DFG-Project (Research Grant)

Title: Functionality and Structure of Evidential Markings in Slavic
             (Integrative Theory with the construction of a database)

Project Director: Björn Wiemer
Research Assistant   Veronika Kampf <mark>(e-mail)</mark>

Start of Project: April 2010
Duration of Grant Support: 36 Months

Abstract
Evidentiality is a conceptual domain whose functions point to a cognitive or communicative foundation based on which the speaker feels entitled to make a statement; hearsay and inference count here in particular. The goal of the proposed project consists of building a database in which conventionalized units with evidential functions will be recorded according to unified semantic and structural criteria and, therewith, a typological comparison made accessible. The database should collate the majority of Slavic standard languages. Based on a distinct definition of the term evidentiality and on an onomasiologically oriented taxonomy, lexemes, morphemes and constructions will firstly be determined and systematically evaluated based on a criteria catalog and using corpora. Questionnaires worked out from this are addressed in one part to "naïve" mother tongue speakers, in another part to linguistically competent informants. This combination of approaches to the data strives to achieve the greatest reliability for the analysis of the entries. Simultaneously, the empirical basis for investigating the hypotheses of syntactic and pragmatic behavior as well as of the semantic structure of evidential markers will be created therewith. Furthermore, based on the Slavic materials, a model will be studied which will allow an expansion to other languages whereby comparability is achieved through methodological uniformity and transparency. The empirical results and the theoretic-methodological foundations of their investigation are recorded in a monograph.

About the State of the Research
'Evidentiality' is to be understood as a conceptual domain that refers to the cognitive or also communicative foundation (source) based on which the speaker feels entitled to make a statement P. The relative foundation can be divided onomasiologically into corresponding functions. A taxonomy or taxonomies resulting from this simultaneously create a starting point within linguistic systematics for registering methods with the help of which a speaker supplies an indication of what information sources or which basis of knowledge ('mode of knowing') he grounds assertion P on. Cf. Willett (1988: 56): "evidentiality is the linguistic means of indicating how the speaker obtained the information on which s/he bases an assertion."

The coarsest division into evidential sub-domains is between reportive and inferential functions, in other words between hearsay (cf. Germ. *Hans **soll** gestern verunglückt sein* 'Peter **is said to** have had an accident yesterday') and inference (cf. *Hier waren **offenbar** Menschen* 'People have **obviously** been here', expressed, for example, at the sight of a forest clearing devoid of people but with the remains of a campfire). This principle comparison depicts the basis for (partial) taxonomies of evidential functions, which have been worked on recently by Aikhenvald (2004); Pungian (2001), but also already by Willett (1988) based on typological data. A finer division within the area of reportive evidentiality dates back to Anderson (1986), cf. also, however, Aikhenvald (2004) and Palmer ([2]2001:40-42). Finer

divisions within the area of inferential evidentiality have been carried out recently by Squartini (2008) based on French and Italian. Astonishingly, these agree with the demonstrated semantic oppositions conducted previously by Jakovleva for Russian (1988; 1994: Chap. 3) between adverbs which allow inferences immediately based on perceptive information (e.g. Russ. *kažetsja*, *javno*) and adverbs which show that an inference is not possible through a perceptible occurrence but rather only through other cognitive backgrounds (e.g. knowledge of habits or a specific factual situation; cf. for instance Russ. *naverno*, *bezuslovno*).

Evidentiality is close to epistemic modality, however is to be principally separated from it; cf. the restrictions and definitions among other things in Aikhenvald (2004), Cornillie (2007), de Haan (1999a), Faller (2002), Kronning (2003), Plungian (2001), van der Auwera/Plungian (1998) as well as, and perhaps most succinctly in de Haan (2005:380): "Evidentiality *asserts* the evidence, while epistemic modality *evaluates* the evidence." (Emphasis by the author). Palmer also differentiated between 'judgments' (= epistemic in the true sense of the word) and 'evidentials' already in the first version of his standard work on modality. Indeed, he still dealt with 'evidentials' as a subclass of epistemic expressions, however, he was aware of the fact the term 'epistemic' was overstretched (Palmer 1986:51ff). In the revised version of his books, this insight resulted in a restructuring according to which epistemic and evidential meanings appeared separately although under the cover of 'propositional modality' (Palmer [2]2001: Chap. 2). Although terminologically different, Boye proceeds in a principally similar manner when he defines evidentiality "in terms of a concept of epistemic justification"; he tries to motivate epistemic and evidential functions together by considering them "in terms of general and socially-communicatively important universal cognitive structures" (Boye 2006: v). Meanwhile, an analytical separation of both these domains is imposed already from the nature of the involved functions: epistemic values are generally portrayed as scalable ('more-less' or values between '0' = 'S is certain, that non-P' and '1' = 'S is certain that P' for subjective (un)certainty which can be graduated in relation to one of either of these limit values) (cf. the overview in Krause 2007: Chap. 3) whilst evidential functions do not allow this very thing (what does a "more" or "less" mean to hearsay or inference?). A separation of both domains is, however, also necessary from a semasiological point of view because, in units of natural language, evidential and epistemic meaning components often (but not always) appear together, that is they form evidential-epistemic syncretisms. In this way, units can be subdivided whether they exhibit such syncretisms as arguably conventionalized meaning (e.g. Germ. *angeblich*, Russ. *jàkoby*, Pol. *rzekomo*, Lit. *esą*) or an epistemic limitation of believability of the information appears next to an evidential one only under certain discourse conditions whilst the relative units are, intrinsically, epistemically neutral (cf. Germ. *sollen*: *Der Antrag **soll** gestellt worden sein* 'The application **is said to** have been submitted', similar, among others, Pol. *podobno*, Lit. *girdì*, both 'one says', Eng. *apparently*, Fren. *paraît-il*, German subjunctive I). This is also equally valid for both grammatical and also lexical units; for lexical markers and auxiliaries, see 2.2, cf. also Cornillie (2007), de Haan (1999b; 2005) and the entries in Squartini (2007), for bound morphemes like, for instance, the Turkish *mIş* suffix, cf. various entries in Johanson/Utas (2000), in relation to the participle constructions based on the perfect tense cf., among others, Plungian (2001:345ff.) who cites the Balkan Slavic so called "renarrative" (Bulg. 'preizkazni formi') as an example of an epistemically modalized evidential system.

