in second-person contexts.

## **6.6. Evidentiality and discourse**

As we have seen, a key aspect of the meaning of the perfect is the relevance of a past event at the moment of speech. Lindstedt (2000: 366) mentions another pertinent element which relates to the use of the perfect in discourse: its non-narrativity – it stands alone, detached from other past facts. Thus the perfect form is frequently utilised to convey background information in discourse, as is the case, for example, in many Turkic languages, including Tatar and Bashkir. However, in a number of languages the perfect form is also used in discourse to advance the plot, that is, to express foreground events that move the story line forward. In such contexts the perfect form has lost its basic perfect function (see Lindstedt 2000: 365), being replaced by a narrative function. Such a narrative use may develop into a marker signalling a particle genre, a genre token.

The development of the perfect meaning into an evidential meaning of inference is widely attested. In such a case the meaning of current relevance of the perfect evolves into inference based on traces of a past event. Such a development takes place, for



example, in Turkish (see Aksu-Koç & Slobin 1986: 164–165) and in Nakh-Daghestanian languages (see Forker 2018b: 493–498), including Aghul and Lezgi. In both languages the perfect has developed an additional meaning of inference. These languages are closely related, but only the Aghul perfect is documented to have evolved further to cover a narrative function. In Lezgi the perfect has not (yet) extended from the meaning of inference (Greed 2017: 19–21), whereas in Aghul inference expressed with the perfect/resultative form has developed into a non-witnessed meaning. Concurrently with this development, the other Aghul past-tense form, the aorist<sup>20</sup>, has acquired the meaning of witnessed with the result that in narrative discourse the two forms are in evidential opposition with each other (see Majsak & Merdanova 2002). Thus the form that was non-narrative, expressing the perfect (backgrounding) meaning, has come to be used as a plot-advancing narrative form.

A similar evolution has been documented in some Romance languages, where the perfect has initially become a narrative tense conveying the perfective past meaning, and has consequently formed an opposition with the imperfective past (see Lindstedt 2000: 374 and Squartini & Bertinetto 2000). In the following sections I will summarise how the function of a narrative tense is manifest in Tatar (Section 6.6.1) and in Even (Section 6.6.2).

### **6.6.1. Discourse genres in Tatar**

In addition to everyday communication, Tatar evidentials play an important part in literary genres: the verb suffixes *-GAn* and *-DY* and the particle *di* function as markers, tokens of genre. Table 5 (Greed 2014 (2016): 83) gives an overview of Tatar evidentials as tokens of genre, with *di*<sub>SP</sub> reported referring to the use of the particle *di* in spoken register, and *di*<sub>LEG</sub> hearsay describing *di* in the folk tale/legend function. Visible in the table is also another tense form, the present, formed with *-A* (not discussed in the current study), which can be used as a similar genre token to the *-GAn* form.

---

<sup>20</sup> This form is called ‘perfect’ in Majsak and Merdanova (2002). The term ‘aorist’ used in this thesis is taken from a more recent article by Majsak (2012: 228).

TABLE 5. Evidentials as tokens of genre (from Greed 2014 (2016): 83)

| Genre                         | Firsthand<br><i>-DY</i> | Non-firsthand<br><i>-GAn (-A)</i> | Reported<br><i>di<sub>SP</sub></i> | Hearsay<br><i>di<sub>LEG</sub></i> |
|-------------------------------|-------------------------|-----------------------------------|------------------------------------|------------------------------------|
| novel, “realistic”<br>fiction | yes                     |                                   |                                    |                                    |
| history writing               |                         | yes                               |                                    |                                    |
| folk tales                    |                         | yes                               |                                    | yes                                |
| spoken language               |                         |                                   | yes                                |                                    |

The Tatar resultative/perfect past-tense form *-GAn* follows the cross-linguistically established path of development of evidential meanings, with the evidential meanings of inference and reported having developed from the resultative meaning. I suggest that the quotative particle *dip*, originating from the converb form of the speech verb *dijü* ‘say’, has through grammaticalisation first become a marker of direct speech, and then developed to also express indirect speech (called semi-direct speech in Greed 2018). It has extended further to marking clausal connections, evident in its complementiser function. All of these different stages and functions are synchronically observable in the language.

### 6.6.2. Narrative function of the Even perfect

As in Tatar, the Even perfect form with *-čA* has acquired a function signalling a narrative genre in western and central dialects of Even (see Malchukov 2000: 444). In both languages the perfect originates from a participle, a non-finite form whose prototypical function is adnominal modification. However, when used in the perfect, evidential, and (consequently) in the narrative function, the originally non-finite participial predicate is behaving like a finite predicate. How has this evolution taken place? Malchukov (2013: 177–178) calls such a development “verbalization”: a non-finite predicate is reanalysed as a verbal finite predicate. In some languages, for example, in the East-Tungusic Nanai, the originally non-finite forms have replaced already existing finite forms and pushed them to a different function (Malchukov 2013: 188). In Even and Tatar, however, such a

development has not taken place, and both the non-finite and finite uses of the participle occur side by side.

