

Linguistic Realization of Evidentiality in European Languages

Empirical Approaches to Language Typology

49

Editors

Georg Bossong

Bernard Comrie

Yaron Matras

De Gruyter Mouton

Linguistic Realization of Evidentiality in European Languages

Edited by

Gabriele Diewald

Elena Smirnova

De Gruyter Mouton

ISBN 978-3-11-022396-5
e-ISBN 978-3-11-022397-2
ISSN 0933-761X

Library of Congress Cataloging-in-Publication Data

Linguistic realization of evidentiality in European languages / edited by
Gabriele Diewald, Elena Smirnova.

p. cm. — (Empirical approaches to language typology ; 49)

“This volume contains a selection of contributions to the workshop
“Linguistic realization of evidentiality in European languages”, held at
the 30th Annual Convention of the German Society of Linguistics in
Bamberg (February 27–29, 2008).”

Includes bibliographical references and index.

ISBN 978-3-11-022396-5 (alk. paper)

1. Evidentials (Linguistics) 2. Typology (Linguistics) I. Diewald,
Gabriele. II. Smirnova, Elena.

P325.5.E96L56 2010
415—dc22

2010042027

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data are available in the Internet at <http://dnb.d-nb.de>.

© 2010 Walter de Gruyter GmbH & Co. KG, Berlin/New York

Printing: Hubert & Co. GmbH & Co. KG, Göttingen

∞ Printed on acid-free paper

Printed in Germany

www.degruyter.com

Table of contents

Introduction	
Evidentiality in European languages: the lexical-grammatical distinction	1
<i>Gabriele Diewald and Elena Smirnova</i>	
Types of verbal evidentiality marking: an overview	15
<i>Vladimir Plungian</i>	
Hearsay in European languages: toward an integrative account of grammatical and lexical marking	59
<i>Björn Wiemer</i>	
Information source in Spanish and Basque: a parallel corpus study	131
<i>Asier Alcázar</i>	
Embedded evidentials in German	157
<i>Mathias Schenner</i>	
Embedding indirective (evidential) utterances in Turkish	187
<i>Hatice Coşkun</i>	
Epistemic modality and evidentiality and their determination on a deictic basis: the case of Romance languages	223
<i>Gerda Haßler</i>	
Evidentiality, polysemy, and the verbs of perception in English and German	249
<i>Richard Jason Whitt</i>	
Evidential markers in French scientific writing: the case of the French verb <i>voir</i>	279
<i>Francis Grossman and Agnès Tutin</i>	
An interactional approach to epistemic and evidential adverbs in Spanish conversation	309
<i>Bert Cornillie</i>	

Revelative evidentiality in European languages: linguistic marking and its anthropological background <i>Alexandra Kratschmer and Adriënne Heijnen</i>	331
Subject index	369

Introduction

Evidentiality in European languages: the lexical-grammatical distinction¹

Gabriele Diewald and Elena Smirnova

The term “evidentiality” refers to a semantic-functional domain, which, beyond being expressed by a great variety of lexical means, is a relevant category in the grammatical systems of numerous of the world’s languages. Like “modality” or “temporality” the notion of evidentiality covers a range of meanings that may serve referential as well as non-referential purposes, or – to use a different diction – evidential meanings range from lexical to grammatical functions. In order to capture this, we speak of a “semantic-functional” domain here.

The basic characteristic of linguistic evidentiality is the explicit encoding of a source of information or knowledge (i.e. evidence) which the speaker claims to have made use of for producing the primary proposition of the utterance. The type of evidence the speaker adduces may be of various kinds (i.e., different modes of perception and cognition), which may be encoded by the evidential expressions but also may be left unspecified. This common core is reflected in most definitions of linguistic evidentiality as the following selection illustrates:

Evidentials may be generally defined as markers that indicate something about the source of the information in the proposition. (Bybee 1985: 184)

Evidentials express the kinds of evidence a person has for making factual claims. (Anderson 1986: 273)

[Evidentials express], how the speaker obtained the information on which s/he bases an assertion. (Willett 1988: 55)

Evidentiality proper is understood as stating the existence of a source of evidence for some information; this includes stating that there is some evidence, and also specifying what type of evidence there is. (Aikhenvald 2003: 1)

In the past decades research on evidentiality and on the expression of evidential meanings in language has made enormous progress. In particular,

there has been growing acceptance of the assumption that evidentiality is a semantic-functional domain in its own right, and not a sub-division of epistemic modality. Accordingly, it is acknowledged by many scholars that evidential markers and evidential systems in languages are in principle – i.e., notwithstanding the empirical fact of frequent overlap – an independent category (cf. de Haan 2001 and Aikhenvald 2004:7).

In much of the prior work on evidentiality it has been generally assumed that European languages typically lack grammatical markers and grammatical systems of evidentiality. Despite the fact that evidential expressions of Balkan languages have been described in several studies (cf. e.g. Friedman 1986, 2000, 2003, Matras 1995, Guentchéva 1993, 1996), a detailed description of the language-specific realizations of evidentiality is yet to be carried out for most of European languages.

This volume is a step towards that aim. It attempts at providing classificatory reflections as well as empirical facts about languages that have various – lexical as well as grammatical – evidential expressions. Most papers of this volume originate in presentations given at the workshop “The linguistic realization of evidentiality in European languages”, held at the 30th Annual Convention of the German Society of Linguistics in Bamberg (February 27-29, 2008). The papers by Schenner and Coşkun are additional contributions, which have been especially commissioned for this volume. The paper by Haßler was not held at the workshop but was submitted afterwards.²

The scope of the ten papers collected here ranges from broad overviews on areal and/or typological issues to in-depth studies of evidential expressions in particular languages, some taking the perspective of synchronic contrastive comparison, others focussing on diachronic investigation. Furthermore, there are contributions that take a broader view by combining linguistic investigation with cultural and anthropological aspects. Thus, beyond the typologically informed contributions, which survey a variety of languages (Plungian, Wiemer), there are studies on particular evidential phenomena in sets of Romance and Germanic languages (Haßler, Kratschmer and Heijnen), as well as detailed examinations of evidentials expressions in French (Grossmann and Tutin), German (Schenner, Whitt), English (Whitt), Spanish (Alcázar, Cornillie), Basque (Alcázar), and Turkish (Coşkun).

Beyond a sample of rich and original empirical data on types of evidential expressions in European languages, their respective degrees of grammaticalization and the kinds of function evidentials fulfil, the volume provides new insights into the following more specific issues:

- evidential (sub)systems and paradigms in different languages: semantic distinctions and pragmatic functions;
- interrelations of evidential expressions with other grammatical categories (e.g. mood, tense, aspect);
- syntactic and pragmatics issues concerning the restrictions on evidentials in subordinate clauses;
- the diachronic rise of evidential markers and the degrees of grammaticalization of evidentials in different languages;
- lexical sources of evidential markers in the domain of verbs of perception;
- frequency, distribution and specific functions of evidential markers in different text types and/or registers;
- metalinguistic devices of marking information sources according to anthropological parameters.

Before introducing the papers individually, one central issue has to be discussed as it is addressed in virtually every paper of this volume, and is – so to speak – the red thread running through its argumentative texture. It is the dispute raised by Aikhenvald (2004) who argues for a very restrictive definition of evidentials and in all declines the justification for assuming that European languages possess evidentials. This point addresses the fundamental question of the grammatical-lexical continuum in expressions of evidentiality, and therefore is worth devoting some attention to from a theoretical viewpoint as well.

Aikhenvald (2004: 3-11) takes a strong stand concerning the distinction between evidential markers proper on the one hand and what she calls evidential strategies, on the other. She claims that European languages do not display evidential markers and that most evidential phenomena in those languages are “mere” evidential strategies. In her argument Aikhenvald points out that the fact that modal verbs in European languages often have evidential overtones and usages is neither an argument for treating evidentiality as a subcategory of epistemic modality nor for claiming that European languages have evidentials. True as this is, the following should be kept in mind: the fact that modals are not evidentials does not mean that European languages lack evidential systems, or evidential constructions.

Furthermore, Aikhenvald emphasizes that the existence of adverbials with evidential meaning, which are the object of research of many scholars, is by no means proof of a grammatical category for evidential distinctions (2004: 5-7). This again has to be agreed upon. Nevertheless, a closer and more differentiated look at European languages is called for. In particular,

it has to be considered that grammatical systems *develop*, thereby producing a full scale of phenomena with intermediate values concerning the cline between lexical and grammatical. Moreover, in the course of developing new evidential functions (which may be grammaticalized later), a particular item (or construction) is bound to run through early developmental stages where evidential functions are contextually induced via conversational implicature and non-evidential functions still prevail to later developmental stages where evidential functions are part of the item's inherent semantic structure (i.e. they are semanticized or conventionalized) no matter whether older, more lexical usages still co-exist beside the new meanings or not.