An initial conclusion can be made: not only theoretical, but also empirical reasons can be presented as to why evidential and epistemic functions cannot be reduced to one another. From the methodological point of view, it is imperatively advisable to only conduct a classification of relevant units from a semasiological viewpoint a f t e r a separation of both domains from an onomasiological viewpoint has occurred by means of clear definitions (see

above). It is exactly this, which should occur, in the proposed project using the example of the Slavic language family.

A second problem complex relevant to the proposed project arises through the classification of evidential markings on a continuum between lexicon and grammar. Not only grammaticalized, but also lexicalized units are conventionalized corner stones of respective languages. However, lexicalized units are not subject per se to any analytic access and are inventoried holistically whilst grammaticalized units assume analytically reducible complex units (constituents) in productive formation (cf. among others Lehmann 2002). Function words like adpositions, conjunctions and particles can, in this sense, be counted as lexical units. An area of transition between grammar and lexicon portrays not only auxiliary verbs, but also constructions with partially auxiliarized verbs such as *drohen (zu)* (cf. *Das Schiff drohte zu kentern* 'The ship is threating to capsize') which Diewald/Smirnova (in print), above all, described for German. In particular, function words without clear paradigm formation such as adpositions and conjunctions and also particles, parentheticals and adverbs are a reason for indecisiveness or changing decisions about whether they should be regarded as lexical (or lexicalized) or grammatical (or grammaticalized) units. Apart from the fact that this decision – insofar as it is justified based on criteria – is very dependent on theory, these types of function words were, until recently, practically not considered as worthy for the description. On the one hand, this tendency was reinforced by Aikhenvald (2004; 2007) who limits 'evidentiality' explicitly to the area of grammatical devices. She contests that lexical markers together with grammatical markers are able to be systematized as belonging to one conceptual domain and introduces the term 'information source' for this conceptual domain. On the other hand, Aikhenvald's point of view led to a counter reaction (e.g. in Cornillie 2007, Squartini 2007, in print, as well as works by the applicant, see below for personal preliminary work). A critical examination of demarcation problems of this kind is currently in full swing. Not sufficiently recognizable, however, is the separation of purely morphosyntactic aspects from functional ones. In this way, it is not clear why Aikhenvald (2007) counts particles, among others, indiscriminately as grammatical markers (thereby recognizing them as evidential resources) and in comparison and just as indiscriminately, not conjunctions (incl. complementizers), adpositions and also parentheticals (rejecting them then as evidential resources). In this case, the fact that particles and parentheticals often exhibit the same semantic-pragmatic functions and identical scope behavior, do not form any strict paradigms, and are often closely associated based on their diachronic development (cf. e.g. SEEM/APPEAR verbs and their complementizers of the AS IF type such as, for instance, Russ. *kažetsja* 'it appears' or *kak budto* 'as if') is overlooked. The arrangements carried out either implicitly or explicitly along a lexicon-grammar continuum prove at least, therefore, to be inconsequential. The attempted exclusion of lexical markings is exposed as being artificial insofar as some of the definitions of 'evidentiality' from Aikhenvald are absolutely able to be understood in a functional-onomasiological way (cf. for instance Aikhenvald 2003: 1) and are not limited to certain form classes or relations between synsemantic and autosemantic morphemes or lexicon units.

Another intermediate conclusion can be made here: until now, a clearly defined division of evidential markings according to morphosyntactic, distributional and lexicological criteria, which allow for an understandable differentiation on a lexicon-grammar continuum, has been missing.

With this desideratum, the following is correlated. A systematic investigation above all of function words with evidential functions – such devices, therefore, which fall under the above-mentioned onomasiological definition but which are not part of grammatical paradigms – has to date, to the best of the applicant's knowledge, not been successful for any language. For Slavic refer to the numerous publications of Kucarov (cf. among others Kucarov 1978; 1993 as well as in the overview from Wiemer 2008: 18). However, these collections of data

never got beyond the stage of a compilation without theoretical processing and almost only touch on translation comparisons between Slavic languages. Isolated approaches also exist for Polish and Czech. However, evidential adverbs, particles and conjunctions were always only named "in the wake" of epistemic modality markers and not treated as a species sui generic (cf. the overview in Wiemer 2006: 14-17 and 2008: 20-22).

Only in the last few years has goal-oriented research of not only the strictly grammatical devices of evidential markers begun; cf. for instance Cornillie (2007) on Spanish or the edited volume from Squartini (2007). Cornillie (2007) also further developed diagnostic processes for the differentiation of inferentive and reportive functions, which – among others – shall be applied in the proposed project. For a general overview of the research regarding Slavic languages cf. Wiemer (2008).

Studies and research complexes offer only relatively small overlapping areas with the nature of the tasks targeted in the project for discourse reproduction such as e.g. those which are introduced in Brendel et al. (2007). It is not only that a special (albeit also central) function area of evidentiality (namely the reportive) is handled in this case and thereby the connection to language-philosophical approaches is searched for, but rather, generally in this strand of research, it is primarily about general syntactic processes, which are connected with the other rules of a core grammar of the respective language (and their egocentric expressions, that is deictics and exclamatives). As in, for instance, Bucalić (2007) and prior to this in Plank (1986) but also in numerous investigations of indirect speech or Free Indirect Discourse (cf. among others Breslauer 1996, Kurt 1999, Socka 2004). The same applies, mutatis mutandis, for studies into syntax and semantics of epistemic or perception verbs (cf. for instance Whitt 2008) regarding an area that overlaps with markings of inferential evidentiality. The nature of the tasks of such investigations is almost complementary to the goal setting of the proposed project: In them, it is primarily about the description of syntactic systems of rules, in the proposed project, however, it is about registering the units from the lexicon and morpheme inventory of individual languages (which, for their part, can become the input of syntactic constructions) with the goal of a unified classification and retrievability of their structural and semantic-pragmatic characteristics as well as their diachronic backgrounds.