In its perfect function, the Even perfect with *-čA* expresses a non-narrative meaning (see Section 3.4). As it evolves through the inferential and non-witnessed evidential meanings to convey the meaning of a narrative genre, at the same time it loses its perfect “nature”, and becomes a plot-advancing narrative form.

Forms with evidential meanings are widely used as markers of genre (see Aikhenvald 2014: 35). Cross-linguistically, the reported evidential is commonly attested in ancestral stories and legends. In a number of languages where the perfect form has evolved to express evidential inference, this meaning has then further progressed to express the evidential reported meaning. This has also taken place in western and central Even with the *-čA* form. In addition to the meaning of inference, this form can also receive the interpretation of non-witnessed and reported, whereas in the eastern dialects the main function of the form is the non-finite attributive function, with the resultative meaning attested infrequently, and the perfect meaning even more rarely (see Zippel 2012). It seems, therefore, a natural development that the narrative function of the perfect form has developed only in the western and central dialects, while being absent in eastern Even.

In the western and central dialects the narrative tense *-čA* occurs in a variety of story types ranging from fairy and folk tales containing supernatural elements (see examples 29a–e below) to (oral) traditional stories passed on from generation to generation. All these narrative discourse types refer to events detached from the moment of speech. In addition, the form *-čA* is used for recounting events and experiences which the narrator had not witnessed. In such contexts, the stance of the story-teller is neutral, “maximally objective” (Zaxarova). The form does not carry additional epistemic modal meanings to do with degree of certainty of the information conveyed, or evaluations about its truth value.

If *-čA* functions as a genre token, marking a piece of discourse as a story, and thus being the main plot-advancing form, other forms, for example, the past with *-Ri* or the non-future with *-RA* are used for other discourse functions: to slow the action down, or to indicate the peak of the story. Examples (29a–e) are from the fairy tale Five Sisters in the

central Momsk dialect, showing the interplay of the different verb forms at one of the peaks of the narrative. The story is about five sisters, the eldest of whom became a monstrous cannibal.

(29a) Arisag            atikan            d'u:la            alačid-ča.  
 monster            old.woman      house:LOC      wait:PROG-PRF.PTCP(SG)  
 The old monster woman waited in the house.

(29b) D'u:duk            ku:ninni            i:-če:  
 house:ABL          shout:POSS.3SG sound-PRF.PTCP(SG)  
 From the house a shout was heard:

(29c) – Ile    jak    gorron imsejur            gele-či-s?  
 where what long fat:POSS2:POSS.PL      search-PST-2SG  
 – Where are you looking for that fat for so long?

(29d) Asi                    ne:kuduk            baroggot-ta-n:  
 woman                  warehouse:ABL      answer-NF-3SG  
 The woman answered from the top of the warehouse:

(29e) – Tek-e            alatli,            geleted-de-m=e!  
 now                  wait:IMP:2SG      search:PROG-NF-1SG=EMPH  
 – Just wait a bit, I am looking for it! (Bokova 2002; cited by Burgherr 2017; central: Momsk; Greed, 2019)

The whole story is narrated with the plot-advancing form with -čA, as is seen in (29a) and (29b). After the direct speech in (29c), the narration in (29d) switches into the non-future tense with -RA. This form highlights a key moment: the monster woman is about to catch the youngest sister and her children and eat them. After (29e) the regular narrative resumes, and verbs with -čA are used until the next peak.

A genre token of the type -čA in the western and eastern Even stories is a defining feature for the discourse. The other forms used in the story have their role in the context of this chosen framework, and may display other functions, not typical for them outside this particular discourse context, as we have seen the marker -čA to do. In its function as a narrative tense marker, -čA is purely a genre marker without any epistemic modal overtones or evaluations to do with the reliability of information. In addition to this function, the form -čA can occur also in its non-narrative basic meaning of perfect (resultative) within a narrative.

### 6.7. Summary of interaction between evidentiality and related categories and domains in Bashkir, Even, Lezgi and Tatar

We have seen that in the context of the four languages investigated, evidentiality interacts with a number of different semantic categories. Figure 2 displays in semantic space the semantic categories and domains playing their part in the expression of evidential and related meanings in Bashkir, Even, Lezgi and Tatar, based on the findings of this study. The meaning labels have been arranged in a manner reflecting the discoveries of the four languages of this study, without reference to wider cross-linguistic studies. The aim of this special semantic map is to portray the complex interrelations of the different categories and domains, and to fit the main functions discussed in this study into a coherent and interconnected whole in semantic space.