Though Aikhenvald (2004: 11) admits that change leads to clines and intermediate values, she insists that only fully grammaticalized evidential systems are worth being treated under the label of evidentiality, and consequently proposes to restrict the term evidentials to Non-European languages which do have fully grammaticalized *inflectional* systems.

The reason for this position may lie in the fact that Aikhenvald herself does not always distinguish properly between structural issues and semantic domains. This becomes clear in the following quotation:

In about a quarter of the world's languages, every statement must specify the type of source on which it is based – for example, whether the speaker saw it, or heard it, or inferred it from indirect evidence, or learnt it from someone else. *This grammatical category, whose primary meaning is information source, is called 'evidentiality'.* (Aikhenvald 2004: 1) [our emphasis]

Equating the grammatical category with the name of the conceptual domain, i.e. “evidentiality”, is analogous to equating the grammatical category “tense” with the conceptual domain “temporality”.³ This tight and impermeable association of conceptual domain and grammatical category prevents an adequate evaluation of European evidential expressions (with their epistemic overtones, i.e. their semantic conflation) as well as an unbiased look at the chronological development, i.e. at the fact that grammaticalization by necessity implies forms not yet fully grammaticalized.

To give an example (which is discussed in detail in Diewald and Smirnova 2010): the German evidential constructions *werden* & infinitive and *scheinen/drohen/versprechen* & *zu*-infinitive, like many analogous constructions in other languages found in the Indo-European family, clearly are of an intermediate stage as concerns the degree of grammaticalization. They are not yet full-fledged grammaticalized evidential systems as compared to those systems invoked by Aikhenvald, which have inflectional or

clitic evidential markers, but they are instances of evidential systems on the rise (e.g. Cornillie 2007, Dendale and Tasmowski 2001, Diewald 2004, Mortelmans 2000, Plungian 2001).

Thus, we contend that, since grammaticalization is a gradual process, it would be counterproductive to restrict attention to fully developed grammaticalized systems. Linguistic research can only account for the make-up and functioning of evidential systems, if it does not exclude evolving systems and their interplay with lexical and other means in expressing evidential values from the agenda.

Therefore, it does not come as a surprise that Aikhenvald's radical position presented in her 2004 book is unequivocally contested by all contributors to this volume. Wiemer explicitly applies the definition given by Aikhenvald (2003: 1) not only to grammatical markers but also to lexical expressions in different European languages. Whitt discovers evidential functions of German and English constructions with verbs of perception which display different degrees of grammaticalization, i.e. which are situated at different points of the lexical-grammatical continuum. Alcázar shows that Basque obviously has a system of evidential markers (particles) which interacts with an array of further evidential strategies in the same language. Schenner proposes a definition of evidential expressions which is meant to equally apply to lexical as well as to grammatical linguistic elements. Schenner argues for treating German verbal periphrastic forms as an evidential system. Coşkun focuses on Turkish which without a doubt does have evidential markers in the strict sense, which however – as she shows – at the same time interact with other phenomena to produce evidential and further effects. Kratschmer and Heijnen draw a clear distinction between grammatical markers and (simple or composed) lexical expressions, and insist that both types of formal expression are valid objects of investigation for those interested in the study of evidential marking. Grossmann and Tutin concentrate on lexical means expressing evidentiality, and emphasize the importance of analyzing lexical elements as they are often potential predecessors of grammaticalized evidentials. In treating epistemic and evidential adverb and adverbials in Spanish, Cornillie assumes that evidentiality is an important linguistic phenomenon in European languages.

Thus, Aikhenvald's verdict has to be considerably relativized: European languages, or better, languages spoken in Europe may not have the most complex or formally and functionally differentiated grammaticalized systems of evidential markers found in the world's languages, but there is a large number of them which clearly have grammaticalized systems of evi-

dential markers, and there is an even larger number with evidential systems in their initial or intermediated stages of development and/or with a great variety of stereotypical, automatized lexical expressions for that purpose. In short, all authors of this volume consent that it is well worth studying evidential markers as well as evidential strategies in European languages, which – needless to mention it – does of course not question the fact that “true” evidential markers – i.e. maximally grammaticalized evidential markers – are (morphologically bound) grammatical formatives.

The order chosen for the contributions starts from the general to the more specific. The first two papers present typological surveys (Plungian, Wiemer), followed by papers focussing on more structural, grammatical issues concerning particular languages (Spanish and Basque by Alcazár, Turkish by Coşkun, German by Schenner). The next contributions concentrate on semantic issues, in particular on evidentials having lexical sources in the domain of perception verbs (Whitt on evidentiality, polysemy, and verbs of perception in English and German; Grossmann and Tutin on French *voir* as a metalinguistic device in written scientific discourse), as well as semantic considerations concerning the borders between evidentiality and modality (Haßler on various issues in Romance Languages). Cornil-lie takes up the topic of functional extensions of evidentials (and epistemic) expressions in spoken discourse, thus shifting the focus towards central aspects of dialog analysis like the turn-taking mechanism. Finally, the field is extended to „revelative evidentiality” in an interdisciplinary approach combining linguistic and anthropological research (Kratschmer and Heijnen on linguistic marking of revelative evidentiality in Icelandic, German, and Italian).

Vladimir Plungian presents a general overview on evidentiality and on the study of evidentiality. In his paper “Types of verbal evidentiality marking: an overview”, he gives a brief overview of grammatical evidential systems which is based on generalizations proposed in the typological studies. After a short description of the areal distribution of evidentiality-marking systems, he discusses the history of linguistic studies in evidentiality. Furthermore, he presents a cross-linguistic classification both of evidential values and evidential systems, whereby a “basic” system is taken as a point of departure. The basic system is considered the prototypical realization of a generalized typological idea of evidentiality in a most neutral way. Plungian concludes with a discussion of the relation between evidentiality and modality. He argues to treat them as two different semantic domains which are closely related synchronically as well as diachronically. He notes that the nature of the relation between evidentiality and

modality cannot be explained in a general way since it depends on the concrete evidential system in a concrete language under investigation. In this connection Plungian proposes to distinguish between “modalized” and “non-modalized” evidential systems.

The contribution by **Björn Wiemer** offers a first general cross-classification of linguistic devices with reportive meaning found in European languages (whereby he takes “reportive” as synonymous to “hearsay” and as a sub-domain of evidentiality). The title of this contribution “Hearsay in European languages: towards an integrative account of grammatical and lexical marking” is programmatic and takes a clear position contra Aikhenvald’s 2004 stance. The layout of this study is empirical and classificatory, and, though acknowledging the distinction between epistemic modality and evidentiality the author aims at a maximally inclusive survey, i.e. at collecting as many reportage expressions as possible. The main body of the papers shows that these – i.e. expressions of reportive functions – range from bound (inflectional and agglutinative) morphology, like e.g. in Georgian languages and Turkish, via functional extensions of various verbal grammatical categories, like tense, mood and aspect markers (e.g. the analytic perfect as in Baltic and Balkan languages), extensions from future grams, past tense grams (e.g. the Italian imperfect), moods (e.g. French and Portuguese) and modal auxiliaries (e.g. German), to lexical means and constructions (like volitional verbs, predicative constructions, sentence adverbials and particles). Due to this overall scheme, the paper is not primarily concerned with language change or grammaticalization-lexicalization clines, but instead takes a purely empirical, “surface-oriented” perspective. Thus, it makes a valuable first move towards a comprehensive view on the areal distribution in Europe’s languages of expressions of reportive functions.

In the paper “Information source in Spanish and Basque: a parallel corpus study” **Asier Alcázar** provides a contrastive study of grammatical evidential markers in Basque and non-grammaticalized evidential strategies in Spanish. Using a Spanish-to-Basque parallel corpus he shows that Basque – unlike Spanish – is a language with several grammaticalized evidential expressions (particles) which fit into typologically acknowledged evidential systems. Special focus is put on reported evidence expressed by the Basque particle *omen* appearing before the tensed verb. The investigation of the translation practice in the Spanish-to-Basque parallel corpus reveals that *omen* is uniformly chosen to translate a variety of Spanish evidential expressions, and, moreover, that *omen* is frequently inserted even if there is no evidential expression in the Spanish original. Furthermore, *omen* is only

used when the relevant Spanish expression has sentential scope. On the basis of these findings the author concludes that Basque *omen* is clearly a grammatical evidential marker with a defined functional spectrum ('hear-say'), obligatory sentential scope, and increased obligatoriness. Further – though less grammaticalized – members of the evidential particle paradigm in Basque are seen in the particles *bide* 'apparently', *edo* 'inferential' ('must'), *ote* 'dubitative', *ohi* 'habitual'. As Basque, in addition to the evidential particles, has further expressions to render evidential meaning, in particular in cases of constituent scope ambiguity or semantic ambiguity, Basque is an exemplar case of the interplay of grammaticalized means and lexical meanings in the expression of evidential values and functions.