The typological as well as the individual language literature on the topic of 'evidentiality' has been constantly getting more comprehensive since the 1990s. In recent years, the boom has not only been shown in collected and special volumes (cf. the above mentioned edited volumes as well as, furthermore, Dendale/Tasmowski 2001, Aikhenvald/Dixon 2003, Guentchéva 1996, Guentchéva/Landaburu 2007, Xrakovskij 2007, Squartini 2007, Diewald/Smirnova, in process, and Wiemer/Plungjan 2008; sect. 2.2), but also in conferences (for instance workshops at the SLE yearly conference in Bremen 2006, the DGfS yearly conference in Bamberg 2008 and during GLOW in Utrecht 2008, furthermore numerous presentations during NRG4 in Leuven 2008) as well as in the form of special PhD Programs.

This overview allows the following conclusion. The term 'evidentiality' has already been known as such for at least 100 years, however, only in relation to North and South American languages as well as in relation to the languages of Central Asia and the Balkans since the 50s. Only since the middle of the 80s has the research into evidentiality gradually intensified and, for approximately 15 years now, has also been carried out in the European language area. The typological as well as individual language literature on this topic and its share in conferences has, since then, constantly become more comprehensive. Limping behind this development, however, is the formation of an integrative theory and the methodology of "making an inventory" of relevant markings.

The proposed project aims to lessen this exact deficit and the collection of the variations within and between individual languages. An integrative theory provides for regulated linking between an onomasiological starting point with an explicit, semasiological analysis in the units in question. It must cover not only grammatical, but also lexical marking (using function

words), furthermore, it must also enable a hierarchization and optimization of criteria comparable to the relevant states of knowledge and, moreover, be able to be extended in relation to further languages (see the diagram below).

It is furthermore regrettable that during the data collection and the verification of hypotheses on the meaning or the behavior of evidential markings, combining different data types and their examinations together has, to date, not been tried. Also to date, no questionnaires on evidential units have been created which are comparable with, for example, the Eurotyp project for the ascertainment of 'grams' (cf. Dahl 1985; 2000 and Lindstedt 2000). Questionnaire drafts for evidential markings, which, however, either primarily target grammatical devices or are limited to hearsay, merely exist in the series from the Petersburg Typology group (cf. Kozinceva 1994: 102f.; 2000: 239f.). In the proposed project, a combination of foundations of verification and the development of a comprehensive questionnaire are strived for (see below).

Cited Literature

Aikhenvald, A.Y. (2003): Evidentiality in typological perspective. In: Aikhenvald, A.Y., Dixon, R.M.W. (eds.): *Studies in Evidentiality*. Amsterdam, Philadelphia: Benjamins, 1-31.

— (2004): *Evidentiality*. Oxford etc.: Oxford U.P.

— (2007): Information source and evidentiality: what can we conclude? In: Squartini, M. (ed.).

Aikhenvald, A.Y., Dixon, R.M.W. (eds.) (2003): *Studies in Evidentiality*. Amsterdam, Philadelphia: Benjamins.

Anderson, Lloyd B. (1986): Evidentials, Paths of Change, and Mental Maps: Typologically Regular Asymmetries. In: Chafe, W., Nichols, J. (eds.): *Evidentiality: The Linguistic Coding of Epistemology*. Norwood, NJ: Ablex Publ., 273-312.

Boye, K. (2006): *Epistemic meaning: a cross-linguistic study*. Copenhagen (Ms. der unveröffenlichten Dissertation).

Brendel, E., Meibauer, J., Steinbach, M. (Eds.) (2007): *Zitat und Bedeutung* (= *Linguistische Berichte*, Sonderheft 15). Hamburg: Buske.

Breslauer, Chr. (1996): *Formen der Redewiedergabe im Deutschen und Italienischen*. Heidelberg: Groos.

Bucalić, T. (2007): Ein typologischer Beitrag zu Formen der Redewiedergabe. In: Brendel et al. (Eds.) …, 45-66.

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

Xrakovskij, V.S. (otv. red.) (2007): *Ėvidencial'nost' v jazykax Evropy i Azii*. Sankt-Peterburg: Nauka.

Cornillie, B. (2007): *Evidentiality and Epistemic Modality in Spanish (Semi-)Auxiliaries (A Cognitive-Functional Approach)*. Berlin, New York: Mouton de Gruyter.

Dahl, Ö. (1985): *Tense and Aspect Systems*. Oxford, New York: Basil Blackwell.

— (2000): The tense-aspect systems of European languages in a typological perspective. In: Dahl, Ö. (ed.): *Tense and Aspect in the Languages of Europe*. Berlin, New York: Mouton de Gruyter, 3-25.

de Haan, F. (1999a): Evidentiality and Epistemic Modality: Setting Boundaries. *Southwest Journal of Linguistics* 18, 83-101.

— (1999b): Evidentiality in Dutch. *Bulletin of the Chicago Linguistic Society* 25, 74-85.

— (2005): Encoding speaker perspective: Evidentials. In: Frajzyngier, Z., Hodges, A., Rood, D.S. (eds.): *Linguistic Diversity and Language Theories*. Amsterdam, Philadelphia: Benjamins, 379-397.

Dendale, P., Tasmowski, L. (eds.) (2001): *Evidentiality* (= *Journal of Pragmatics* 33, Special issue).

Diewald, G., Smirnova, E. (in print): The German evidential constructions and their origins: a corpus based analysis. In: de Mulder, W., Mortelmans, J., Mortelmans, T. (eds.): *Papers presented at the International conference on tense, aspect, mood, and modality (CHRONOS 7), 18-20 September 2006, University of Antwerp, Belgium.*

—— (eds.) (in preparation.): *Linguistic Realization of Evidentiality in European Languages.* Berlin, New York: Mouton de Gruyter.

Faller, M. (2002): *Semantics and Pragmatics of Evidentials in Cuzco Quechua.* Stanford, CA. (Ms. der unveröff. PhD)

Guentchéva, Zl. (ed.) (1996): *L'énonciation médiatisée.* Louvain, Paris: Peeters.

Guentchéva, Zl., Landaburu, J. (éds.) (2007): *L'énonciation médiatisée II,* Louvain etc.: Peeters.

Hansen, B. (2001): *Das slavische Modalauxiliar (Semantik und Grammatikalisierung im Russischen, Polnischen, Serbischen/Kroatischen und Altkirchenslavischen).* München: Sagner.