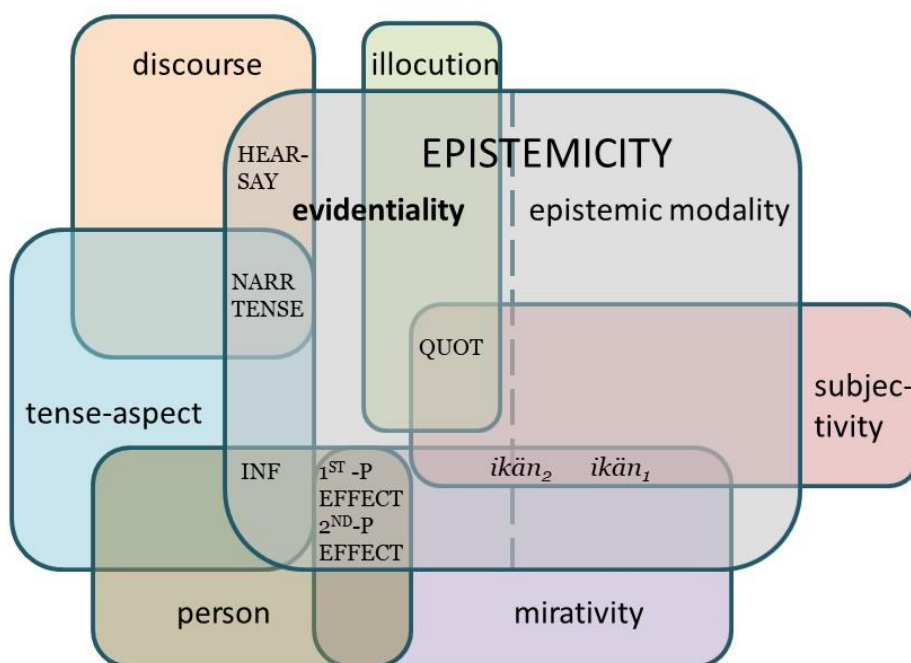


FIGURE 2. Interface between evidentiality and related categories and domains in Bashkir, Even, Lezgi and Tatar

Evidentiality and epistemic modality are closely related categories and together they are subsumed under the supercategory of epistemicity. In the four languages of this study, the two categories are in general distinct, with evidentiality expressing source of information and epistemic modality conveying the way the speaker relates to information. However, they also have some overlap.

One of the key ways of expressing evidential meanings in the four languages is as extensions of the tense-aspectual meanings. This development is particularly visible with the evidential meaning of inference. As shown in Figure 2 inference (“INF”) appears at the intersection of tense-aspect and evidentiality, and this development was attested in the western and central dialects of Even, in Lezgi, Tatar and Bashkir. A further development of inference takes place in first-person contexts (“1<sup>st</sup>-P EFFECT”) where inference interacts with mirativity. Such a case occurs with western and central Even, Tatar and Bashkir (see Greed 2018: 28). In Even mirative meanings interact with tense and inference also in second-person contexts (“2<sup>nd</sup>-P EFFECT”). The meaning of inference can develop to express the non-witnessed or reported meaning, and this meaning may cross over to the domain of discourse, with the earlier perfect meaning having evolved through the evidential path into a narrative tense (“NARR TENSE”), thus circling back to overlap with tense-aspect. This development was attested in western and central Even and Tatar.

The quotative (“QUOT”), as described in Bashkir and Tatar, is the meaning where illocution, evidentiality and subjectivity intersect. The Lezgi quotative also displays behaviour similar to Bashkir and Tatar in the area of complementation, and may therefore occupy a similar position in semantic space. Even though the labels displayed in Figure 2 are all semantic, I chose to include in it also one language-specific grammatical marker, the Tatar modal/subjective particle *ikän*. This marker codes the most intricate meaning of this study, and I find it helpful to display visually the location of the meaning complex represented by this particle. The marker occupies two positions in the space: one where epistemic modal, subjective and mirative meanings overlap (“*ikän<sub>1</sub>*”), and another where epistemic modal and subjective meanings intersect with the evidential meaning of assumption (“*ikän<sub>2</sub>*”).

## 7. Conclusion

This thesis has investigated the expression of evidentiality and related categories in four indigenous non-Slavonic languages spoken in the Russian Federation, the languages in question being the Turkic Tatar and Bashkir, the Nakh-Daghestanian Lezgi, and the North-Tungusic Even. The focus of the study has been on non-lexical expression of evidential meanings, centring on verbal expression and particles, as these are the main ways in which the four languages convey evidential and related meanings.