Based on corpus as well as questionnaire data and using formal semantics as a descriptive tool, **Mathias Schenner** in his paper "Embedded evidentials in German" tackles two questions concerning evidentials in German. First, he takes up the discussion on whether German (representative of further „European languages") does have evidentials in the true sense of the word. Here, again, Aikhenvald's position is contradicted, and Schenner argues for accepting a semantically strict but formally liberal notion of evidentiality. Separating epistemic notions from evidential ones (i.e. arguing for treating evidentiality as a category of its own), he shows that the "reportive usage" of *sollen* in German qualifies this item as an evidential. With reference to Diewald and Smirnova [to appear] he further argues that there may be an even more elaborate paradigm of evidentials in German. Second, he takes up the question of the possibility of embedding evidentials in subordinate clauses. He demonstrates that – contrary to mainstream assumptions – embedded reportive uses of *sollen* can be found in German, whereby a range of different meanings of this item in embedded clauses has to be taken into account. His conclusion is that embedded reportive *sollen* can be used only with certain types of embedding predicates, whereby a particular type of the matrix predicate licenses a particular reading of *sollen*. Communication predicates correspond to the assertive reading of *sollen*, (semi-)factive predicates license the global reading, and certain negative predicates embed *sollen* in its concord reading.

Hatice Coşkun, in her paper on "Embedding indirective (evidential) utterances in Turkish", focuses on the interpretative span of the highly grammaticalized evidential marker *-mIs* in Turkish. This marker is usually labelled as expressing "indirectivity", i.e. an evidential category indicating that the propositional content of the utterance is reached in an *indirect way* ('it appears to the recipient that X is the case'). The author takes up the question which kinds of subordinate clauses license evidential meaning in

this marker. She demonstrates that – contra prior research, e.g. by Johanson 2000, 2006, according to which evidential sentence types in Turkic languages are restricted to asserted main clauses – in analytic types of embedding of subordinate clauses evidential readings arise in specific contextual constellations, whereby the semantic type of the matrix verb and other factors of the discourse-pragmatic type are essential for the actualization of different meanings. She suggests – similarly to Schenner – that there are five types of matrix predicates selecting finite complement clauses with indirective content, namely: communication verbs (e.g. *de-* ‘say’ or *anlat-* ‘explain’), cognition verbs (e.g. *bil-* ‘know, think’), perception verbs (*duy-* ‘hear’), attitude verbs (*san-/zannet-* ‘think, believe, consider’), evaluative or descriptive nominal predicates (*belli* ‘known’, *iyi* ‘good’). Beyond the result that evidential meanings may be embedded, this paper underpins the fact that very often meanings which elsewhere are realized by grammatical markers may arise due to what we might call “contextual collaboration”, i.e. the interplay of syntactic, morphological and semantic-pragmatic functions. Thus, although the author does not argue herself in this direction, the data and results presented in this paper strongly point towards the usefulness of constructional approaches in the interpretation of grammatical(izing) meanings.

Gerda Haßler’s contribution is entitled “Epistemic modality and evidentiality and their determination on a deictic basis: the case of Romance languages”. Beyond a succinct recapitulation of descriptions of evidential systems in typological research this paper discusses the relation between epistemic modality and evidentiality and its various expressions in Romance languages. The author applies a dialogically enriched, non-canonical notion of deixis to identify evidential meanings and demonstrates her findings using examples from French, Spanish, Italian and Portuguese corpora. The author claims that while evidentiality is a deictic category, epistemic modality is not; and that therefore “recognising the deictic character of evidentiality and describing it as deictic phenomenon helps to determine the place of evidentiality within the language and provides a basis for distinguishing it from epistemic modality.”

The contribution by **Richard Jason Whitt** on “Evidentiality, polysemy, and perception verbs: a corpus-based analysis of English and German” summarizes the results of a corpus-based analysis of evidential meanings signified by English and German perception verbs from the Early Modern period to the present, focusing on polysemy of such expressions on one hand and on the differences between German and English on the other. Beyond semantic issues the paper also addresses the importance of con-

structions for particular readings and specific features of perception verbs, in particular with respect to the distinction between subject-oriented and object-oriented perception (*ich schmecke das Salz – die Suppe schmeckt salzig*). The author shows that verbs of perception in English and German are polysemous and can express a variety of evidential meanings, i.e. their readings range from direct/perceptual to inferential and to hearsay evidential values, some of which are restricted to specific construction types. Furthermore, the similarities and differences between German and English perception verbs are highlighted, as well as the great importance of perception verbs for the expression of the speaker's epistemology and for the development of evidential markers.

The role of perception expressions within the domain of evidentiality is also the topic of the next paper. **Francis Grossman** and **Agnés Tutin** in their paper "Evidential markers in French scientific writing: the case of the French verb *voir*" analyze corpus data of French academic writings in linguistics and economics, thereby focusing on the different uses of the French verb *voir*. Different construction types of this perceptual verb are taken into account. The authors propose that *voir* displays five clearly distinct usages: (i) *voir* as a "statement marker", where it has the clear evidential inferential meaning which can be paraphrased as "inference based on observation"; (ii) *voir* as a "reference marker", where it refers to sources of knowledge and to textually present additional material; (iii) *voir* in the sense of 'to examine', which is not evidential; (iv) *voir* as a judgement – and non-evidential – verb; and (v) other non-evidential usages of the verb. The authors concentrate on the first two usages of the verb and their distribution and particular functions in linguistic and economic scientific texts. As *voir* is often used to specify the source of information in scientific texts, it is said to have an important role in validating research work. Admitting that the core semantics of the verb *voir* is its perceptual meaning, the authors argue that – used in academic writings – it may be used as an inferential evidential strategy, indicating the source of information the researcher has for making the claim and serving as a inter- and intratextual device to direct the reader's attention to pieces of evidence.

Shifting focus from written – monologic – language to spoken – dialogic – discourse, **Bert Cornillie** presents "An interactional approach to epistemic and evidentials adverbs in Spanish conversation". Although epistemic and evidential adverbs in the majority of uses in written language have the function of qualifying the proposition according to factuality values (in the case of epistemic markers), or according to information sources (in the case of evidentials), they display a variety of additional, interactional func-

tions in conversational contexts. Using corpus data on the epistemic adverb(ial)s *a lo mejor*, *igual*, *quizá*, *tal vez* (all meaning ‘maybe’, ‘perhaps’) and on the evidential expressions *aparentemente*, *evidentemente*, *obviamente*, *supuestamente*, *visiblemente*, the study aims at classifying various discourse functions of epistemic and evidential adverbs. The author shows that in both groups the items can be distinguished according to frequency as well as the prominences of additional interactional functions they take in dialogic discourse. Of the four the epistemics *a lo mejor* and *igual* have the additional function of marking the hypothesis as subject to a confirmation or refutation by the interlocutor, of the group of evidential adverbs, *obviamente* and *evidentemente* display interactional functions in contributing to the on-line planning concerning the turn-taking mechanism (in signalling that the speaker wants to keep the floor).

Adrienne Heijnen and **Alexandra Kratschmer** are concerned with “Revelative evidentiality in European languages: linguistic coding and its anthropological background”. Using data from biblical texts in Romance and Germanic languages as well as interviews and other data sources, they show the interrelations between linguistic coding of revelative evidentiality, which means “seeing in a dream”, on one hand and cultural and social traditions on the other. The authors treat revelative evidentiality as a sub-category of the – semantic – domain of evidentiality and propose to define it as follows: “information linguistically marked as created inside the mind of a subject without input from the outside world”. Using this definition, Heijnen and Kratschmer investigate which linguistic means (including grammatical as well as lexical elements) can be used in different languages to mark this semantic sub-category of evidentiality. For this they use large corpus data of biblical texts and interviews with native speakers of modern Icelandic, German and Italian. The authors conclude that there are no independent grammatical markers for revelative evidentiality in European Languages. Revelative evidentiality is often associated with visual evidential expressions and with inferential evidential expressions. Most often, however, there are lexical linguistic means to mark that information was obtained in a dream. With the help of their broad and valuable data material the authors are able to connect anthropological and linguistic questions, laying open the anthropological and cognitive foundations of linguistic structure. Furthermore, the authors raise methodological questions concerning the motivational and argumentative connection between language and culture, thus touching on the issue of linguistic relativity. Moreover, the authors propose that it is important to take the *interactive* aspect of evidentiality into account: they encourage typologist to investigate whether

languages have different markers for evidential meanings like ‘overhearing something’ vs. ‘(inter-)actively being told’.