Jakovleva, E.S. (1988): Soglasovanie modusnyx xarakteristik v vyskazyvanii. In: Arutjunova, N.D. (red.): *Pragmatika i problemy intensional'nosti.* Moskva: AN SSSR, 278-302.

—— (1994): *Fragmenty russkoj jazykovoj kartiny mira (modeli prostranstva, vremeni i vosprijatija).* Moskva: Gnozis.

Johanson, L., Utas, B. (eds.) (2000): *Evidentials. Turkic, Iranian and Neighbouring Languages.* Berlin, New York: Mouton de Gruyter.

Kozinceva, N.A. (1994): Kategorija ėvidencial'nosti (problemy tipologičeskogo analiza). *Voprosy jazykoznanija* 1994-3, 92-104.

—— (2000): K voprosu o kategorii zasvidetel'stvovannosti v russkom jazyke: kosvennyj istočnik informacii. In: Bondarko, A.V., Šubik, S.A. (otv. red.): *Problemy funkcional'noj grammatiki: Kategorii morfologii i sintaksisa v vyskazyvanii.* Sankt-Peterburg: Nauka, 226-240.

Krause, M. (2007): *Epistemische Modalität (Zur Interaktion lexikalischer und prosodischer Marker).* Wiesbaden: Harrassowitz.

Kronning, H. (2003): Modalité et évidentialité. In: Birkelund, M., Boysen, G., Kjærsgaard, P.S. (éds.): *Aspects de la Modalité.* Tübingen: Niemeyer, 131-151. (Linguistische Arbeiten 469.)

—— (2007): Les auxiliaires „médiatifs" en suédois à la lumière du conditionnel épistémique et du futur aléthique en français. In: Begioni, L., Muller, Cl. (éds) (2007): *Problèmes de sémantique et de syntaxe. Hommage à André Rousseau.* Lille: Éd. du Conseil Scientifique de l'Université Charles-de-Gaulle (Lille 3) (= UL3 Travaux et recherches), 287-309.

Kucarov, I. (1978): Izrazjavane na preizkaznost v slavjanskite ezici črez vmjatane na modificirašči dumi, izrazi i izrečenija. *Godišnik na Sofijskija universitet. Fakultet po slavjanski filologii, Bălgarski ezik v săpostavka s drugi slavjanski ezici,* t. LXIX, 3, 83-116.

—— (1993): Izrazjavane na konkluzivnost v slavjanskite ezici. *Slavjanska filologija* 21. Sofija, 185-193.

Kurt, S. (1999): *Erlebte Rede aus linguistischer Sicht: Der Ausdruck von Temporalität im Französischen und Russischen (Ein Übersetzungsvergleich).* Bern etc.: Lang. (Slavica Helvetica 64.)

Lehmann, C. (2002): New reflections on grammaticalization and lexicalization. In: Wischer, I., Diewald, G. (eds.): *New reflections on grammaticalization.* Amsterdam, Philadelphia: Benjamins, 1-18.

Lindstedt, J. (2000): The perfect – aspectual, temporal and evidential. In: In: Dahl, Ö. (ed.): *Tense and Aspect in the Languages of Europe*. Berlin, New York: Mouton de Gruyter, 365-383.

Palmer, F.R. (1986): *Mood and modality*. Cambridge etc.: Cambridge U.P.

— (²2001): *Mood and Modality*. Cambridge etc.: Cambridge U.P.

Plank, F. (1986): Über den Personenwechsel und den anderer deiktischer Kategorien in der wiedergegebenen Rede. *Zeitschrift für germanistische Linguistik* 14, 284-308.

Plungian [= Plungjan], V.A. (2001): The place of evidentiality within the universal grammatical space. *Journal of Pragmatics* 33, 349-357.

Rakhilina, E.V. (1996): *Jakoby* comme procédé de médiatisation en russe. In: Guentchéva, Zl. (éd.): *L'énonciation médiatisée*. Louvain, Paris: Peeters, 299-304.

Ramat, P. (1996): "Allegedly, John is ill again": strategies pour le médiatif. In: Guentchéva (1996): Guentchéva, Zl. (éd.): *L'énonciation médiatisée*. Louvain, Paris: Peeters, 287-298.

Socka, A. (2004): *Sprachliche Merkmale der erlebten Rede im Deutschen und Polnischen*. Tübingen: Niemeyer. (LA 485.)

Squartini, M. (2001): The internal structure of evidentiality in Romance. *Studies in Language* 25-2, 297-334.

— (2004): Disentangling evidentiality and epistemic modality in Romance. *Lingua* 114, 873-895.

— (ed.) (2007): *Evidentiality between lexicon and grammar (= Rivista di Linguistica* 19-1).

— (2008): Lexical vs. grammatical evidentiality in French and Italian. *Linguistics* 46-5, 917-947.

van der Auwera, J., Plungian, V.A. (1998): On modality's semantic map. *Linguistic typology* 2-1, 79-124.

Weiss, D. (o.J.): *Semantyka konstrukcji 'mieć + bezokolicznik'. Próba rozstrzygnięcia polisemii*. Hamburg (Ms.).

Whitt, R.J. (2008): *Evidentiality and Perception Verbs in English and German: A Corpus-Based Analysis from the Early Modern Period to the Present*. Berkeley: Univ. of California (Ms. der unveröff. Dissertation)

Wiemer, B. (2006): Particles, parentheticals, conjunctions and prepositions as evidentiality markers in contemporary Polish (A first exploratory study). *Studies in Polish Linguistics* 3, 5-67.

— (2008): Lexikalische Markierungen evidenzieller Funktionen: zur Theoriebildung und empirischen Erforschung im Slavischen. In: Wiemer, B., Plungjan, V.A. (Hgg.): *Lexikalische Evidenzialitätsmarker im Slavischen*, 5-49. (= Wiener Slawistischer Almanach, Sonderband 72.)

Willett, Th. (1988): A cross-linguistic survey of the grammaticization of evidentiality. *Studies in Language* 12-1, 51-97.

Personal Preliminary Works

The previous works of the applicant on the topic of evidentiality were aimed, in essence, at lexical markings (function words). This is contingent on the interest in languages which, in the true sense of the words, possess no grammatical devices (bound morphology or TMA paradigms) to mark evidential functions like, for instance, Russian and Polish. In addition to this, certain works on Lithuanian, which deal with both grammatical and also lexical markers, also come from the applicant. The vast majority of works touch on the analysis of corpus data.