While the four peer-reviewed articles of this thesis each focused on one individual language, in this introductory article the goal has been, firstly, to investigate the results of each study and to compare the languages with one another, and, secondly, to discuss the topics found relevant for the four languages in a cross-linguistic context. For this purpose, the interaction of evidentiality with categories of tense-aspect, epistemic modality, mirativity, as well as the domains of discourse and illocution were considered.

The fact that the four languages are in many ways typologically similar has found reflection in the way they express the explored meanings. All four languages use an illocutionary quotative to mark direct speech, and the inferential evidential was also discovered to be a unifying factor for the four languages. However, each language also displays its own idiosyncrasies.

For Even the key for its expression of evidential meanings with verb forms is dialect-based semantics, as the eastern dialect is located at a different position on the path of development of the resultative/perfect to a narrative function than the western and central dialects. On the basis of comparing the results of the four articles and the wider cross-linguistic and theoretical discussion, this thesis offers proposals on the path of development of the resultative/perfect of Even, and brings together the semantic parameters of evidentiality and their grammatical markers of Lezgi in graphic form.

The functions of evidential markers have been shown in a number of languages to cross over to the domain of discourse, as they develop into markers signalling a particular genre. This was found to be the case with Even and Tatar. In both languages the resultative meaning has developed through evidential meanings of inference and non-witnessed to a narrative function, where the form is used as the main plot-advancing form

in genres, narrating events distant from the moment of speech. The meaning of the marker has thus experienced two different evolutions: in the domain of discourse from a non-narrative, backgrounding function to plot-advancing narrative use; and a circular path from its erstwhile tense-aspect meaning through evidential and discourse functions back to expression of tense-aspect. Information structure also plays a role in the interaction of different categories, as it was found to be the explanation for the syntactic change which takes place with the subject argument in the switch from direct speech to semi-direct speech in Bashkir. Such findings pertaining to discourse and information structure are further indications of the complexity of which one needs to be aware when investigating phenomena like evidentiality.

Two key discoveries are worthy of special mention. Firstly, through the exploration of the quotative *tip* of Bashkir and the modal/subjective marker *ikän* of Tatar, subjectivity was found to be another important semantic dimension interacting with these categories. In Bashkir I discovered that the subjective expression of *tip* expanded to convey multisubjectivity in conjunction with some verbs of perception and cognition, as, in addition to the current speaker, the marker signals the involvement of the original speaker(s) by merging the different speakers into a polyphonic expression. I would envisage that the semantic dimension of (multi)subjectivity will find its application in many languages in the future.

Another discovery, one which is commended for further investigation cross-linguistically, is the second-person effect found to occur in Even in inferential contexts. Across languages inference in first-person contexts has been attested to bring about a mirative interpretation, but in Even the same takes place also in the context of the second person. The evidential origo remains the same in both cases: it is the speaker whose viewpoint is conveyed by the expression.

While the four articles had some overlap in the viewpoint from which the studies were conducted, each one of them concentrated on a particular aspect relevant to each individual language. Therefore discoveries were not always compatible. For example, it was possible to portray the features of Bashkir only very briefly, as the study concentrated on the quotative. Therefore, a number of the features described in the other three languages, especially in the related language Tatar, await more detailed study.



A key outcome of the thesis is the visual display of the semantic categories involved in the expression of evidential and related meanings in the four languages in the form of Figure 2 sketched in section 6.7 of this introductory article. In this illustration, the interactions of the different categories and domains are displayed, and, at the same time, some key meaning labels are assigned a place appropriate for them on the basis of this study. Some of these labels have cross-linguistic application, while others are language-specific. My hope is that the proposed sketch is but a beginning of further cross-linguistic investigation into the complex interaction of semantic categories involved in human communication.

## **Abbreviations**

ABL = ablative  
ABS = absolutive  
ACC = accusative  
ADD = additive  
ADJ = adjectivizer  
ADV = adverbial  
ALI = alienable possession  
ASS = assumption  
CAT.FUT = categorical past  
CVB = converb  
COP = copula  
DAT = dative  
DEF.PST = definite past  
DIST = distal  
ELAT = elative  
EMPH = emphasis  
EP = epenthesis  
ERG = ergative  
EVID = evidential  
EXCL = exclusive (1PL)  
FOC = focus  
FUT = future  
HAB = habitual  
HEARSAY = hearsay evidential  
IMP = imperative