With this array of studies the present volume makes an attempt to show that, if one is concerned with the notion of evidentiality, European Languages should not be excluded and are worth investigating. European Languages with their different ways of the linguistic realization of evidentiality often display not only lexical means for expressing evidential meanings but also have grammaticalized markers and even tightly organized grammatical paradigms which are in line with other typologically acknowledged evidential systems. By presenting theoretical considerations and by providing empirical evidence the papers of this volume aim at establishing a coherent view on the notion of evidentiality. Namely, they all emphasize that evidentiality should be seen as a semantic-functional (conceptual) domain which is not restricted to grammatical(ized) markers but can be realized by different linguistic expressions. Furthermore, in claiming that evidentiality and epistemic modality are two categories which are largely independent from each other, though often intertwined in individual languages and in individual expressions, this volume contributes to a better understanding of this particularly complex issue.

Notes

1. The workshop and this publication were funded by the Fritz Thyssen Stiftung as part of the research project “Evidential markers in German”. The work of the editors on this topic was furthermore supported by the Belgian Federal Grant P6/44 within the program of interuniversity attraction poles (IAP). The editors express their gratitude to the sponsors.
2. Cf. the workshop discussion platform under www.gabrielediewald.de
3. In other places in the book, Aikhenvald clearly is aware of this difference, cf.: “Saying that English parentheticals are ‘evidentials’ is akin to saying that time words like ‘yesterday’ or ‘today’ are tense markers”. (Aikhenvald 2004: 10).

References

- Aikhenvald, Alexandra Y.
 2003 Evidentiality in typological perspective. In *Studies in Evidentiality*, A. Y. Aikhenvald and R. M. W. Dixon (eds.), 1-31. Amsterdam/Philadelphia: John Benjamins.
 2004 *Evidentiality*. Oxford: Oxford U.P.

- 2007 Information source and evidentiality: what can we conclude? In *Evidentiality between lexicon and grammar*, Squartini, M. (ed.), *Rivista di Linguistica* 19-1: 209-227.
- Anderson, Lloyd B.
 1986 Evidentials, Paths of Change, and Mental Maps: Typologically Regular Asymmetries. In *Evidentiality: The Linguistic Coding of Epistemology*, Wallace Chafe and Johanna Nichols (eds.), 273-312. Norwood, NJ: Ablex.
- Bybee, Joan L.
 1985 *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam/Philadelphia: John Benjamins.
- Cornillie, Bert
 2007 *Evidentiality and epistemic modality in Spanish (semi-)auxiliaries: a cognitive-functional approach*. Berlin/New York: Mouton de Gruyter.
- Dendale, Patrick and Tasmowski, Liliane (eds.)
 2001 *Evidentiality* (Journal of Pragmatics 33, special issue). Amsterdam: Elsevier.
- Diewald, Gabriele
 2004 Faktizität und Evidentialität: Semantische Differenzierungen bei den Modal- und Modalitätsverben im Deutschen. In *Tempus/Temporalität und Modus/Modalität im Deutschen – auch in kontrastiver Perspektive*, Oddleif Leirbukt (ed.), 231-258. Tübingen: Stauffenburg.
- Diewald, Gabriele and Elena Smirnova
 2010 *Evidentiality in German: Linguistic Realization and Regularities in Grammaticalization*. Berlin/New York: Mouton de Gruyter.
- Friedman, Victor A.
 1986 Evidentiality in the Balkans: Bulgarian, Macedonian, and Albanian. In *Evidentiality: The Linguistic Coding of Epistemology*, Wallace Chafe and Johanna Nichols (eds.), 168-187. Norwood, NJ: Ablex.
- 2000 Confirmative/nonconfirmative in Balkan Slavic, Balkan Romance, and Albanian with additional observations on Turkish, Romani, Georgian and Lak. In *Evidentials*, Lars Johanson and Bo Utas (eds.), 329-366. Berlin/New York: Mouton de Gruyter.
- 2003 Evidentiality in the Balkans with special attention to Macedonian and Albanian. In *Studies in Evidentiality*, Alexandra Y. Aikhenvald and R.M.W. Dixon (eds.), 189-218. Amsterdam/Philadelphia: John Benjamins.
- Guentchéva, Zlatka
 1993 La catégorie du médiatif en bulgare dans une perspective typologique. *Revue des études slaves* 65(1): 57-72.

- 1996 Le Médiatif en Bulgare. In *L'énonciation médiatisée*, Zlatka Guentchéva (ed.), 47-70. Louvain/Paris: Peeters.
- Haan, Ferdinand de
 - 2001 The relation between modality and evidentiality. In *Modalität und Modalverben im Deutschen*, Marga Reis und Reiner Müller (eds.), 201-216. Hamburg: Buske.
- Johanson, Lars
 - 2000 Turkic indirectives. In *Evidentials: Turkic, Iranian and neighbouring languages*, Lars Johanson and BoUtas (eds.), 61-87. Berlin/New York: Mouton de Gruyter.
 - 2006 Indirective sentence types. *Turkic Languages* 10: 72-87.
- Lehmann, Christian
 - 1995 *Thoughts on Grammaticalization*. München/Newcastle: LINCOM Europa.
 - 2002 New reflections on grammaticalization and lexicalization. In *New reflections on grammaticalization*, Ilse Wischer and Gabriele Diewald (eds.), 1-18. Amsterdam/Philadelphia: John Benjamins.
- Matras, Yaron (ed.)
 - 1995 *Romani in Contact: The history, structure and sociology of a language*. Amsterdam/Philadelphia: John Benjamins.
- Mortelmans, Tanja
 - 2000 On the evidential nature of the epistemic use of the German modals *müssen* and *sollen*. In *Modal verbs in Germanic and Romance languages*, Johan van der Auwera and Patrick Dendale (eds.), 131-149. Amsterdam/Philadelphia: John Benjamins.
- Plungian, Vladimir A.
 - 2001 The place of evidentiality within the universal grammatical space. *Journal of Pragmatics* 33: 349-357.
- Willett, Thomas
 - 1988 A cross-linguistic survey of the grammaticalization of evidentiality. *Studies in Language* 12: 51-97.

Types of verbal evidentiality marking: an overview

*Vladimir A. Plungian*¹

The paper provides a brief overview of verbal systems with grammatical marking of evidentiality, based on previous case studies and generalizations proposed in the vast literature on the subject. The following topics are discussed: definitions of evidentiality as a verbal grammatical category; areal distribution of evidentiality-marking systems; the history of linguistic studies in evidentiality; possible cross-linguistic classifications both of evidential values and evidential systems; and the relation of evidentiality to other verbal categories (mainly, person and modality).

1. Introduction

1.1. Goals and outline

The primary goal of the present study is to provide an overview of the main types of existing systems where evidentiality is grammatically marked as a verbal category. This goal naturally requires a certain limitation of the material, which means that, as will be shown below, one can by no means expect an all-embracing description of the problem of evidentiality. We will primarily be concerned with the possible set of semantic parameters that are characteristic for evidentiality as a universal-linguistic *grammatical* category (or “cross-linguistic gram type” in terms of Bybee and Dahl 1989 or Bybee et al. 1994). Furthermore, evidentiality is construed here as a *verbal* category (i.e. as a category expressed by means of morphological or analytic modifications of the verbal forms). Not all generally possible types and means of expression of evidential values are accounted for with this definition. Thus, the various means of *lexical* expression of evidentiality in the languages of the world are not taken into consideration (for a discussion of this problem, which in the last time began to attract much more attention of typologists, see, for instance, the recently published edited volumes by Squartini 2007 and Wiemer and Plungjan 2009, as well as Squartini 2008 and Wiemer 2010). Similarly, also those markers will be

ignored which cannot be clearly analyzed as modifiers of the verb (for the possibility of expressing evidential values on nouns by means of different types of “sentential” markers whose categorial membership is not entirely clear, see, for example, Aikhenvald 2004; the problem needs further research).

However, the remaining material, i.e. that in which evidentiality occurs as a grammatical category of the verb, is still extensive, so that in the present overview we may concentrate only on some basic phenomena that seem to be particularly important and interesting for the problem in question. There are two main questions to which we would like to provide some answers: (i) possible types of *grammatical values* that may be found within the universal semantic domain of evidentiality, and (ii) possible types of *grammatical systems* that are used for expressing grammatical values (grammemes) of evidentiality in the languages of the world.