    The applicant's articles refer to the gaps in research presented above in the following way:

    1) In practically all the works (see the list below), the applicant points out the analytical differentiation between epistemic and evidential functions and the necessity for a

separation of onomasiological and semasiological viewpoints and uses them for recording the relative units. A more exactly intensional determination of the area of evidentiality takes place in particular in Wiemer (2006b, 2007a-b, 2008a, in print1).

2) In Wiemer (2006b: 20-51), it is reasoned, based on individual lexemes, that it is necessary to differentiate between 'epistemic reserve' (doubt and other similar things) and 'epistemic agnosticism' (that is, that the speaker refrains from making judgment about the believability or reliability of the source of information).

3) Criteria for a differentiation of units on a lexicon-grammar continuum are carried out explicitly in Wiemer (in print2) whereby the exemplary application is also carried out in other European languages in addition to Slavic.

4) The initial foundations for a single language description and systematization of evidential markings are laid out in Wiemer (2006b) for Polish, for lexical markers of Lithuanian in Wiemer (2007b). Additionally, lexicographical portraits (in the style of Wierzbicka) are created for the studied function words in Wiemer (2006b).

5) A general systematization of the Lithuanian inventory of evidential markings against a background of the newest typological work is carried out in Wiemer (2006a; 2007a). In Wiemer (2006a), it was particularly shown that the system of Lithuanian, which builds on participles, is tripartite (in terms of Aikhenvald 2004) and thus differentiates itself not only from Balkan Slavic, but also from its nearest relative, Latvian (Balkan Slavic and Latvian exhibit a bipartite system).

6) A review of the state of research on evidentiality in Slavic languages at all as well as an overview of the lexical evidentiality markers mentioned to date for Slavic languages were presented recently in Wiemer (2008a).

7) The case study Wiemer (2009) analyzes two evidential predicatives of Polish (*słychać*, *widać*) in relation to other predicatives and in the context of the diachronic syntax of Polish.

Furthermore, the applicant has accomplished further preliminary works in published works, which are necessary for the formulation and systematization of criteria in the database and which to some extent have been mentioned in the research overview above. In this context,theses were also developed which had up to that point not appeared in the research literature. In detail;

8) In Wiemer (2005), diachronic change patterns for common lexical markers from the epistemic range and from comparative constructions (e.g. Pol. *podobno*, Russ. *kak budto*, Lit. *lyg*, *tarsi*, all in their original meaning 'like, as though, similar') were compiled. In doing this, the following observations could be made:

(i) Only a minority of reportive markers in these three languages stem etymologically from speech verbs (e.g. Russ. *mol*, Pol. *rzekomo*). Given that, in the literature, the importance of illocutive verbs as source expressions for grammaticalization and lexicalization are often alluded to as reportive units, whether the Slavic-Baltic language area, in this regard, perhaps portrays a peculiarity, which is areally justified, remains to be investigated.

(ii) The majority of reportive units are derived diachronically from expressions which describe a comparison (e.g. Russ. *budto*, *jàkoby*, Pol. *jakòby*, *podobno*). A cursory overview of other European languages substantiates the impression that we are dealing here with a frequent pattern (at least within Europe) insofar as petrified verb forms of appearing should count here too (cf. For instance Fren. *paraît-il*, Ital. *a quanto pare*, Alem. *schints*, Eng. *apparently*, analog to Russ. *kažetsja*). At least in Slavic and Baltic, original complementizers often appear with reportive functions, which in the meantime are able to be used as particles (cf. Russ. *budto*, *jàkoby*, Pol. *jakòby*, Cze. *že*, Lit. *esą*, Lat. *it kā*).

(iii)   Commentaries in historical dictionaries of Polish permit the assumption that, at least in some cases, the expansion of an inferentially and epistemically used unit in the area of hearsay (reportive) occurs because speakers want to support their epistemic assessment (assumption, certainty or something similar) by calling on general knowledge. This is based then, from their perspective, on hearsay.

9) The diachronic development of the Lithuanian system for grammatical marking of evidential functions is followed by the applicant in Wiemer (1998; 2007a). In the last article, certain lexical devices (conjunctions, particles) are regarded diachronically.

10) The article from Wiemer (2008b) is directed at the development of meaning from inferentive to reportive function for Russian and Polish evidential markers. The attention is actually paid here to current occurrences, however, a diachronic background is also thoroughly investigated, which ties into the study by Wiemer (2005)

The communicative mechanism is uncovered that leads to units which, initially only used as an indication of perception-based inferences, then started also being used with reference to hearsay. This extended, corpus- based case study substantiates the hypothesis that a figure-ground relationship exists between the inferential and reportive function and that it can lead to a tipping effect between them, which is comparable with metonymic changes. Contrastingly, the communicative motivation for this change consists of epistemic distancing from the content of speech, which remains identical in its inferential and reportive use and which becomes salient in certain discourse conditions. Units arise in the results which are indifferent in this way as to whether they appear as the cognitive basis for claims of perception based inferences or hearsay (e.g. Russ. *kažetsja*); such units develop themselves one step further if they lose their inferential function and only indicate a hearsay function (e.g. Russ. *jàkoby* or Pol. *podobno*; cf. more exactly also Wiemer 2005, see above point 8-iii).

Whether the unit concerned additionally demonstrates a constantly epistemic component (like e.g. Russ. *jàkoby* 'apparently') or not (cf. for instance Pol. *podobno* 'one says') plays no role in this meaning impetus. Exactly here it is very important to demonstrate the difference between 'epistemic reserve' (doubt, among others) and 'epistemic agnosticism' (see above).

As an expansion from an inferential to a reportive function is not only limited to Slavic languages but rather is to be observed often in other European languages in both lexical and grammatical domains (cf. Ital. *a quanto pare*, Fren. *paraît-il*, Engl. *apparently*, Alem. *schints*, Dut. *schijnbaar*, Georg. *turme* as well as the Romance conditional of certain modal auxiliaries and also the Turk. *mIş*-suffix; cf. Giacalone Ramat/Topadze 2007, Squartini 2008 among others), results from Slavic research into the reconstruction of the semantic development and description of evidential marking can also achieve great use in other languages (cf. also Wiemer, in print2).