INCL = inclusive (1PL)  
INDEF = indefinite  
INF = inference  
INS = instrumental  
MIR = mirative  
MOM = momentative  
LOC = locative  
NEG = negation  
NF = non-future  
NFIRSTH = non-firsthand  
NOM = nominative  
NMLZ = nominalizer  
NWIT = non-witnessed  
PL = plural  
POSS = possessive  
POSTP = postposition  
PRF = perfect  
PROG = progressive  
PROP = proprietive ('provided with')  
PROX = proximal  
PRS = present  
PST = past  
PST1 = past 1 of Komi  
PST2 = past 2 of Komi  
PTCL = particle  
PTCP = participle  
QUOT = quotative  
R = Russian copy  
RECP = reciprocal  
REFL = reflexive (possessive)  
REP.HAB = repeated habitual  
RESTR = restrictive clitic  
RES = resultative  
S = Sakha copy  
SG = singular  
WIT = witnessed

## References

- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 2008. Semi-direct speech: Manambu and beyond. *Language Sciences* 30, 383–422.
- Aikhenvald, Alexandra Y. 2014 (2016). The grammar of knowledge: a cross-linguistic view of evidentials and their expression of information source. In: Aikhenvald & Dixon (eds), 1–51.
- Aikhenvald, Alexandra Y. 2018. Evidentiality: The framework. In: Aikhenvald (ed.) 2018, 1–43.
- Aikhenvald, Alexandra Y. (ed.) 2018. *The Oxford Handbook of Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y. 2018. Evidentiality: the framework. In: Aikhenvald (ed) 2018, 1–43.
- Aikhenvald, Alexandra Y. & R. M. W. Dixon (eds) 2014 (2016). *The Grammar of Knowledge*. Oxford: Oxford University Press.
- Aksu-Koç Ayhan A. & Dan I. Slobin 1986. A Psychological Account of the Development and Use of Evidentials in Turkish. In: Chafe & Nichols (eds), 159–222.
- Anderson, Lloyd 1985. The 'Perfect' as a Universal and as Language-Particular Category. In Hopper, Paul J. (ed.) *Tense-Aspect: Between Semantics & Pragmatics*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 227–264.
- Anderson, Lloyd 1986. Evidentials, path of change, and mental maps: Typologically regular asymmetries. In: Chafe, Wallace & Johanna Nichols (eds), 273–312.
- Boas, Franz 1938. Language. In: Boas, Franz (ed.), *General Anthropology*. Boston, New York: D. C. Heath and Company. 124–145.
- Bokova, E. N. 2002. *Èvèdy fol'klor 5–6* [Even folklore 5–6]. Jakutsk: Bičik.
- Boye, Kasper 2012. *Epistemic meaning: A cross-linguistic and functional-cognitive study*. Berlin: Mouton de Gruyter.
- Boye, Kasper 2018. Evidentiality: The notion and the term. In: Aikhenvald (ed.) 2018, 261–272.
- Burgherr, Esther 2017. Prominence in Even Narratives. MA Dissertation, Redcliffe College, University of Gloucester.
- Bybee, Joan, Revere Perkins & William Pagliuca 1994. *The Evolution of Grammar: tense, aspect, and modality in the languages of the world*. Chicago and London: The University of Chicago Press.
- Cardona, George 2003. Sanskrit. In: Cardona, George & Dhanesh Jain (eds), *The Indo-Aryan Languages*. London and New York: Routledge. 115–178.
- Chafe, Wallace L. & Johanna Nichols (eds). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex.
- Comrie, Bernard 1976. *Aspect*. Cambridge: Cambridge University Press.