The discussion will be structured as follows. In Section 1.2 a preliminary definition of evidential values will be introduced. Furthermore, we will discuss the problem related to the distinction between phenomena that can be related to evidential functions on the one hand and those that have to be regarded as part of other functions on the other hand. (Since the distinction between evidential values and some other values is rather difficult, we will address it again at a more advanced stage, in Section 4). There will also be a brief overview of those languages in the world in which evidentiality is expressed grammatically (Section 1.3). Later, in Section 1.4, a brief outline of the history of the study of evidentiality that covers the main stages will be given. Sections two and three form the central part of the present study. Section two includes a detailed discussion of the inventory of the main types of evidential values and different suggestions concerning their classification (including our own ones, which are partly based on the works by Plungian 2000: 321-325 and Plungian 2001). Section 3 deals with a typology of evidential systems (taking the suggestions made in Aikhenvald's 2004 book into consideration, with some modifications). Section 4 offers a detailed discussion of the interrelation between evidentiality and two other grammatical categories: person (Section 4.1) and modality (Section 4.2). The distinction between evidentiality and modality is, as it is known, difficult to be drawn and often disputed in the typological literature. In the concluding Section some further tasks of the typological description of the category of evidentiality will be formulated.

Since the present study offers a generalization of the material collected by other researchers in the first way (and thus represents the result of an “indirect access” to the sources of information in terms of the category of

evidentiality) we will include a minimal quantity of linguistic examples only, restricting our references mainly to those works in which the material of the respective languages was analyzed by specialists in these languages. Our main goal is the documentation of the history of research and, related to it, the “history of ideas” on evidentiality, but also a discussion of the problem of finding the typologically most adequate universal classification of markers of evidentiality.

1.2. What is evidentiality: a preliminary characterization

According to the common view in current grammatical typology, markers are related to the semantic domain of evidentiality if their basic meaning is the indication of the “source of information” of the situation spoken about. In other words, while using a marker of evidentiality speakers communicate how they came to know about what they are talking about. If the system of such marking is obligatory (for instance, if it occurs as a grammatical category of the verb) speakers have to indicate on which basis they make an assertion about a respective situation with each use of a verbal form. To be more precise, one could say that in such cases the speakers indicate in what way they had access to the information referred to in a particular speech situation. As we shall see below when discussing the classification of evidential values, the conception of “access” appears to be somewhat more elastic than the conception of “knowledge” or that of an “acquisition of knowledge”, although they are very close to each other. However, the conception of access places more emphasis on the way in which a situation is perceived than on the mere epistemic aspect of the problem, and this obviously corresponds better to the way in which the category of evidentiality is organized in the languages of the world. From this point of view, the title (or better: the subtitle) of one of the most well-known edited volumes on the typology of evidentiality, Chafe and Nichols’ (1986) *The linguistic coding of epistemology*, appears to be somewhat imprecise since the problem of evidentiality is more related to the way in which an individual perceives the reality surrounding him than to the problem of epistemology as such.

In order to illustrate what evidentiality is we will reproduce the well-known example by Aikhenvald from a language which is usually considered one with a maximally differentiated expression of the values of this category, namely the languages of the Tucano family (Amazon basin).² According to Aikhenvald (2004: 52) the meaning of the phrase ‘the dog

stole the fish' can (or better: must) be translated into Tucano by means of at least one of four possible constructions, depending on the way in which the speakers had access to this information:

- (1) Tucano [Aikhenvald 2004: 52]; evidential markers are highlighted
- a. *diâyĩ wa 'i-re yaha-ámi*
'the dog stole the fish' (I saw it)
 - b. *diâyĩ wa 'i-re yaha-ásĩ*
'the dog stole the fish' (I heard the noise)
 - c. *diâyĩ wa 'i-re yaha-ápĩ*
'the dog stole the fish' (I inferred it)
 - d. *diâyĩ wa 'i-re yaha-úpĩ'*
'the dog stole the fish' (I was told)

All sentences listed in example (1) exhibit the order SVO (therefore the agent is unmarked, but the patient is marked by means of the special suffixal marker *-re*, which Aikhenvald calls the "topic non-subject case"). As we can see, the way in which the information was accessed is clearly reflected in the morphology of the verb. The four different grammemes of evidentiality distinguished in Tucano represent rather typical instances of evidential values (they will be characterized more in detail in Section 4). They either indicate that the speakers themselves directly observed the situation as viewers (a), or that the speakers observed the situation themselves and directly earlier, but not as viewers, i.e. they did not see it, but only heard what was going on (or perceived it in any other way) (b), or that the speakers had no direct access to the situation, but observed some facts which they interpret as having caused a particular situation, i.e. based on what they observed they can assume that the situation took place (c), or, finally, that the speakers themselves had neither access to the situation, nor to its consequences, but somebody else informed them that the situation took place (d).

A translation of the various sentences in example (1) into a language without grammatical markers of evidentiality would require the use of lexical equivalents in order to express the respective meaning. The most typical lexical equivalent to the meaning in (a) is the construction 'I saw that/how P' or 'P in my eyes'; an equivalent to the meaning of type (b) is 'I heard that P'; one expressing the meaning of type (c) is 'obviously/one may notice that P'; and one to type (d) – 'it is said that/it can be heard said that P or through the ears, P'. As lexical units such expressions are not very frequent. They are used only in such cases in which the speaker explicitly

wants to emphasize the respective meaning. As already mentioned, grammatical markers of evidentiality must occur with all forms of the verb (or at least with the majority of verbal forms marked for any other category, such as all finite forms, all forms of the past tenses, etc.), often independently of the speaker's intention.

1.3. Evidentiality in the languages of the world: the main areas

The obligatory indication of the source of information about an uttered fact in a linguistic system might seem entirely exotic to speakers of the majority of the European languages. Nevertheless, in a rather large number of languages on earth this strategy forms part of the grammatical rules of the use of verbal forms. Moreover, the number of languages in which evidentiality is grammaticalized is, as studies of the past years have shown, much larger than it is often assumed. Nowadays, evidentiality is not regarded as a rare and unusual category, as it was the case in the period following its "discovery" by European linguists. It will perhaps soon be more correct to say that the lack of grammatical marking of evidentiality in the majority of the languages of Western and Central Europe is an important linguistic particularity of this area, which distinguishes it from many others.

Currently, the following linguistic areas are known as agglomerations of languages in which evidentiality is expressed grammatically with verbs:³

A. The extended *Euro-Asian area*, including the Balkan languages (except for Modern Greek), Asia Minor and, from there, a broad strip going to the Far East over the Caucasus, Southern Asia, the Volga district, and Southern Siberia; in the linguistic literature this area is usually called the "Great Evidential Belt" and considered to be the most significant geographical locality of evidentiality, both with respect to the size of the territory and the number of languages and their genetic diversity (the Great Evidential Belt includes Southern Slavic, Albanian, Iranian, Indo-Arian, Armenian, Kartvelian, Abchaso-Adygei, Nachsko-Daghestan, Turkish, Finno-Ugric, and some other languages). It should, however, be mentioned that the type of grammaticalized evidentiality which is mainly represented in the languages of the Great Belt belongs to the group of more simple evidential systems (the so-called binary type, cf. Section 3), that is, the area does by far not represent the full range of possible types and thus the full typological diversity. An exception are the Tibetan languages (especially those which will be discussed below), which have very unusual and com-

plex systems of evidentiality that are without direct analogues in other areas.

The existence of evidential markers in the languages of the area sketched above has long been known in modern linguistics, and it is the profile of these languages that originally formed the point of departure for the „discovery” of the category of evidentiality in typology (next to the profile of the languages spoken in North America, see Section 1.4). Therefore, earlier descriptions of the grammatical system of the languages of the “Balkan type” usually did not treat evidentiality as an independent grammatical category, but analyzed evidential meanings as a special type of modality marking (or the marking of tense with a modal shade). Thus, the grammars of these languages referred to it by means of all-embracing categories of the type “non-obvious tenses”, “retell mood”, and so on. The motivation for such descriptions was based on the peculiar occurrence of the evidential markers in the languages of the Balkan: with respect to their expression they exhibit a close relation to markers of the perfect, regarding their content they display a close relation to the semantic domain of modality. This problem will be treated more in detail in Section 4.2.

Evidentiality in the languages of the “Great Belt” has been described in many studies (beginning with the monograph by Haarmann 1970). An important role for a comprehension of the nature of these markers in the languages of the Balkan played the studies by Friedman 1986, 2000, 2003 et al. For the Balkan languages and other languages of the area see also the articles in the edited volume by Guentchéva 1996, especially those about Bulgarian (Demina 1959, Aronson 1967, Guentchéva 1993, 1996a, and Nizolova 2006). For the languages of Asia and the Caucasus cf. the specific works by Lazard 1957 (one of the first theoretical studies of evidentiality), Friedman 1979, Slobin and Aksu 1982, Tatevosov 2001, the edited volumes by Johanson and Utas 2000, Aikhenvald and Dixon 2003, Guentchéva and Landaburu 2007, Xrakovskij 2007, and others. For the languages of the Volga area, see Leinonen and Vilks 2000.