11) In Wiemer (2007b: 4.2), based on Lithuanian corpus data, it is ascertained that within combinations of evidential markings in the same expression, hidden epistemic meaning components can be intensified. A change in the evidential function, however, does not take place. This observation remains to be more exactly followed up on as it supports the essential supposition that evidential and epistemic functions demonstrate a principally different structure.

In addition to works that are published or are already in print, the applicant has also produced the following works:

Edited Volumes:

1. Wiemer (2008a) presents the introduction to a volume edited together with Vladimir Plungjan (Wiemer/Plungjan 2008). It is the first on this topic from the Slavic perspective and simultaneously one of the first at all dealing with lexical markers.
   - Abstract (Link)
2. As a follow up publication from a workshop that took place in April 2009 (see below), a thematic issue is planned for the journal STUF (edited together with Katerina Stathi, FU Berlin). The release of the manuscript for printing is scheduled for February 2010.
   - Abstract (Link)

The methodological approach to an integrative theory of evidential markings can be outlined as follows (cf. also the introduction to the aforementioned STUF-issue):

Towards an Integrative Theory of Evidentiality Markings
The integrative moment of a theory of evidential markings consists on the one hand in the fact that not only − as was the case until a few years ago − grammatical (that is paradigmatic and/or strongly morphologized) markings are primarily investigated, but also lexical units (function words). On the other hand, the integrative approach consists in a coupling of onomasiological and semasiological actions: the starting point is onomasiological (= conceptual) insofar as, under the guideline of a clear definition (= intensional regulation) of what 'evidentiality' is to be understood as (above all in distinction to epistemic modality), a taxonomy of evidential functions is to be estimated. Based on the already existing state of knowledge, this taxonomy typologically requires sufficient validity. The features taken into account there find their entry in the first part of a unified template. The semasiological aspect constitutes the description of the individual units which, per language, are to be understood as parts of the relative inventories of evidential markings and which should record both semantic-pragmatic and also structural characteristics of these markings in a unified manner and with explicitly formulated theoretical foundations. The verification of the description takes place finally through the questionnaires and questioning.

The connection between onomasiological and semasiological perspectives firstly requires their analytical separation. In its explicitness, this portrays a novelty insofar as an uncritical mixture of both viewpoints in the research to date has often led to confusion not only regarding the functional domain, but also the extension of the class of units. Simultaneously, the typological comparability of the Slavic materials is ensured through this approach without a "scaffolding" of functions or form classes or, in other words, units developing which, once set, does not later allow for any modifications as soon as new empirical findings (also from other language groups) make an expansion of the taxonomy and/or the inventory of markers seem appropriate. Rather, based on a representative selection of approx. 10 units per language, hypotheses are tested.

The integrative approach represented here demonstrates a similarity to the procedures of the Leningrad/St. Petersburg typology groups, in which, based on a manageable number of languages, a "calculus" of theoretical form-function relations is worked out for one domain ("phenomenon α"; see below), which will be used as onomasiological guidelines for a detailed and, above all, homogeneous representation of marking procedures which is adequate for the data of each individual language. Cf. here Nedjalkov/Litvinov (1995: 222): „We have in mind a reduction of language data and descriptions which is equally adequate for all the data and languages under investigation [$\rightarrow$ homogeneity; BW] and therefore can be accepted as an integral element of general typological knowledge [$\rightarrow$ expandable and comparable; BW]. (…) We speak of 'phenomenon α' in the sample of languages under investigation when we have achieved its definition and classification [$\rightarrow$ clear intensional guidelines with a resulting structure; BW]." The two most essential differences to the Petersburg Typology Group consist in the fact that (i) only languages from one language family will be studied,

however within the study then (ii) lexical units will also be integrated. As there is no reason to assume a priori that propositional markers, like the evidential ones, behave in a more unified manner than, for example, affixal morphology or functional extensions of verbal or nominal paradigms, the language genetic background most probably does not play a great role; rather, areal clusters allow the assumption that the Slavic languages are simply divided differently (Balkan language group in the (South)East, CBA in the North, areal clusters with Germanic and/or Romance languages in the (South)West).

Furthermore, for the integrative approach represented here, a certain model exists in Seiler's Cologne Universals Project (Unityp). Also firstly carried out in this project deductively were the subdivisions of domains in which certain mental basic operations were assumed; responses to the morphosyntax of real languages were only looked out for subsequently. However, the individual domains (functional dimensions) were not correlated and how the connection between 'indication' and 'predication' (in Seiler's terminological sense) for propositional operators like evidential markings should be used is not immediately obvious; furthermore, these are also difficult to conceptually separate from epistemic components in this model.

The questionnaire method is based on the one hand on translational equivalence, which, as far as possible, gets by without theoretical presuppositions e.g. about synonymy or polyfunctionality (cf. here Dahl 1985; 2000: 5f.; Koptjevskaja-Tamm 2008: 8-13 or also Berthele 2006) and on the other hand, it creates a standard for comparison between languages, which allows the meaning components to be established (or excluded) using as few contexts as possible through a comparable extension of reference objects (= situations or in other words, propositions which are modified by evidential markers).

Quoted Works

Personal Works

- Reviewed publications

Wiemer, B. (1998): Pragmatic inferences at the threshold to grammaticalization – The case of Lithuanian predicative participles and their functions. *Linguistica Baltica* 7, 229-243.

— (2005): Conceptual affinities and diachronic relationships between epistemic, inferential and quotative functions (preliminary observations on lexical markers in Russian, Polish and Lithuanian). In: Hansen, B., Karlík, P. (eds.): *Modality in Slavonic languages. New perspectives*. München: Sagner (= Slavolinguistica 6), 107-131.

— (2006a): Grammatical evidentiality in Lithuanian (a typological assessment). *Baltistica* 41-1, 33-49.

— (2006b): Particles, parentheticals, conjunctions and prepositions as evidentiality markers in contemporary Polish (A first exploratory study). *Studies in Polish Linguistics* 3, 5-67.

— (2007a): Kosvennaja zasvidetel'stvovannost' v litovskom jazyke. In: Xrakovskij, V.S. (otv. red.): *Ėvidencial'nost' v jazykax Evropy i Azii*. Sankt-Peterburg: Nauka, 197-240.