- Cornillie, Bert 2009. Evidentiality and epistemic modality. On the close relationship between two different categories. *Functions of Language* 16:1, 44–62.
- Crystal, David. 2003 (2007). *A Dictionary of Linguistics & Phonetics*, Fifth Edition. Malden: Blackwell Publishing.
- Curnow, Timothy Jowan 2003. Nonvolitionality expressed through evidentials. *Studies in Language* 27:1, 39–59.
- Cysouw, Michael, Martin Haspelmath & Andrej Malchukov 2010. Introduction to the Special Issue: “Semantic Maps: Methods and Applications”. *Linguistic Discovery* 8:1, 1–3.
- Dahl, Östen (ed) 2000. *Tense and Aspect in the Languages of Europe*. Berlin, New York: Mouton de Gruyter.
- Dankoff, Robert 1982, editor and translator in collaboration with James Kelly: al-Kāšyarī, Maḥmūd, *Compendium of the Turkic Dialects (Dīwān Luyāt at-Turk)*. Part I. Cambridge, Mass.: Harvard University Printing Office. Available at [https://ia802800.us.archive.org/32/items/CompendiumOfTheTurkicDialectsPart1-MahmudAl-Kashghari/Al-K%C4%81%C5%A1%D0%B3ar%C4%AB%2C%20Mah%CC%A3m%C5%ABd.-Compendium%20of%20the%20Turkic%20dialects%20%28D%C4%ABw%C4%81n%20lug%C4%81t%20at-Turk%29%2C%20Part%20I%20%281\\_3%29.pdf](https://ia802800.us.archive.org/32/items/CompendiumOfTheTurkicDialectsPart1-MahmudAl-Kashghari/Al-K%C4%81%C5%A1%D0%B3ar%C4%AB%2C%20Mah%CC%A3m%C5%ABd.-Compendium%20of%20the%20Turkic%20dialects%20%28D%C4%ABw%C4%81n%20lug%C4%81t%20at-Turk%29%2C%20Part%20I%20%281_3%29.pdf)
- DeLancey, Scott. 2001. The mirative and evidentiality. *Journal of Pragmatics* 33, 369–382.
- Desnitsky, Andre, Fanuza Gimaeva, Teija Greed, Elena Kazenina & Alfiya Shamgunova 2002. Comments on Tatar Discourse Structure. In: Clifton, John M. & Deborah A. Clifton (eds), *Comments on Discourse Structures in Ten Turkic Languages*. Dallas: North Eurasia Group, SIL International. 81–103.
- Dixon, R. M. W. 2006. Complement Clauses and Complementation Strategies in Typological Perspective. In: Dixon, R. M. W. & Alexandra Y. Aikhenvald (eds), *Complementation*. Oxford: Oxford University Press. 1–48.
- Donet, Charles, M. 2014. The Importance of Verb Saliency in the Followability of Lezgi Oral Narratives. MA Thesis. Graduate Institute of Applied Linguistics.
- Dooley, Robert A. & Stephen H. Levinsohn 2001. *Analyzing Discourse: A Manual of Basic Concepts*. Dallas, Texas: SIL International.
- Dutkin, X. I. 1995. *Allaixovskij govor évenov Jakutii*. Sankt-Peterburg: Nauka.
- Forker, Diana 2018a. Evidentiality and its relations with other verbal categories. In: Aikhenvald (ed.) 2018, 65–84.
- Forker, Diana 2018b. Evidentiality in Nakh-Daghestanian languages. In: Aikhenvald (ed.) 2018, 490–509.
- Ganenkov, Dmitry, Timur Maisak & Solmaz Merdanova 2009. The origin and use of quotative markers in Agul. Quotative markers: origins and use, a workshop at the 42<sup>nd</sup> Annual Meeting of the Societas Linguistica Europaea, University of Lisbon, 9–12 Sept., 2009.

- Georgakopoulos, Thanasis & Stéphane Polis 2018. The semantic map model: State of the art and future avenues for linguistic research. *Language and Linguistics Compass*, 12: 2, 1–33.  
Available at <https://onlinelibrary.wiley.com/doi/abs/10.1111/lnc3.12270>.
- Gipper, Sonja 2011. *Evidentiality and Subjectivity in Yurakaré: An Interactional Account*. PhD Dissertation, Radboud University Nijmegen. <[www.mpi.nl/publications/escidoc-1086666](http://www.mpi.nl/publications/escidoc-1086666)>, accessed 2 June 2018.
- Greed, Teija 2009. Evidentiality in Tatar. Secondary MA Thesis for General Linguistics. University of Helsinki.  
Available at [https://www.academia.edu/6988647/Evidentiality\\_in\\_Tatar\\_2009](https://www.academia.edu/6988647/Evidentiality_in_Tatar_2009).
- Greed, Teija 2014 (2016). The expression of knowledge in Tatar. In: Aikhenvald & Dixon (eds), 69–88.
- Greed, Teija 2017. Evidential coding in Lezgi. *Languages of the Caucasus*. Volume 2. Available at <https://escholarship.org/uc/item/0w59z5t4>
- Greed, Teija 2018. The quotative in Bashkir. *Studia Orientalia Electronica*, Volume 6, 23–55. Available at <https://journal.fi/store/article/view/69685>.
- Greed, Teija 2019 (in press). From perfect to narrative tense: The development of an evidential meaning examined generally and in the Even language. *Studies in Language* 42:4, 923–966.  
Will be available via <https://www.benjamins.com/catalog/sl>.
- Grice, H. Paul 1989. *Studies in the way of words*. Cambridge, Mass: Harvard University Press.
- Guentchéva, Zlatka (ed.) 1996. *L'énonciation médiatisée. Volume 1*. Louvain, Paris: Éditions Peeters.
- Guentchéva, Zlatka & Jon Landaburu (eds) 2007. *L'énonciation médiatisée II. Le traitement épistémologique de l'information: illustrations amérindiennes et caucasiennes*. Louvain, Paris: Éditions Peeters.
- Güldemann, Tom 2008. *Quotative Indexes in African Languages: A Synchronic and Diachronic Survey*. Berlin, New York: Mouton de Gruyter.
- Haspelmath, Martin 1993. *A Grammar of Lezgian*. Berlin, New York: Mouton de Gruyter.
- Haspelmath, Martin 1994. The Tense System of Lezgian. In: Thieroff, Rolf & Joachim Ballweg (eds), *Tense systems of European languages*. Tübingen: Max Niemeyer Verlag. 267–277.
- Haspelmath, Martin 2003. The Geometry of Grammatical Meaning: Semantic Maps and Cross-Linguistic Comparison. In: Tomasello, Michael (ed.), *The New Psychology of Language. Cognitive and Functional Approaches to Language Structures. Volume 2*. Mahwah, London: Lawrence Erlbaum Associates, Publishers. 211–242.
- Hill, Jane H. and Irvine, Judith T. 1992. Introduction. In: Hill, Jane H. and Judith T. Irvine (eds), *Responsibility and evidence in oral discourse*. Cambridge University Press: Cambridge. 1–23.