B. Northern Siberia and Eurasia’s circumpolar zone. This includes, above all, languages of the Samodic group, which have a rather complex system of the grammatical expression of evidentiality, but also the Ob-Ugric languages, Yukagir and some others. For a discussion of the particularities of the evidential systems in the languages of this area see, in particular, Perrot 1996, Nikolaeva 1999, Künnap 2002, Maslova 2003, Burkova 2004 and especially the edited volume by Xrakovskij 2007. As one can see, an intense typological study of these systems has started only during the past years. Unfortunately, it is accompanied by a similarly intense loss of

these languages, which results in a destruction of grammatical systems and a complete or partial loss of the category of evidentiality with the younger generation of speakers (see Il'ina 2002 and Kazakevič 2005).

C. *The Baltic Region* includes, above all, Lithuanian and Latvian (and their dialects), which are Baltic languages of the Indo-European family, and Estonian, a Uralic language of the Balto-Finnic group. Evidentiality in the Baltic languages has, in all probability, a totally areal origin, although not all details of this process are entirely clear (see Koptjevskaja-Tamm and Wälchli 2001, Holvoet 2007: 81-105). Moreover, the system of the expression of evidentiality in Lithuanian exhibits a number of differences from that found in Latvian and Estonian, which are both more closely related to each other, and it generally seems to be somewhat more complex although it is, at present, also less grammaticalized than in the two other languages (for further details cf. Wiemer 2006, Wiemer 2007 and Holvoet 2007). Typologically, the systems of evidential marking found in these languages resemble the one used in the Balkan languages, both with respect to their form (use of participle forms of the verb, see Wälchli 2000) and content (the systems of evidentiality in the Baltic languages also belong to the “binary” type). However, there are also some important fine differences between the two (see Kehayov 2002, 2009, and Wiemer 2006).

D. The languages of *North America, Central America and South America*. The American continent is one of the richest areas with respect to the types and means used to express evidentiality. Furthermore, the most complex grammatical systems of the expression of evidentiality with verbs that we know at present are located in this area (as such one may consider the systems found in the languages of California, particularly in those of the Wintu and Pomo families). Systems of evidentiality of different types (all of them including, as a rule, a large number of grammatical oppositions) are found also in the Salish and Wakashan language family and in other languages of the North-Western coast of North America (including Eskimo languages) and in a number of other areas, such as in the languages of the Caddoan family, the Muskogean family, and others (Sherzer 1976, Mithun 1999: 181-186). In South America evidential marking is particularly widespread in the languages of the Amazon Basin (Aikhenvald and Dixon 1998), which are equally complex and diverse with respect to their structure (see the example from Tucano above). Evidentiality is also well attested in the neighboring Andean languages of Colombia, Ecuador, Peru and Bolivia, in various dialects of Quechua and Aymara, in the Barbacoan languages, and others.

One should not forget that the knowledge of evidentiality as a special category of the verb (the peculiarity of these languages seems to reflect cultures and ways of thinking unknown to the Europeans) acquired by the European linguists began with the study of the American Indians in the works of Boas and Sapir (although the languages of the Balkan type had, of course, been known earlier in Europe). Also the term “evidentiality” originated from studies of the languages spoken in this area. More precisely, the beginning of the typological study of evidentiality in the 1980s in general turned out to be linked with the languages of the American Indians basically. Therefore, the importance of the languages of the American continent for the study of evidentiality cannot be overestimated. Except for the studies of the languages of North and South America referred to above we would also like to mention the edited volume by Aikhenvald and Dixon 1999. Specific aspects of evidential systems in these languages are also discussed in many articles included in the widely known volume by Chafe and Nichols 1986, in the study by Willett 1988 and, most importantly, in the edited volumes by Aikhenvald and Dixon 2003 and Guentchéva and Landaburu 2007.

E. *Australia and New Guinea.* The languages spoken in this area have, without any doubt, markers of evidentiality, which exhibit different degrees of grammaticalization. However, the form of evidential marking is entirely different in the various regions. As it seems, the most complex of all the systems that are attested at present are located in New Guinea, especially in the languages of Southern Nagaur (Foe, Faso, and others; see Foley 1986: 166, Aikhenvald 2004: 62-63, 293). Next to these, there are also languages with simpler systems in New Guinea and in Australia, among them “binary” systems, and in many languages the grammatical expression of evidentiality is entirely absent. In the Australian languages in general evidentiality is weakly grammaticalized and usually expressed by means of a number of adverbial (or phrasal) units whose meaning often includes more than just a clearly identifiable element of evidentiality. The languages of New Guinea, on the other hand, are characterized by rarely occurring systems of the expression of evidentiality with untypical grammemes, a fact that requires a more detailed study.

The presented survey did not include all linguistic areas in which evidentiality can be found, but all those in which the grammatical expression of evidential values is most intensely used and/or in which it is highly grammaticalized. Areas that are comparatively poor with respect to the expression of evidentiality are, for instance, Oceania (where evidentiality is attested in some single Philippinean languages only), but also tropical

Africa, although the problem of the expression of evidential values in the verb system of the African languages requires further investigation. Probably, the suggestion that there is no evidentiality in the African languages (Aikhenvald 2004: 291) will have to be corrected in the future. At present, the grammatical expression of clearly definable evidential values can be found only in some few African languages (among them Dogon, spoken in Mali, which I studied some time ago and which is one of the central dialects that have perfective forms with clearly inferential components, see Plungian 1988). However, some of the grammatical strategies that are broadly spread across Africa and which are related to the phenomenon in question deserve much more attention. Among them are, above all, the highly frequent use of logophoric pronouns and the category of temporal distance that occurs in various forms in the verbal systems of many languages. Indeed, logophoric pronouns, which indicate whether the subject of the speech (i.e. the speaker) and the subject of the governing verb of saying in a proposition coincide or not, may, in some specific contexts, be used (similar to the German *Konjunktiv*) for the indication of whether a given fragment of a text is the reproduction of the words uttered by a third individual and, consequently, whether given information is known to the speaker through the words of some other speaker.

The close relation between logophoric markers and markers of evidentiality has been noticed in a number of typological studies (see e.g. Dimmendaal 2001 and Aikhenvald 2004: 132 ff.). Concerning the category of temporal distance, whose use (especially in discourse) has by far not been studied sufficiently, one may observe that in many African languages the verb system is characterized by a special marking of events that go back far in the past and which are not directly accessible to the perception of the speaker (for further details see Dahl 1985: 120-128, Aksenova 1997, Nurse 2008: 80-124). A relation to evidential values seems to be very likely in this case or at least deserves some more detailed investigation. For the important semantic potential of many forms that are traditionally labelled “discontinuous past tense” see also Plungian and van der Auwera 2006.

1.4. From the history of the study of evidentiality

As it has already been said, in the tradition of “western” linguistics the grammatical category of evidentiality was clearly recognized only during the second half of the twentieth century. The first typological studies of this category began to be conducted towards the end of the 1960s until the

beginning of the 1980s. A more intense study of evidentiality in various languages of the world began even later, towards the end of the 1980s. This clearly distinguishes evidentiality from other typologically important verbal categories, which were known in the European grammatical tradition much earlier, in some cases (such as with tense or mood) as early as in the antiquity.

Nevertheless, the history of the study of evidentiality is not that short that one could represent it in a nutshell. As it is often assumed, the first occurrence of evidential oppositions in forms of the past tense of Turkish verbs was noticed and analyzed already in the well-known “Collection of Turkish languages” (Arabic: *Dīwānu Luġāto t-Turki*) by Mahmud al-Kashghari in the eleventh century (see Friedman 2003: 189), with comments by Robert Dankoff, the English translator of the *Dīwānu* (Dankoff 1982). Another pioneer of the modern study of evidentiality was, undoubtedly, the French folklorist and poet Auguste Dozon, who at the end of the nineteenth century focused on special forms of the Albanian verb (see e.g. Dozon 1879). In order to describe them he introduced the term ‘admirative’ (French *admiratif*), which came to be widely used. Dozon took the emotional evaluation of an uttered fact as the basis of meaning (therefore the unusual internal form of the term, which indicates the meaning of “affection”), which is indeed clearly observable in the evidential forms of the Albanian verb, together with more usual inferential and reportative values that form one of the distinctive features of Albanian (for further details see Friedman 1986, 2000, 2003; Duchet and Pěrnaska 1996). Considering this shift of the semantic focus on modal meanings of evaluation, one may conclude that the observations by Dozon were highly important and had a strong impact on the terminology and the direction of subsequent investigations in the tradition of Balkan philology.