— (2007b): Lexical markers of evidentiality in Lithuanian. In: Squartini, M. (ed.): *Evidentiality between lexicon and grammar*. (= *Rivista di Linguistica* 19-1.)

— (In Print1): Evidenzialität aus kognitiver Sicht. In: Anstatt, T., Norman, B. (Hgg.): *Slavjanskie jazyki v kognitivnom aspekte – Die slavischen Sprachen im Licht der kognitiven Linguistik*. Wiesbaden: Harrassowitz. (Ms. von 25 Seiten)

— (In Print2): Hearsay in European languages: toward an integrative account of grammatical and lexical marking. In: Diewald, G., Smirnova, E. (eds.): *Linguistic Realization of Evidentiality in European Languages*. Berlin, New York: Mouton de Gruyter. (Ms. from 44 pages)

- Non reviewed publications

— (2008a): Lexikalische Markierungen evidenzieller Funktionen: zur Theoriebildung und empirischen Erforschung im Slavischen. In: Wiemer, B., Plungjan, V.A. (Hgg), 5-49.

— (2008b): Pokazateli s citativnoj i inferentivnoj funkcijami v russkom i pol'skom jazykax – kommunikativnye mexanizmy semantičeskogo sdviga. In: Wiemer, B., Plungjan, V.A. (Hgg.), 335-376.

— (2009): *Widać* und *słychać*: zum Schicksal zweier erstarrter Infinitive. In: Berger, T., Giger, M, Kurt, S., Mendoza, I. (Hgg.): *Von grammatischen Kategorien und sprachlichen Weltbildern – Die Slavia von der Sprachgeschichte bis zur Politsprache (Festschrift für Daniel Weiss zum 60. Geburtstag)*. München—Wien, 615-632. (= Wiener Slawistischer Almanach, Sonderband 73.)

Wiemer, B., Plungjan, V.A. (Hgg.) (2008): *Lexikalische Evidenzialitätsmarker im Slavischen*. (= Wiener Slawistischer Almanach, Sonderband 72.)

Works of other Authors

Aikhenvald, A.Y. (2004): *Evidentiality*. Oxford etc.: Oxford U.P.

Berthele, R. (2006): *Ort und Weg (Die sprachliche Raumreferenz in Varietäten des Deutschen, Rätoromanischen und Französischen)*. Berlin, New York: Mouton de Gruyter.

Dahl, Ö. (1985): *Tense and Aspect Systems*. Oxford, New York: Basil Blackwell.

— (2000): The tense-aspect systems of European languages in a typological perspective. In: Dahl, Ö. (ed.): *Tense and Aspect in the Languages of Europe*. Berlin, New York: Mouton de Gruyter, 3-25.

Giacalone Ramat, A., Topadze, M. (2007): The coding of evidentiality: a comparative look at Georgian and Italian. In: (ed.): *Evidentiality between lexicon and grammar* (= *Rivista di Linguistica* 19-1).

Koptjevskaja-Tamm, M. (2008): Approaching Lexical Typology. In: Vanhove, M. (ed.): *From Polysemy to Semantic Change. Towards a typology of lexical semantic associations*. Amsterdam/Philadelphia: Benjamins, 3-52.

Nedjalkov, V.P., Litvinov, V.P. (1995): The St Petersburg/Leningrad Typology Group. In: Shibatani, M., Bynon, Th. (eds.): *Approaches to Language Typology*. Oxford: Clarendon Press, 215-271.

Plungian, V.A. (2001): The place of evidentiality within the universal grammatical space. *Journal of Pragmatics* 33, 349-357.

Squartini, M. (2008): Lexical vs. grammatical evidentiality in French and Italian. *Linguistics* 46-5 , 917-947.

Goals

The project begins with the above-mentioned desiderata insofar as it is based on an integrative approach. This means that, in accordance with an exact onomasiological qualification of evidentiality and a taxonomical division of their sub-functions, a data base of evidential markers from 7-8 Slavic languages is formed in which, based on a criteria catalog which is to be worked through consistently and, when necessary, modified, information about semantics and range of function, pragmatic characteristics, morphological status and syntactic behavior, possible register limitations, lexical or categorial conditional limitations as well as, if possible, the etymological origin and diachronic development is recorded. The unified acquisition of data, identically defined criteria and their value allocations should guarantee the comparability of the entries also between languages. Units, which are examined, should be easily locatable using unified terminology and put together with other units according to certain criteria. This enables a better testing of the hypotheses of patterns for more than one language in behavior, structure and development of evidential markings, and should also help

in determining conclusions about implicative relationships between the individual criteria, which explain regularities in the development and functionality of evidential units.

In a heuristic process, the criteria and test procedures, firstly proven in the specialist literature and then extended in the preliminary work of the applicant, will be applied in order to, based on clear onomasiological guidelines, preliminarily examine the investigated units (above all, function words) based on criteria of the type named above and also conduct an initial inventory. Subsequently, first hypotheses (of a general nature and regarding individual units) will be examined based on corpora. Above all, however, questionnaires will be produced with which two groups of informants – one consisting of "naïve" mother tongue speakers and the other of linguistically competent speakers – from 7-8 Slavic languages will be questioned. The evaluation of the questionnaires flows into the final form of the data entries. A systematic explanation of the often implicit foundation, which is imperative for the description of evidential markings and for understanding the composition of the data entries, together with the empirical results and conclusions which result from the comparison of data entries for an integrative theory of evidential markers, go into a monograph.

The Slavic data base and its transparently explained theoretic-methodological premises can be understood as the starting point for the extension to other languages (or language groups) and the monograph as a type of handbook and signpost for such an extension. The integrative approach which the data acquisition and the monograph use as a basis comprises not only grammatical devices in the truest sense of the word, but also, to a greater degree, function words and syntagms with a conventionalized status whereby it is necessary to pay attention to the optimization and hierarchization of the relevant criteria.