- Isxakova, X. F., D. M. Nasilov, I. A. Nevskaja & I. V. Šencova 2007. Èvidencial'nost' v tjurkskix jazykax. In: Xrakovskij (ed.), 467–518.
- Jacobsen, William H. Jr. 1986. The Heterogeneity of Evidentials in Makah. In: Chafe & Nichols (eds), 3–28.
- Jakobson, Roman 1971. Shifters, Verbal Categories, and the Russian Verb. In: Jakobson, Roman, *Selected Writings. Vol II*. The Hague: Mouton. 130–147.
- Janhunen, Juha 2005. Tungusic: An endangered language family in Northeast Asia. In: Bradley, David (ed.), *Language Endangerment in the Sinosphere*. International Journal of the Sociology of Language 173. Berlin: Mouton de Gruyter. 37–54.
- Johanson, Lars 2000. Turkic indirectives. In: Johanson & Utas (eds), 61–87.
- Johanson, Lars 2003. Evidentiality in Turkic. In: Aikhenvald, Alexandra Y. & R. M. W. Dixon (eds), *Studies in Evidentiality*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 273–290.
- Johanson, Lars & Bo Utas (eds) 2000. *Evidentials. Turkic, Iranian and Neighbouring Languages*. Berlin: Mouton de Gruyter.
- Juldaš 2017 = <http://yuldashfm.ru/ba/guzl-mostafina-tteldy>, accessed 10 June 2017.
- Langacker, R. W. 1990. Subjectification. *Cognitive Linguistics* 1, 5–38.
- Langacker, Ronald W. 1995. Raising and Transparency. *Language*, 71:1, 1–62.
- Leinonen, Marja & Maria Vilkuna 2000. Past tenses in Permic languages. In: Dahl (ed.), 495–514.
- Levin, V. I. 1936. *Kratkij èvensko-russkij slovar'*. Moskva, Leningrad: Učpedgiz.
- Lewis, Geoffrey 2000 [1967]. *Turkish Grammar*. Second edition. Oxford: Oxford University Press.
- Lezgian Corpus 2011. Dmitrij S. Ganenkov.  
[http://daglanguages.org/LezgianCorpus/search/index.php?interface\\_language=r](http://daglanguages.org/LezgianCorpus/search/index.php?interface_language=r).
- Lindstedt, Jouko 2000. The Perfect – aspectual, temporal and evidential. In: Dahl (ed.), 365–383.
- Majsak, Timur 2012. Pričastnye formy v vido-vremennoj sisteme agul'skogo jazyka. In: Plungjan, Vladimir A. (ed.), *Acta Linguistica Petropolitana. Vol. VIII, part 2*. St Petersburg: Nauka. 228–289.
- Majsak, Timur A. & Solmaz R. Merdanova 2002. Kategorija èvidencial'nosti v agul'skom jazyke. *Kavkazovedenie* 1, 102–112.
- Majsak, Timur A. & Solmaz R. Merdanova 2016. Perfekt i smežnye značeniya v agul'skom jazyke. In: Majsak, Timur A., Vladimir A. Plungjan & Ksenija P. Semenova, eds. *Acta Linguistica Petropolitana. Vol. XII, part 2*. St Petersburg: Nauka. 379–422.
- Mushin, Ilana 2001. *Evidentiality and Epistemological Stance*. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Musin, Noyman 1987. *Keše huqmayy*. Ufa: Bašqortostan kitap nāšriäte.
- Malchukov, Andrej L. 2000. Perfect, evidentiality and related categories in Tungusic languages. In: Johanson & Utas (eds), 441–469.