Apparently, these two names – Mahmud al-Kashghari and Auguste Dozon – also represent the two most important landmarks of the history of the study of evidentiality. However, the establishment of this category as a research object in modern linguistics begins, as it is usually claimed, with Franz Boas. Even the term “evidential” itself goes back to Boas,⁴ as well as the first exhaustive description of the functioning of this category in Kwa-kiutl (Wakashan family) and a number of other languages of the North-Western coast, which was carried out at the beginning of the twentieth century. In subsequent works Boas repeatedly focused on the role of evidential meanings in what today we could label “folk semantics” (or “naïve conceptual system”) characteristic of American Indian languages. A crucial element in Boas’ argumentation was the grammatical (that is, obligatory)

character of the expression of evidential values within the verb systems of many North American languages, a phenomenon that, from his point of view, formed a marked difference between American Indian languages and languages like English concerning the grammatical strategies used to create forms of the verb.

After Boas, evidential markers were discovered in and described for many languages of the American Indians, particularly in the works of Edward Sapir, Morris Swadesh and other specialists in American languages. The terminology was certainly not entirely uniform and fixed, and a conception of these markers as a special type of “mood” predominated (similarly, in earlier works in Bulgarian philology the phenomenon is often referred to as the “renarrative mood” of the Bulgarian verb). However, one may state that the specific grammatical properties of evidentiality in this period had been understood rather well, although the position of this category among other grammatical categories of the verb and its typological status in general were not subject to more detailed analyses. This was, however, a natural consequence of the fact that in the first half of the twentieth century a conception of a typology of grammatical categories in the truest sense of the word had not been established yet.

The turning point in the history of the typological study of evidentiality was the year 1957 when, at the same time and independently of each other, two works appeared, both determining the direction of subsequent studies for a long time. One of them was the seminal article by Jakobson (1957), in which he introduces the term “shifter” and, based on it, the first fundamental classification of verbal categories (with an illustration of their different properties on the basis of data from Russian). The other, not less important (though, perhaps, less known) work was an article by the French specialist in Iranian (and later one of the leading French typologists) Gilbert Lazard 1957, which was highly relevant for a conception of the problem. In this article the author studied the verbal system of Tadjik (including the expression of the category of evidentiality in this language, which he particularly focused on).

Both studies do not exclusively focus on evidentiality, but in each of them particular aspects of this category play an important role.

In Jakobson’s article evidentiality is one of the examples used to illustrate the peculiar type of verbal categories, which include in their meaning the joint reference to what he labelled “narrated event”, “speech event” and “narrated speech event”. In his own words, “[t]he speaker reports an event on the basis of someone else’s report (quotative, i.e. hearsay evidence), of a dream (revelative evidence), of a guess (presumptive evidence), or of his

own previous experience (memory evidence)” (Jakobson 1971: 135)⁵. Furthermore, Jakobson did not only establish the term “evidentiality” as one denoting that grammatical category which indicates the source of information of the speaker’s utterance, but also integrated the category into the broader typological context, establishing its place next to a number of other verbal categories. Thus (and probably for the first time), various phenomena found in typologically different languages could be described under one label, that of “evidentiality”. Generally speaking, one could say that it was now possible to find similarities, for example, between the verbal system of Bulgarian and that of Kwakiutl, both of which are constructed in an entirely different way otherwise. Thus, we owe it to Jakobson that a general label was found for phenomena that occurred in different languages, but which until that time had not been considered as similar in nature. One may regard this moment as the beginning of the typological study of evidentiality. Furthermore, it was in the works of Jakobson that evidentiality was first clearly distinguished from modality. This distinction had not been drawn in such a clear-cut way before, neither in the tradition of the Balkan linguistics nor in American linguistics.

The article by Lazard is entirely different in character. It is a detailed analysis of the verbal system of Tadjik, which has grammaticalized markers of evidentiality. However, the way evidentiality is expressed in Tadjik (which belongs to the zone of the “evidential belt” of Eurasia) differs fundamentally from the one observed by Boas and his successors in the languages of the North American Indians: it is a binary system (see below) in which evidential markers are highly polysemous. For the description of the semantic invariant of such markers Lazard suggested the term “meditative” (French *méditatif*), which is not entirely identical with Jakobson’s “evidentiality”: it denotes any form of “indirect” reflection of a situation, which is the case when the perception of situation is not based on direct, personal experience of the speakers or on their own conception of the world. During the following decades, Lazard’s suggestions were not immediately acknowledged in linguistic typology. However, through the accumulation of linguistic data and a deeper understanding of evidential systems in the languages of the world it became clear that Lazard’s intuitions were in many respects accurate, and the term “meditative” came to be widely used for the description of systems of the Balkan type, especially in the French linguistic tradition (see e.g. Guentchéva 1993, 1996, but also later works of Lazard himself: Lazard 1996, 1999 and 2000).

The publication of Jakobson’s works gave a new impulse to the study of evidentiality in various languages, and the period from the beginning of the

1960s to the middle of the 1980s was characterized by a continuous accumulation of new findings and the publication of a remarkable number of works. In these works, the authors intended to provide a deeper and more precise description of systems of evidentiality found in individual languages and to establish a more comprehensive typological conception of evidentiality, i.e. one that was not restricted to a particular tradition, but which accounted for all languages in the world that were known at this time. Some of the most important studies of this period are Aronson 1967 (elaborating on Jakobson's ideas of the Bulgarian system of the expression of evidentiality), Haarmann 1970 (one of the first monographs on evidentiality in the languages of the Great Evidential Belt), Friedman 1979 (analyzing data from the Caucasian languages), Hardman 1981 and Barnes 1984 (both providing a detailed analyses of data from South American languages), Slobin and Aksu 1982 (one of the first detailed descriptions of evidentiality in Turkish), and Givón 1982 (one of the first theoretical works accounting for the relation between evidentiality and modality).

The studies of this period laid the foundation of the publication of the famous edited volume by Chafe and Nichols 1986 and, closely related to it, the article by Willett 1988, in which the material from the edited volume was systemized and commented by the author. Until the present day, this edited volume is one of the most frequently cited books with respect to evidentiality. It does not only include a collection of very interesting and reliable data from languages of different linguistic areas (although the languages of North America predominate), but also proposes a reasoned classification of evidential values in the languages of the world (which was supplemented in Willett's article). Practically all subsequent studies were based on this classification in one way or the other, and many ideas suggested in the edited volume and in Willett's article were taken up later by other authors who wrote about the classification of evidential values (for further details see Section 2.2).

The following years were characterized by a rapidly growing interest in the category of evidentiality (which, strictly speaking, ceased to be "exotic" after the edited volume by Chafe and Nichols). It was a period in which more data were accumulated and in which more attention was given to the languages of the „Old World". Such a shift of the focus was only to be expected in view of the fact that these languages were given much less room in the edited volume by Chafe and Nichols. However, the universal classification of evidential values proposed in the volume could now be applied to the new material, both to new and to known data. It does not come as a surprise that in this period a number of edited volumes dealing

with evidentiality in the languages of the Old World were published, such as Guentchéva 1996 and its continuation Guentchéva and Landaburu 2007, Bo and Utas 2000, but also Dendale and Tasmowski 2001 (published as a special issue of the *Journal of Pragmatics*). Each of these works does not only provide a collection of new data, but also attempts to make a contribution to the theory of evidentiality and the classification of evidential values in general. These attempts often end up in more or less polemic statements with reference to Willett's classification, which had drawn largely on material from the languages of the New World. Especially important with respect to theoretical aspects are the re-introduction of Lazard's "meditative" (or "indirective" in the terminology of Johanson and others) as a cover term for the semantic cluster formed by evidential markers in binary systems, and also the more profound discussion of the relation between modality and evidentiality (we will return to this problem in Section 4.2). All these discussions are characterized also by a renewed interest in the conception of admirativity, which some authors exclude from the semantic domain of evidentiality, whereas others do include it (see the different points of view in De Lancey 1997, 2001, Lazard 1999 and Xrakovskij 2007a; for the modal component of evidential values see especially de Haan 1999 and 2005, Xrakovskij 2007a, but also van der Auwera and Plungian 1998).

The results of the study of evidentiality at this stage were analyzed in two recently published works of more general nature: the edited volume by Aikhenvald and Dixon 2003 and the monograph by Aikhenvald 2004. Currently, the latter can be regarded as the most comprehensive guide through the history of the study of evidentiality and the semantics of this category in the languages of the world. Further research on evidentiality will largely draw on this book. In the present overview we will repeatedly refer to Aikhenvald's views and also include those ideas into the discussion which appear debatable to us.