Some Methodical Explanations About the Procedure

Evidential markers are discourse pragmatically easy to influence. Among other things, it is hardly noticeable for a great number of markings within their discursive contexts whether (i) in addition to an evidential function, they also demonstrate an epistemic function per se and (ii) they let themselves be limited to specific evidential functions. Additionally, for the objectification of conclusions, weaknesses of one data type (or data access) should be counter balanced by the strengths of another. For these reasons, the empirical research of evidential function markers can only lead to reliable results if different methods of verification are applied: (a) critical excerption from the existing secondary literature (based on clearly intentional, or in other words, defining specifications), (b) testing hypotheses with the help of corpus data and their variation with the help of consultants and informants, (c) surveying with the help of questionnaires. For this a difference must be made between mother tongue informants and linguists, who are available as consultants, and "naïve" mother tongue speakers who react spontaneously to the input of supporting evidence from the corpus and in correspondingly prepared questionnaires. A three-way division arises between (i) consultants (linguists) who are not mother tongue speakers but are specialists in the languages concerned, (ii) consultants who are simultaneously mother tongue speakers and are linguistically competent, (iii) mother tongue speakers who are not trained in linguistics, but with an "impartial" relationship to their language. In the project, all three types of respondents are called on. The questionnaires worked on in the project differentiate themselves from those used in the Eurotyp project and the Petersburg Typology Group (see above) in that they do not require translations from a middle language (English or similar languages) (such as in the Eurotyp project)or require them only partially and that not only linguistic consultants are questioned (as was normal within the Petersburg Typology Group).

As the description of evidential units (incl. those from a rather incidental lexicographical survey) to date hardly provides operative criteria with which their functional (semantic) range

can be tested, the project must deduce, formulate and test operative criteria. To date, a separation of stable meaning components from merely pragmatically associated components has often been missing. For this purpose, the operative differentiation of these component parts shall be tied in with the procedure dating back to Grice of differentiating asserted vs. implicit components. These are central to the test procedure with which the 'cancelability' and 'detachability' of the corresponding components can be ascertained. This procedure is particularly recommended for answering the question of whether, in addition to an evidential function, the given unit also demonstrates a stable (= not cancelable) epistemic component. Central to this is the difference already developed by the applicant between 'epistemic reserve' and 'epistemic agnosticism' (see above, preliminary works). This difference can also be illustrated in German: the auxiliary verb *sollen* has, among others, a reportive function; an often observed epistemic nuance according to which the speaker distances themselves with the help of *sollen* from the expressed assertion but which cannot be ascribed to the semantics of *sollen* themselves but rather to a contextual implicature, which can be easily canceled (or suppressed); cf. the continuation in (1a-b):

(1)   *Peter **soll** zehn Pfannkuchen in einem Rutsch gegessen haben.*
      'Peter **is said to** have eaten ten pancakes in one go.'
   (a)   *Und das glaube ich auch.*
         'And I believe it too.'
   (b)   *Frag mich aber nicht, ob das stimmt.*
         'But don't ask me whether or not that is correct.'

A speaker who uses *sollen* reportively portrays himself as 'epistemic agnostic' (in a formalized manner: 'I don't know if p', or '¬ {I know that p} ∧ ¬ {I know that ¬ p}'). The same applies to the subjunctive (e.g. in news broadcasts). It is different for the evidential adverb (or particle) *angeblich*: the speaker cannot step back from the doubt (its epistemic reserve compared with the communicated statement) simultaneously communicated with the use of this marker, which is why the continuations in (2a-b) sound inadequate:

(2)   *Peter hat **angeblich** zehn Pfannkuchen in einem Rutsch gegessen.*
      'Peter **apparently** ate 10 pancakes in one go.'
   (a)   ?*Und das glaube ich auch.*
         '? And I believe it too.'
   (b)   ?*Daran gibt es keinen Zweifel.*
         '? There is no doubt about it.'

The overwhelming majority of questioned mother tongue speakers confirm this assessment. However, it is not supported through an analysis of corpus data without restrictions.

Test procedures, which build on the incompatibility of a direct consequence expression with a test sentence are, however, also suitable to, for example, demonstrate whether units which go back to SEEM verbs or their complementizers (from the AS THOUGH type) are limited to inferential evidentiality and allow only a reportive reading qua implicature, develop themselves into indifferent (indifferential-reportive) units (like is obviously the case for Germ. *anscheinend* 'apparently'), or are even limited to hearsay (see above for these differentiations). Among others, Cornillie (2007) has developed a corresponding diagnostic based on Spanish. A further procedure is applied by Squartini (2008) for the differentiation on a finer taxonomical level, namely in order to differentiate between perception based ('circumstantial') and purely deductive inferences. Corresponding examples and questioning methods are included in the questionnaires.

About the corpora and the respondants

The Slavic languages examined in the project are Russian, Ukranian, Polish, Czech or Slovak, Bulgarian, Macedonian and Serbian/Croatian; Upper Sorbian can also potentially be included here. This choice is confined on the one hand to the areal distribution (the widest possible distribution) and the state of research, on the other hand by the possibility of accessing the primary data and the possibility of cooperation with colleagues (at least within countries where Slavic languages are spoken).

As an extension to the above-mentioned corpora, Google can also be used with necessary care. It must, however, be stressed that the actual check of the evidential functions of the preliminarily compiled markings per language is to be carried out using the help of consultants and informants; corpus evidence at best forms the starting point and can (in the case of missing pieces of evidence for the units) be completed by embedding them in predetermined contexts using questionnaires.

About the structure and handling of the data base

The relation between data sources, data basis, theoretical review and integrative theory are illustrated in the following schema:

Schema

Data Sources/ Verification:
1. Secondary literature
2. Dictionaries
3. Corpora
4. Consultants
5. Questionnaire Answers
    4a. from linguistically competent informants
    4b. from "naïve" informants (also in the Pre-Test, however with other people)

Database (FileMaker): / per language criteria (= fields)
1. Labels (allocation of values)
2.
…
n…

Σ Entries (Lemmata)
= Inventory of evidential markers in language X

Theoretical Review:
- Semantics / Pragmatics
- Morphological Status
- Syntactic behavior
…..
- Diachronic aspects

+ Question catalog (as a guideline for further research)

Integrative Theory (monograph), applied to the classification of Slavic evidential markers:

1. onomasiologically defined, semasiologically applied

2. integrates grammatical and lexical markers

3. hierarchizes and optimizes criteria on different levels

4. Suitable for expansion regarding further languages