- Narrog, Heiko 2012. *Modality, Subjectivity, and Semantic Change*. A Cross-Linguistic Perspective. Oxford: Oxford University Press.
- Nedjalkov, V. P. & S. J. Jaxontov 1988. The typology of resultative constructions. In Nedjalkov, V. P. (ed.), *Typology of resultative constructions*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 3–62.
- Nuyts, Jan 2001. Subjectivity as an evidential dimension in epistemic modal expressions. *Journal of Pragmatics* 33, 383–400.
- Pakendorf, Brigitte 2015. A comparison of copied morphemes in Sakha (Yakut) and Èven. In: Gardani, Francesco, Peter Arkadiev & Nino Amiridze (eds), *Borrowed Morphology*. Berlin/Boston/Munich: Walter de Gruyter, Inc. 157–187.
- Palmer, Frank. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.
- Plungian, V. A. 2001. The place of evidentiality within the universal grammatical space. *Journal of Pragmatics* 33, 349–358.
- Robbek, V. A. & M. E. Robbek (compilers) 2009. *Ètév nimkarin* [My grandfather's stories]. Text edited into Even by E. V. Nesterova. Jakutsk: Bičik.
- Russian Census 2010.  
[http://www.gks.ru/free\\_doc/new\\_site/perepis2010/croc/perepis\\_itogi1612.htm](http://www.gks.ru/free_doc/new_site/perepis2010/croc/perepis_itogi1612.htm).  
 Accessed 17 March 2018.
- Schultze-Berndt, Eva & Nikolaus P. Himmelmann 2004. Depictive secondary predicates in cross-linguistic perspective. *Linguistic Typology* 8, 59–131.
- SG = Korkina, E. I., E. I. Ubrjatova, L. N. Xaritonov & N. E. Petrov 1982. *Grammatika sovremennogo jakutskogo literaturnogo jazyka*. Moskva: Izdatel'stvo Nauka.
- Simons, Gary F. & Charles D. Fennig (eds) 2018. *Ethnologue: Languages of the World, Twenty-first edition*. Dallas, Texas: SIL International. Online version: <http://www.ethnologue.com>.
- Speas, Margaret 2018. Evidentiality and formal semantic theories. In: Aikhenvald (ed.) 2018, 286–311.
- Storch, Anne 2018. Evidentiality and the expression of knowledge: An African perspective. In: Aikhenvald (ed.) 2018, 610–628.
- Squartini, Mario & Pier Marco Bertinetto 2000. The Simple and Compound Past in the Romance Languages. In: Dahl (ed.), 402–439.
- Sun, Jackson T.-S. 2018. Evidentials and person. In: Aikhenvald (ed.) 2018, 47–63.
- TG = *Tatarskaja grammatika* 1993. Tom II, *Morfologija*. Kazan: Tatarskoe knižnoe izdatel'stvo.
- Tatevosov, Sergei. 2007. Evidentiality and Mirativity in the Mishar dialect of Tatar. In: Guentchéva & Landaburu (eds), 407–433.
- Willett, Thomas 1988. A Cross-linguistic Survey of the Grammaticization of Evidentiality. *Studies in Language* 12-1, 51–97.
- Williams, Michael J. 2001. *Problems of knowledge. A critical introduction to epistemology*. Oxford: Oxford University Press.

- Xrakovskij, V. S. 2007, *Ėvidencial'nost' v jazykax Evropy i Azii. Sbornik statej pamjati Natalii Adreevny Kozincevoj*. Sankt-Peterburg: Nauka.
- Zakiev, M. Z. 1992. *Tatarskaja grammatika*, Tom III, *Sintaksis*. Kazan: Tatarskoe knižnoe izdatel'stvo.
- Zakirova, Rämzija 2012. 'Universiadada tatar tele kiräk ikän...' [The Tatar language is apparently needed in the University Games...], at <http://tatclub.net/2012/03/universiadaga-tatar-tele-kir%D3%99k-ik%D3%99n/>, 13 March 2012, accessed 28 May 2012.
- Zippel, Luise 2012. *Finitheit im Ewenischen: Eine Untersuchung verbaler (In)finität unter besonderer Berücksichtigung des dialektalen Vergleichs des Perfektpartizips auf -čE*. Magisterarbeit. Institut für Linguistik der Universität Leipzig.



**PART II:**  
**Articles**

**List of Articles:**

Article 1: The expression of knowledge in Tatar

Article 2: Evidential coding in Lezgi

Article 3: The quotative in Bashkir

Article 4: From perfect to narrative tense: The development of an evidential meaning examined generally and in the Even language