2. Classification of evidential values

2.1. The main types of opposition

As it was shown above, the generalization of existing data of the category of evidentiality suggests that the main opposition inherent in the various evidential values is based on the types of access to the information uttered by speakers, i.e. the means by which the speakers got to know about the

situation they speak about. This opposition can be of two kinds. On the one hand, one may distinguish between *direct* (firsthand, witnessed) and *indirect* access to information. Direct access refers to any means of obtaining information that presuppose a direct perception of a situation by the speakers (mainly visual, but not necessarily) and/or a direct participation of the speakers in a situation. The difference between these two types of direct access is needed because not all forms of direct participation in a situation necessarily involve its observation “from the outside”. Moreover, sometimes occurrences that are described from the “first person” perspective are generally not observable, for instance in the case of personal physiological sensations and emotions. These two types of direct access may be expressed by means of different grammatical constructions in different evidential systems, which will be discussed more in detail below. Accordingly, an indirect access to information presupposes that the knowledge about a situation was obtained through other means: the speaker did not observe the situation in a direct way and did not participate in it.

On the other hand, one may distinguish between a *personal* and a *non-personal* access to information. This opposition is not identical with the preceding one, although it seems to be very similar to it. A personal access to information presupposes that the speakers know about a situation on the basis of facts that they personally got to know, whereas a non-personal access to information presupposes that the speakers received their knowledge, roughly, through a report in somebody else’s words. Thus, the source of a retold assertion may be either known or unknown to the speakers. The non-personal access to information forms the basis of the semantics of one of the most widespread types of evidential markers, which, in accordance with current practice, is called ‘reportative’ (one may also find the terms ‘renarrative’, ‘quotative’, ‘hearsay’, and others; for more details on the typology of these markers, see especially Wiemer to appear).

For a significant number of situations the conception of direct/indirect and personal/non-personal access appears to be identical: a direct access is, for obvious reasons, always personal, and an indirect access is non-personal. (Note that a non-personal access is, per definition, also indirect.) However, these features may also be independent of each other.

Indeed, a direct (and personal) access presupposes that the speakers perceived the situation which they speak about themselves (in person) and in a direct way. A non-personal (and indirect) access presupposes that the speakers, talking about a situation, base their utterance on the information of another person. Another possible combination is the one of an indirect and personal access to the situation. In this case the speakers obtained

knowledge of a situation themselves, without other persons being involved, but the knowledge of this situation has not been obtained in a direct way since the speakers did not observe the situation directly. They may, for instance, either observe any results of a situation and take these as the basis for the conclusion that it took place or argue on the basis of some other data. In any case, the speakers talk about a situation which they did not observe personally and in which they did not take part, but about which they have some evidence that allows them to assume that it took or takes place. In (1) this type of access to a situation is reflected in (c): it may be uttered by speakers if, for instance, they do not see the fish on the table, but see how the dog guiltily has its tail between its legs. Such values are usually called *inferential* (or *inferentive*), since the speakers are required to draw a logical conclusion (an inference) on the basis of observed results. Another widespread type of meaning of a personal indirect access is the so-called *presumptive* value, which refers to cases where speakers produce an utterance in which they refer to a situation about which they do not know through concrete observed results, but through their knowledge about particular cause-and-effect relations: compare the context of the form *at this time of the year the berries should already be ripe*, where the utterance about the ripeness of the berries is not based on direct observation, but, so to speak, on the natural order of things. The difference between inferential and presumptive contexts may also be illustrated in the following example. One and the same utterance, such as *The neighbour is already at home*, may, in a language that has the respective set of grammatical means, be produced either by using presumptive or by using inferential markers of evidentiality. The first would be used for instance in the context of *eight o'clock - the neighbour should already be at home* (if the speakers know that this is time when the neighbour usually comes home). The second construction would be used in the context of *the light is on in the house - it seems that the neighbour is already at home* (in this case the speakers conclude from an observed situation, which they interpret as evidence enough for making the utterance). In this sense, the main difference between presumptive and inferential markers is that the first are based on the speakers' knowledge about the world and their capacity to draw logical conclusions, whereas the second are based on direct observation by the speakers (and, indirectly, on their capacity to draw logical conclusions from it). For more details on this problem, see also Tatevosov 2003.

From the two main oppositions characterized above the one between the direct and the indirect access proves to be hierarchically (and typologically) more important. The indirect access stands for a very broad range of

ways to get access to information, but excludes a direct synchronous perception of a situation or participation in it. One of these ways may be a logical conclusion from an observed result, or a reasoning based on general properties of the world, or a reproduction of what other people said. In many languages, this variety of types of indirect access is reflected in the existence of finer grammatical distinctions: here, specialized markers of the inferential, presumptive, reportative and other specific categories are possible. However, in the majority of languages a different strategy can be observed, namely the use of one single marker with a diffuse meaning for the expression of the indirect access to a situation. Its basic value includes only the indication of the fact that the speakers did not have direct access to a situation, but that they got informed about it anyhow, or perhaps things cleared up (if they did at all) through the context or through pragmatic factors. Exactly this uniform type of evidential markers, i.e. the one which indicates indirect access with a wide range of values (which are, however, conceptually alike), was the type discussed by Lazard 1957 under the label “meditative” (French *médiatif*), which was also called “indirective” in his studies on the Turkish languages, “non-confirmative” in many studies of the Balkan languages, and in many other ways. Aikhenvald 2004 uses the term “non-firsthand” for values that are very close to the one discussed above. Since languages with this type of evidential markers form, as already said, the majority of all languages spoken on earth, the number of terms that are used to capture their meaning and that arose in various linguistic traditions independently of each other is large as well. In the Russian linguistic tradition, the term *zaglaznost'* (referring literally to what happens “behind one’s back”) has been used to refer to this type of uniform markers that indicate indirect access (especially in the works of Caucasian linguistics, see, e.g., Kibrik 1977 for Archi). During the past few years also the descriptive term “indirect evidence” (*kosvennaja zasvidetel'stovannost'*) came to be widely used (see Kozinceva 1994 and 2007, Xrakovskij 2007a).

Markers of indirect access, which have a broad range of values in the languages of the world, have, as a rule, one property in common: they do not only express evidential values, but also different types of modal values. In other words, their inherent evidential meaning is, as a rule, not separated from the modal meaning, more precisely, from the meaning of epistemic assessment (i.e., the assessment of the degree of certainty of what is said). The pragmatic basis for this proximity is evident: speakers tend to evaluate information that they did not obtain as the result of a direct synchronous observation of a situation as less certain or less reliable. A more careful

formulation would probably not make use of the term “incertainty”, but rather speak of “epistemic distance”: the speakers refrain from taking over responsibility of the truth value of an utterance since the respective information did not enter into their personal sphere.

It should be noted that the relation between the values of indirect witness and of epistemic uncertainty (or rather epistemic “distance”) is generally not extremely close: there are evidential systems in which the fact that the speakers had no personal access to the respective information does not necessarily result in a lower degree of certainty of this information. A useful criterion for the evaluation of this relation is the type of markers used in a given language to formulate utterances of the type “general truths” (i.e. those which are unconditionally taken for granted in a given community) or utterances that refer to well-known facts (for example, facts related to the past of a given community), i.e. types of information which speakers cannot observe personally. In languages with “meditative” evidentiality, though not only in these, it often occurs that the use of markers of indirect access is omitted in such cases (although the speakers has no direct access to the described situation) in order to avoid an unnecessary epistemic load of these markers. An analysis of data according to these criteria can be found e.g. with reference to Bulgarian in Guentchéva 1996a and in Nicolova 2007; cf. also Friedman 2000 for other languages with such semantic effects.

However, in languages with more complex evidential systems markers of direct access often lack such epistemic load: the use of “indirect” evidential markers does not mean that the reliability of the description of a situation that has not been observed personally is in any way affected. Rather, these markers may be used also to describe situations known as true (this fact has, for example, been considered as characteristic for Tibetan by Tournadre 1996). Confer also Mithun 1999’s observation that in Central Yup’ik, in contrast to a number of other languages, markers of indirect access may render a verbal statement even more credible than markers of a direct access since their semantics does not include reference to subjective personal experience, but to more reliable collective experience (≈‘This is not what I myself thought through, this is what everyone knows.’). Various aspects of this type of marking are discussed also in Aikhenvald 2004, where the term “epistemic extensions” is used to refer to this aspect of the semantics of evidential markers. Thus, generally speaking, assertions referring to common knowledge may be expressed by means of entirely different strategies in languages of the world: in some rare cases specialized markers are used in this case (such as in the Tibetan and the Pomo lan-